Mismatch and Resolution in CS

Albert Fannjiang, UC Davis

Coworker: Wenjing Liao, UC Davis

SPIE Optics + Photonics, San Diego, August 21-25, 2011
Outline

• Mismatch: gridding error
• Band exclusion
• Local optimization
• Numerical results
• Comparison
• Conclusion
**Example: spectral estimation**

Noisy signal:

\[
y(t) = \sum_{j=1}^{s} c_j e^{-i2\pi \omega_j t} + n(t)
\]

where \( \omega_j \) are the frequencies, \( c_j \) are the amplitudes and \( n(t) \) is the external noise.

**Main problem: the frequencies.**

**Vectorization:** \( \Phi x + e = y \)

Set \( y = (y(t_k)) \in \mathbb{C}^N \) to be the data vector where \( t_k, k = 1, ..., N \) are the sample times in the unit interval \([0, 1]\).

\( \implies \) We can only hope to recover \( \omega_j \) are separated by at least 1 (resolution)
Approximate $\omega_j$ by the closest subset of cardinality $s$ of a regular grid $G = \{ p_1, \ldots, p_M \}, M \gg s$.

Write $x = (x_j) \in \mathbb{C}^M$ where $x_j = c_j$ whenever the grid points are the nearest grid points to the frequencies and zero otherwise.

The measurement matrix

$$W = \begin{bmatrix} a_1 & \cdots & a_M \end{bmatrix} \in \mathbb{C}^{N} \times M$$

with

$$W \in \mathbb{C}^{N} \times M, \quad \forall x \in \mathbb{C}^{N} \times M, \quad (W \cdot x = \phi)$$

Errors:

$$e = n + d, \quad n = \text{external noise}, \quad d = \text{gridding error}, \quad e = n + d$$
Gridding error is inversely proportional to refinement factor $F$

$$G = \frac{Z}{F}$$
Coherence pattern $\Phi^*\Phi$ for $100 \times 4000$ matrix with $F = 20$ (left).
Let $\eta > 0$. Define the $\eta$-coherence band of the index $k$ to be the set

$$B_\eta(k) = \{i \mid \mu(i, k) > \eta\},$$

and the $\eta$-coherence band of the index set $S$ to be the set

$$B_\eta(S) = \bigcup_{k \in S} B_\eta(k).$$

Due to the symmetry $\mu(i, k) = \mu(k, i)$, $i \in B_\eta(k)$ if and only if $k \in B_\eta(i)$.

Denote

$$B_\eta^{(2)}(k) \equiv B_\eta(B_\eta(k)) = \bigcup_{j \in B_\eta(k)} B_\eta(j)$$
$$B_\eta^{(2)}(S) \equiv B_\eta(B_\eta(S)) = \bigcup_{k \in S} B_\eta^{(2)}(k).$$
We make the following change to the matching step

\[ i_{\text{max}} = \arg \min_i | \langle r^{n-1}, a_i \rangle |, \quad i \notin B^{(2)}_{\eta}(S^{n-1}) \]

meaning that the double \( \eta \)-band of the estimated support in the previous iteration is avoided in the current search. This is natural if the sparsity pattern of the object is such that \( B_{\eta}(j), j \in \text{supp}(x) \) are pairwise disjoint.

**Algorithm 1. Band-Excluded Orthogonal Matching Pursuit (BOMP)**

Input: \( \Phi, y, \eta > 0 \)

Initialization: \( x^0 = 0, r^0 = y \) and \( S^0 = \emptyset \)

Iteration: For \( n = 1, \ldots, s \)

1) \( i_{\text{max}} = \arg \min_i | \langle r^{n-1}, a_i \rangle |, i \notin B^{(2)}_{\eta}(S^{n-1}) \)
2) \( S^n = S^{n-1} \cup \{i_{\text{max}}\} \)
3) \( x^n = \arg \min_z \| \Phi z - y \|_2 \) s.t. \( \text{supp}(z) \in S^n \)
4) \( r^n = y - \Phi x^n \)

Output: \( x^s \).
Theorem 1 Let $x$ be $s$-sparse. Let $\eta > 0$ be fixed. Suppose that

$$B_\eta(i) \cap B_\eta^{(2)}(j) = \emptyset, \quad \forall i, j \in \text{supp}(x)$$

and that

$$\eta(5s - 4) \frac{x_{\max}}{x_{\min}} + \frac{5\|e\|_2}{2x_{\min}} < 1$$

where

$$x_{\max} = \max_k |x_k|, \quad x_{\min} = \min_k |x_k|.$$ 

Let $\hat{x}$ be the BOMP reconstruction. Then $\text{supp}(\hat{x}) \subseteq B_\eta(\text{supp}(x))$ and moreover every nonzero component of $\hat{x}$ is in the $\eta$-coherence band of a unique nonzero component of $x$.

BOMP can resolve 3 RLs. Numerical experiments indicates resolution close to 1 RL when the dynamic range is close to 1 RL.
**Algorithm 2. Local Optimization (LO)**

**Input:** $\Phi, y, \eta > 0, S^0 = \{i_1, \ldots, i_k\}$.

**Iteration:** For $n = 1, 2, \ldots, k$.

1) $x^n = \arg\min_z \|\Phi z - y\|_2, \text{supp}(z) = (S^{n-1}\backslash \{i_n\}) \cup \{j_n\}$, for some $j_n \in B_\eta(\{i_n\})$.

2) $S^n = \text{supp}(x^n)$.

**Output:** $S^k$.

**Algorithm 3. BLOOMP**

**Input:** $\Phi, y, \eta > 0$

**Initialization:** $x^0 = 0, r^0 = y$ and $S^0 = \emptyset$

**Iteration:** For $n = 1, \ldots, s$

1) $i_{\max} = \arg\min_i |\langle r^{n-1}, a_i \rangle|, i \notin B_\eta^{(2)}(S^{n-1})$

2) $S^n = \text{LO}(S^{n-1} \cup \{i_{\max}\})$ where LO is the output of Algorithm 2.

3) $x^n = \arg\min_z \|\Phi z - y\|_2 \text{ s.t. supp}(z) \in S^n$

4) $r^n = y - \Phi x^n$

**Output:** $x^s$. 
**Theorem 2** Let $\eta > 0$ and let $x$ be a $s$-sparse well-separated vector. Let $S^0$ and $S^k$ be the input and output, respectively, of the LO algorithm.

If

$$x_{\min} > (\varepsilon + 2(s - 1)\eta) \left(\frac{1}{1 - \eta} + \sqrt{\frac{1}{(1 - \eta)^2} + \frac{1}{1 - \eta^2}}\right)$$

and each element of $S^0$ is in the $\eta$-coherence band of a unique nonzero component of $x$, then each element of $S^k$ remains in the $\eta$-coherence band of a unique nonzero component of $x$. 
Band-Excluded Thresholding (BET)

Two variants: Band-excluded Matched Thresholding (BMT) and Band-excluded Locally Optimized (BLOT).

Algorithm 4. BMT

Input: $\Phi, y, \eta > 0$.
Initialization: $S^0 = \emptyset$.
Iteration: For $k = 1, \ldots, s$,
1) $i_k = \arg \max_j |\langle y, a_j \rangle|, \forall j \notin B^2(\eta)(S^{k-1})$.
2) $S^k = S^{k-1} \cup \{i_k\}$
Output $\hat{x} = \arg \min_z \| \Phi z - y \|_2$ s.t. $\text{supp}(z) \subseteq S^s$

Algorithm 5. BLOT

Input: $x = (x_1, \ldots, x_M)$, $\Phi, y, \eta > 0$.
Initialization: $S^0 = \emptyset$.
Iteration: For $n = 1, 2, \ldots, s$.
1) $i_n = \arg \min_j |x_j|, j \notin B^2(\eta)(S^{n-1})$.
2) $S^n = S^{n-1} \cup \{i_n\}$.
Output: $\hat{x} = \arg \min \| \Phi z - y \|_2$, $\text{supp}(z) \subseteq \text{LO}(S^s)$. 
**BLO-based algorithms**

BLO Subspace Pursuit (BLOSP)

BLO Iterative Hard Thresholding (BLOIHT)

---

**Algorithm 6. BLOSP**

**Input:** $\Phi, y, \eta > 0$.

**Initialization:** $x^0 = 0, r^0 = y$

**Iteration:** For $n = 1, 2, ...$, 

1) $\tilde{S}^n = \text{supp}(x^{n-1}) \cup \text{supp}(\text{BMT}(r^{n-1}))$
2) $\tilde{x}^n = \arg\min_{z} \| \Phi z - y \|_2 \text{ s.t. } \text{supp}(z) \subseteq \tilde{S}^n$.
3) $S^n = \text{supp}(\text{BLOT}(\tilde{x}^n))$
4) $r^n = \min_{z} \| \Phi z - y \|_2, \text{supp}(z) \subseteq S^n$.
5) If $\|r^{n-1}\|_2 \leq \epsilon$ or $\|r^n\|_2 \geq \|r^{n-1}\|_2$, then quit and set $S = S^{n-1}$; otherwise continue iteration.

**Output:** $\hat{x} = \arg\min_{z} \| \Phi z - y \|_2 \text{ s.t. } \text{supp}(z) \subseteq S$. 

---
Algorithm 7. BLOIHT

Input: $\Phi, y, \eta > 0$.
Initialization: $\hat{x}^0 = 0, r^0 = y$.
Iteration: For $n = 1, 2, ...$
1) $x^n = \text{BLOT}(x^{n-1} + \Phi^*r^{n-1})$.
2) If $\|r^{n-1}\|_2 \leq \epsilon$ or $\|r^n\|_2 \geq \|r^{n-1}\|_2$
then quit and set $S = S^{n-1}$; otherwise continue iteration.
Output: $\hat{x}$.

In addition, the technique BLOT can be used to enhance the recovery capability with unresolved grids of the $L^1$-minimization principles, Basis Pursuit (BP)

$$\min_z \|z\|_1, \quad \text{subject to} \quad y = \Phi z.$$ 

and the Lasso

$$\min_z \frac{1}{2}\|y - \Phi z\|_2^2 + \lambda \sigma \|z\|_1,$$

where $\sigma$ is the standard deviation of the each noise component and $\lambda$ is the regularization parameter.
Numerical results

For two subsets \( A \) and \( B \) in \( \mathbb{R}^d \) of the same cardinality, the **Bottleneck distance** \( d_B(A, B) \) is defined as

\[
d_B(A, B) = \min_{f \in \mathcal{M}} \max_{a \in A} |a - f(a)|
\]

where \( \mathcal{M} \) is the collection of all one-to-one mappings from \( A \) to \( B \).

For dynamic range greater than 3, BOMP has the best performance.
LO dramatically improves the performance w.r.t. dynamic range
Duarte-Baraniuk 2010: Spectral Iterated Hard Thresholding (SIHT)

\[ y = \Phi x + e = \Phi \Psi \alpha + e \]

where \( \Phi \) is i.i.d. Gaussian matrix and \( \Psi \) is an oversampled, redundant DFT frame.

Assumption: \( \alpha \) is widely separated.

Performance metric:

\[
\frac{\| \Psi (\alpha - \hat{\alpha}) \|}{\| \Psi \alpha \|}
\]
Coherence bands of the DFT frame $\Psi$ (left) and $\Phi = \Phi \Psi$ (right).
Relative errors versus relative noise (left, dynamic range=1) and number of measurements (right, dynamic range=10)
Frame-adapted BP: synthesis approach

Candès et al 2010:

$$\min_z \|\Psi^* z\|_1, \quad \|\Phi z - y\|_2 \leq \varepsilon$$

Assumption: $\Psi^* z$ is sparse.

Analysis coefficients $\Psi^* z$ reorganized according to magnitudes.
Conclusion