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

\[
\Phi = \begin{bmatrix} a_1 & \ldots & a_M \end{bmatrix} \in \mathbb{C}^{N \times M}
\]

with

\[
a_j = \frac{1}{\sqrt{N}} \left( e^{-i2\pi t_k p_j} \right) \in \mathbb{C}^N, \quad j = 1, \ldots, M.
\]

Errors:

\[
e = n + d, \quad n = \text{external noise}, \quad d = \text{gridding error}.
\]
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).
Coherence band

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

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

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) \( i_{\text{max}} = \arg \min_i |\langle r^{n-1}, a_i \rangle|, i \notin B_\eta^{(2)}(S^{n-1}) \)
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: \( x^s \).
Two-dimensional case
**Performance guarantee**

**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.
**Local optimization**

### 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) $x^n = \arg \min_z \| \Phi z - y \|_2$ s.t. $\text{supp}(z) = (S^{n-1} \setminus \{i_n\}) \cup \{j_n\}$, 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) $x^n = \arg \min_z \| \Phi z - y \|_2$ s.t. $\text{supp}(z) \in S^n$

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

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

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 $x$. 
Figure 2: Reconstruction of the real part of 20 widely separated spikes \((R = 1, \text{ minimum distance } 3\rho)\) with \(F = 50, \epsilon = 5\%\) by (a) OMP (b) BLOOMP (c) BPDN (d) BPDN-BLOT.
Band-Excluded Thresholding (BET)

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

**Algorithm 4. BMT**

<table>
<thead>
<tr>
<th>Input: $\Phi, y, \eta &gt; 0$.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initialization: $S^0 = \emptyset$.</td>
</tr>
<tr>
<td>Iteration: For $k = 1, \ldots, s$,</td>
</tr>
<tr>
<td>1) $i_k = \arg \max_j</td>
</tr>
<tr>
<td>2) $S^k = S^{k-1} \cup {i_k}$.</td>
</tr>
<tr>
<td>Output $\hat{x} = \arg \min_z |\Phi z - y|_2$ s.t. $\text{supp}(z) \subseteq S^s$.</td>
</tr>
</tbody>
</table>

**Algorithm 5. BLOT**

<table>
<thead>
<tr>
<th>Input: $x = (x_1, \ldots, x_M), \Phi, y, \eta &gt; 0$.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initialization: $S^0 = \emptyset$.</td>
</tr>
<tr>
<td>Iteration: For $n = 1, 2, \ldots, s$.</td>
</tr>
<tr>
<td>1) $i_n = \arg \min_j</td>
</tr>
<tr>
<td>2) $S^n = S^{n-1} \cup {i_n}$.</td>
</tr>
</tbody>
</table>
| Output: $\hat{x} = \arg \min \|\Phi z - y\|_2$, $\text{supp}(z) \subseteq \text{LO}(S^s)$.
Corollary 1. Let \( \hat{x} \) be the output of BLOOMP. Under the assumptions of Theorems 1 and 2, \( \text{supp}(\hat{x}) \subseteq B_\eta(\text{supp}(x)) \), and, moreover, every nonzero component of \( \hat{x} \) is in the \( \eta \)-coherence band of a unique nonzero component of \( x \).

Even though we cannot improve the performance guarantee for BLOOMP, in practice the LO technique greatly enhances the success probability of recovery that BLOOMP has the best performance among all the algorithms tested with respect to noise stability and dynamic range (see section 5). In particular, the LO step greatly enhances the performance of BOMP with respect to dynamic range. Moreover, whenever Corollary 1 holds, for all practical purposes we have the residual bound for the BLOOMP reconstruction \( \hat{x} \)

\[
\|b - A\hat{x}\|_2 \leq c\|e\|_2, \quad c \sim 1.
\]

On the other hand, it is difficult to obtain bounds for the reconstruction error since \( \|x - \hat{x}\|_2 \) is not a meaningful error metric without exact recovery of an overwhelming majority of the object support.

4. Band-Excluded Thresholding (BET). The BE technique can be extended and applied to selecting \( s \) objects all at once in what is called the Band-Excluded Thresholding (BET).

We consider two forms of BET. The first is the Band-excluded Matched Thresholding (BMT), which is the band exclusion version of the One-Step Thresholding (OST) recently shown to possess CS capability under incoherence conditions [1].

For the purpose of comparison with BOMP, we give a performance guarantee for BMT under similar but weaker conditions than (15)–(16).


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

Theorem 3. Let \( x \) be \( s \)-sparse. Let \( \eta > 0 \) be fixed. Suppose that

\[
B_\eta(i) \cap B_\eta(j) = \emptyset \quad \forall i, j \in \text{supp}(x)
\]

and that

\[
\eta(2s - 1) \frac{x_{\max}}{x_{\min}} + \frac{2\|e\|_2}{x_{\min}} < 1,
\]

where \( x_{\max} = \max_k |x_k|, \quad x_{\min} = \min_k |x_k| \).

Let \( \hat{x} \) be the BMT 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 \( x \).
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 $n = 1, 2, \ldots$,  

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

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

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.
For the rest of simulations, we show the percentage of successes in 100 independent trials. A reconstruction is counted as a success if every reconstructed object is within $1\ell_R$ of the object support. This is equivalent to the criterion that the Bottleneck distance between the true support and the reconstructed support is less than $1\ell_R$. The result is shown in Figure 4. With 10 objects of dynamic range 5, BLOOMP requires the least number of measurements, followed by BOMP and then OMP, which does not achieve high success rate even with 100 measurements (left panel). With 100 measurements ($N = 100$) and 1% noise, BLOOMP can handle dynamic range up to 120 while BOMP and OMP can handle dynamic range about 5 and 1, respectively.

For the second example (11)-(12), we test, in addition to our algorithms, the method proposed by Duarte and Baraniuk$^5$ and the analysis approach of frame-adapted Basis Pursuit$^2,6$.

The algorithm, Spectral Iterative Hard Thresholding (SIHT)$^6$, assumes the model-based RIP which, in spirit, is equivalent to the assumption of well separated support in the synthesis coefficients and therefore resembles closely to our approach.

While SIHT is a synthesis method like BOMP and BLOOMP, the frame-adapted BP

$$
\min \|\Psi^*z\|_1 \quad \text{s.t.} \quad \|\Phi z - b\|_2 \leq \|e\|_2,
$$

is the analysis approach$^6$. Candès et al.$^2$ have established a performance guarantee for (19) provided that the measurement matrix $\Phi$ satisfies the frame-adapted RIP:

$$
(1 - \delta)\|\Psi z\|_2 \leq \|\Phi \Psi z\|_2 \leq (1 + \delta)\|\Psi z\|_2, \quad \|z\|_0 \leq 2s
$$

for a tight frame $\Psi$ and a sufficiently small $\delta$ and that the analysis coefficients $\Psi^*y$ are sparse or compressible.

Instead of the synthesis coefficients $x$, however, the quantities of interest are $y$. Accordingly we measure the performance by the relative error $\|\hat{y} - y\|_2/\|y\|_2$ averaged over 100 independent trials. In each trial, 10 randomly phased and located objects (i.e. $x$) of dynamic range 10 and i.i.d. Gaussian $\Phi$ are generated. We set $N = 100$, $R = 200$, $F = 20$ for test of noise stability and vary $N$ for test of measurement compression.

As shown in Figure 5, BLOOMP is the best performer in noise stability (left panel) and measurement compression (right panel). BLOOMP requires about 40 measurements to achieve nearly perfect reconstruction while the other methods require more than 200 measurements. Despite the powerful error bound established in [2], the analysis approach (19) needs more than 200 measurements for accurate recovery because the analysis coefficients $\Psi^*y$ are typically not sparse. Here redundancy $F = 20$ produces about $2F = 40$ highly coherent columns around each synthesis coefficient and hence $\Psi^*y$ has about 400 significant components.
LO dramatically improves the performance w.r.t. dynamic range
Spectral CS

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).
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.
Figure 5: Relative errors versus relative noise (left) and number of measurements (right, zero noise) for dynamic range 10.

In general, the sparsity of the analysis coefficients is at least $2sF$ where $s$ is the sparsity of the *widely separated synthesis* coefficients and $F$ is the redundancy. Thus according to the error bound of [2] the performance of the analysis approach (19) would degrade with the redundancy of the dictionary.

To understand the superior performance of BLOOMP in this set-up let us give an error bound using (18) and (20)

$$
\|\Psi(x - \hat{x})\|_2 \leq \frac{1}{1 - \delta} \|A(x - \hat{x})\|_2 \leq \frac{1}{1 - \delta} \|b - e - A\hat{x}\|_2 \leq \frac{1 + c}{1 - \delta} \|e\|_2
$$

(21)

where $\hat{x}$ is the output of BLOOMP. This implies that the reconstruction error of BLOOMP is essentially determined by the external noise, consistent with the left and right panels of Figure 5, and is independent of the dictionary redundancy if Corollary 1 holds. In comparison, the BOMP result appears to approach an asymptote of nonzero ($\sim 10\%$) error. This demonstrates the effect of local optimization technique in reducing error. The advantage of BLOOMP over BOMP, however, disappears in the presence of large external noise (left panel).

5. CONCLUSION

We have proposed algorithms, BOMP and BLOOMP, for sparse recovery with highly coherent, redundant sensing matrices and have established performance guarantee that is *redundancy independent*. These algorithms have a sparsity constraint and computational cost similar to OMP's. Our work is inspired by the redundancy-independent performance guarantee recently established for the MUSIC algorithm for array processing.7

Our algorithms are based on variants of OMP enhanced by two novel techniques: band exclusion and local optimization. We have extended these techniques to various CS algorithms, including Lasso, and performed systematic tests elsewhere8.

Numerical results demonstrate the superiority of BLO-based algorithms for reconstruction of sparse objects separated by above the Rayleigh threshold.

Acknowledgements. The research is partially supported in part by NSF Grant DMS 0908535.
Band-Excluded Thresholding (BET)

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

---

**Algorithm 4. BMT**

<table>
<thead>
<tr>
<th>Input: $\Phi, y, \eta &gt; 0$.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initialization: $S^0 = \emptyset$.</td>
</tr>
<tr>
<td>Iteration: For $k = 1, \ldots, s$,</td>
</tr>
<tr>
<td>1) $i_k = \arg \max_j</td>
</tr>
<tr>
<td>2) $S^k = S^{k-1} \cup {i_k}$</td>
</tr>
</tbody>
</table>

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

---

**Algorithm 5. BLOT**

<table>
<thead>
<tr>
<th>Input: $x = (x_1, \ldots, x_M), \Phi, y, \eta &gt; 0$.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initialization: $S^0 = \emptyset$.</td>
</tr>
<tr>
<td>Iteration: For $n = 1, 2, \ldots, s$.</td>
</tr>
<tr>
<td>1) $i_n = \arg \min_j</td>
</tr>
<tr>
<td>2) $S^n = S^{n-1} \cup {i_n}$.</td>
</tr>
</tbody>
</table>

Output: $\hat{x} = \arg \min \|\Phi z - y\|_2$, $\text{supp}(z) \subseteq \text{LO}(S^s)$. 

Algorithm 7. BLOIHT

Input: $\Phi, y, \eta > 0$.
Initlization: $\hat{x}^0 = 0, r^0 = y$.
Iteration: For $n = 1, 2, \ldots$,
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.
Figure 2: Reconstruction of the real part of 20 widely separated spikes ($R = 1$, minimum distance $3\rho$) with $F = 50, \epsilon = 5\%$ by (a) OMP (b) BLOOMP (c) BPDN (d) BPDN-BLOT.
Figure 3: Relative error with noise level $\epsilon = 1\%$ (top) 5\% (middle) and 10\% (bottom) and filter width $\delta = 0$ (left) and 0.05 (right).
Figure 1: Reconstruction of closely spaced spikes ($R = 3$, minimum distance 0.2$\rho$) with $F = 100$, $\epsilon = 5\%$ by (a) OMP, (b) BLOOMP, (c) BPDN, (d) BPDN-BLOT.