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:

$$y(t) = \sum_{j=1}^{s} c_j e^{-\mathrm{i}2\pi\omega_j t} + n(t)$$

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

Main problem: the frequencies.

Vectorization: $\Phi\mathbf{x} + \mathbf{e} = \mathbf{y}$

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

$\Longrightarrow$ We can only hope to recover ω_j are separated by at least 1 (resolution)

Approximate ω_j by the closest subset of cardinality s of a regular grid $\mathcal{G} = \{p_1, \ldots, p_M\}, M \gg s,$.

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

The measurement matrix

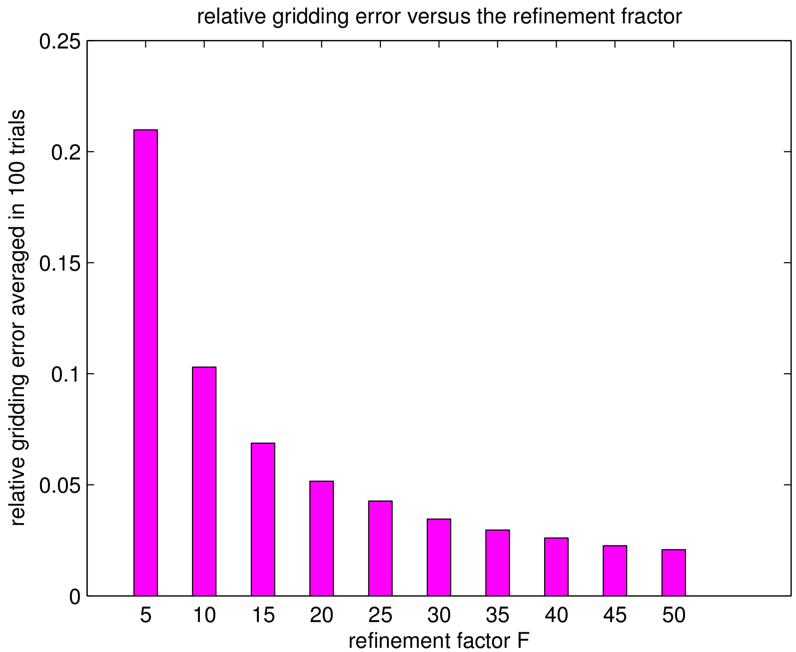
$$\mathbf{\Phi} = \begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_M \end{bmatrix} \in \mathbb{C}^{N \times M}$$

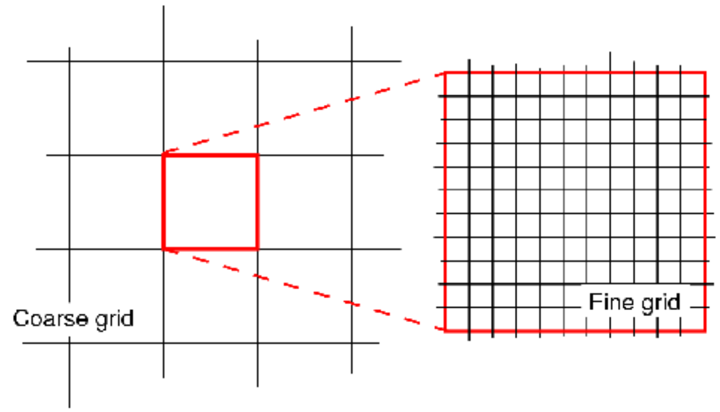
with

$$\mathbf{a}_j = \frac{1}{\sqrt{N}} \left(e^{-i2\pi t_k p_j} \right) \in \mathbb{C}^N, \quad j = 1, ..., M.$$

Errors:

$$\mathbf{e} = \mathbf{n} + \mathbf{d}, \quad \mathbf{n} = \text{external noise}, \quad \mathbf{d} = \text{gridding error}.$$

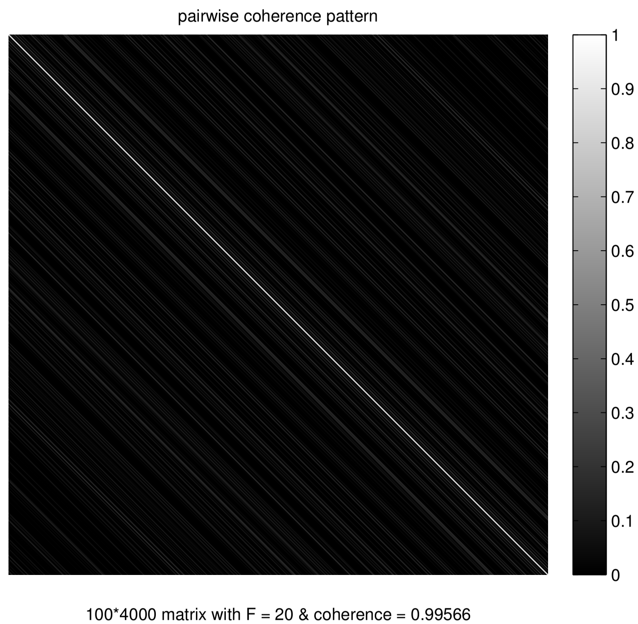


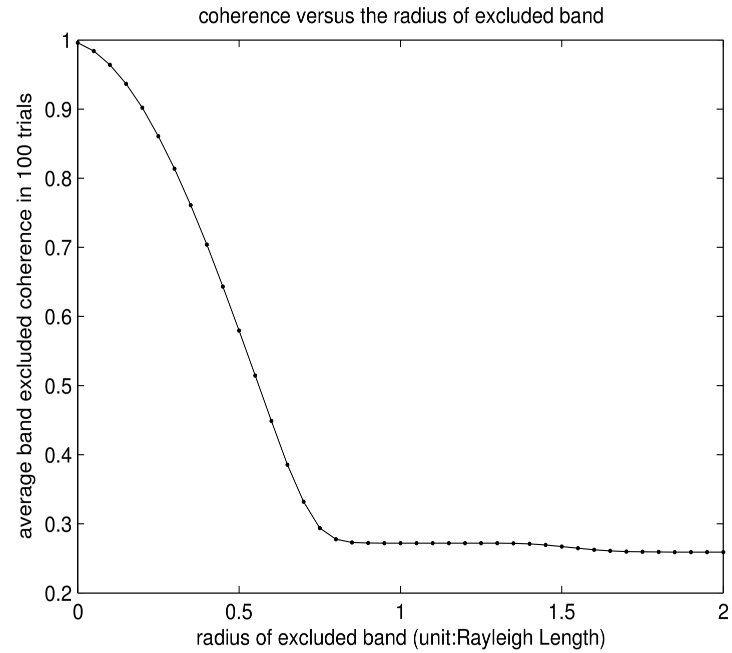


Gridding error is inversely proportional to refinement factor F

$$\mathcal{G} = \mathbb{Z}/F$$

.





Coherence pattern $\Phi^*\Phi$ for 100×4000 matrix with $F = 20$ (left).

Coherence band

Let $\eta > 0$. Define the η-coherence band of the index k to be the set

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

and the η-coherence band of the index set S to be the set

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

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

$$\begin{aligned} B_\eta^{(2)}(k) &\equiv B_\eta(B_\eta(k)) = \cup_{j \in B_\eta(k)} B_\eta(j) \\ B_\eta^{(2)}(S) &\equiv B_\eta(B_\eta(S)) = \cup_{k \in S} B_\eta^{(2)}(k). \end{aligned}$$

Band-excluded OMP

We make the following change to the matching step

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

meaning that the double η-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 $B_\eta(j), j \in \text{supp}(\mathbf{x})$ are pairwise disjoint.

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

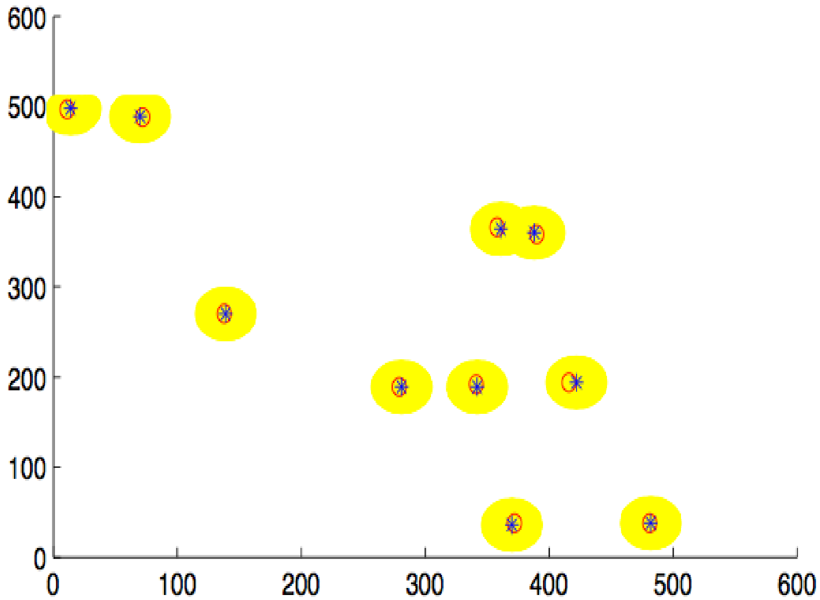
Input: $\mathbf{\Phi}, \mathbf{y}, \eta > 0$

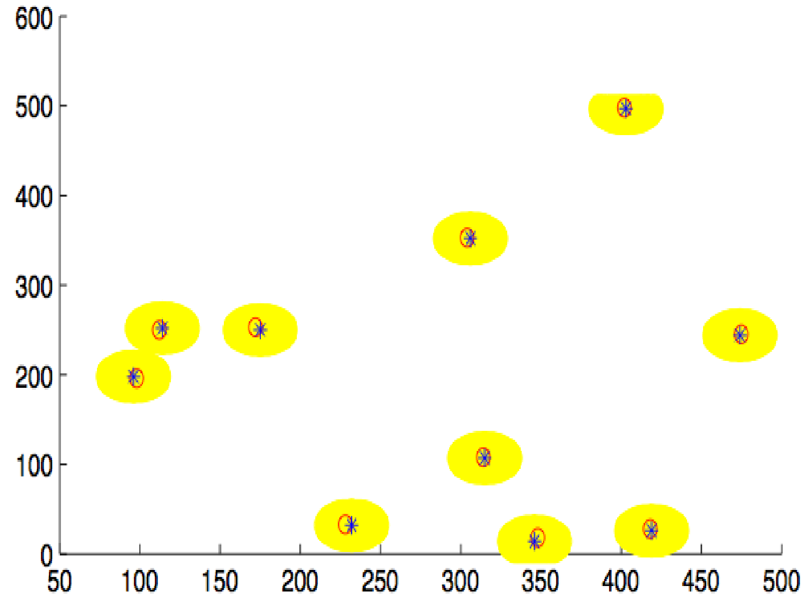
Initialization: $\mathbf{x}^0 = 0, \mathbf{r}^0 = \mathbf{y}$ and $S^0 = \emptyset$

Iteration: For $n = 1, ..., s$

1) $i_{\text{max}} = \arg\min_i |\langle \mathbf{r}^{n-1}, \mathbf{a}_i \rangle|, i \notin B_\eta^{(2)}(S^{n-1})$
2) $S^n = S^{n-1} \cup \{i_{\text{max}}\}$
3) $\mathbf{x}^n = \arg\min_{\mathbf{z}} \|\mathbf{\Phi}\mathbf{z} - \mathbf{y}\|_2$ s.t. $\text{supp}(\mathbf{z}) \in S^n$
4) $\mathbf{r}^n = \mathbf{y} - \mathbf{\Phi}\mathbf{x}^n$

Output: $\mathbf{x}^s$.





Two-dimensional case

Performance guarantee

Theorem 1 *Let $\mathbf{x}$ be s-sparse. Let $\eta > 0$ be fixed. Suppose that*

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

and that

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

where

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

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

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

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

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

Output: S^k.

Algorithm 3. BLOOMP

Input: $\mathbf{\Phi}, \mathbf{y}, \eta > 0$

Initialization: $\mathbf{x}^0 = 0, \mathbf{r}^0 = \mathbf{y}$ and $S^0 = \emptyset$

Iteration: For $n = 1, ..., s$

1) $i_{\mathrm{max}} = \arg\min_i |\left\langle \mathbf{r}^{n-1}, \mathbf{a}_i \right\rangle|, i \notin B_\eta^{(2)}(S^{n-1})$
2) $S^n = \mathrm{LO}(S^{n-1} \cup \{i_{\mathrm{max}}\})$ where LO is the output of Algorithm 2.
3) $\mathbf{x}^n = \arg\min_{\mathbf{z}} \|\mathbf{\Phi z} - \mathbf{y}\|_2$ s.t. $\mathrm{supp}(\mathbf{z}) \in S^n$
4) $\mathbf{r}^n = \mathbf{y} - \mathbf{\Phi x}^n$

Output: $\mathbf{x}^s$.

BLOOMP: performance guarantee

Theorem 2 *Let $\eta > 0$ and let $\mathbf{x}$ 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 η-coherence band of a unique nonzero component of $\mathbf{x}$, then each element of S^k remains in the η-coherence band of a unique nonzero component of $\mathbf{x}$.

Band-Excluded Thresholding (BET)

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

Algorithm 4. BMT

Input: $\mathbf{\Phi}, \mathbf{y}, \eta > 0$.
Initialization: $S^0 = \emptyset$.
Iteration: For $k = 1, ..., s$,
1) $i_k = \arg\max_j | \left\langle \mathbf{y}, \mathbf{a}_j \right\rangle |, \forall j \notin B_\eta^{(2)}(S^{k-1})$.
2) $S^k = S^{k-1} \cup \{i_k\}$

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

Algorithm 5. BLOT

Input: $\mathbf{x} = (x_1, \ldots, x_M)$, $\mathbf{\Phi}, \mathbf{y}, \eta > 0$.
Initialization: $S^0 = \emptyset$.
Iteration: For $n = 1, 2, ..., s$.
1) $i_n = \arg\ \min_j |x_j|, j \notin B_\eta^{(2)}(S^{n-1})$.
2) $S^n = S^{n-1} \cup \{i_n\}$.

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

BLO-based algorithms

BLO Subspace Pursuit (BLOSP)

BLO Iterative Hard Thresholding (BLOIHT)

Algorithm 6. BLOSP

Input: $\mathbf{\Phi}, \mathbf{y}, \eta > 0$.

Initialization: $\mathbf{x}^0 = 0, \mathbf{r}^0 = \mathbf{y}$

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

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

then quit and set $S = S^{n-1}$; otherwise continue iteration.

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

Algorithm 7. BLOIHT

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

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)

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

and the Lasso

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

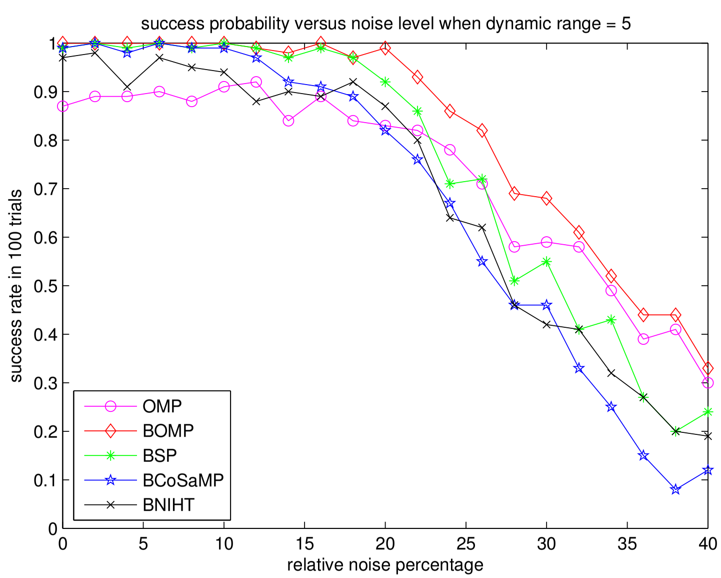
where σ is the standard deviation of the each noise component and λ is the regularization parameter.

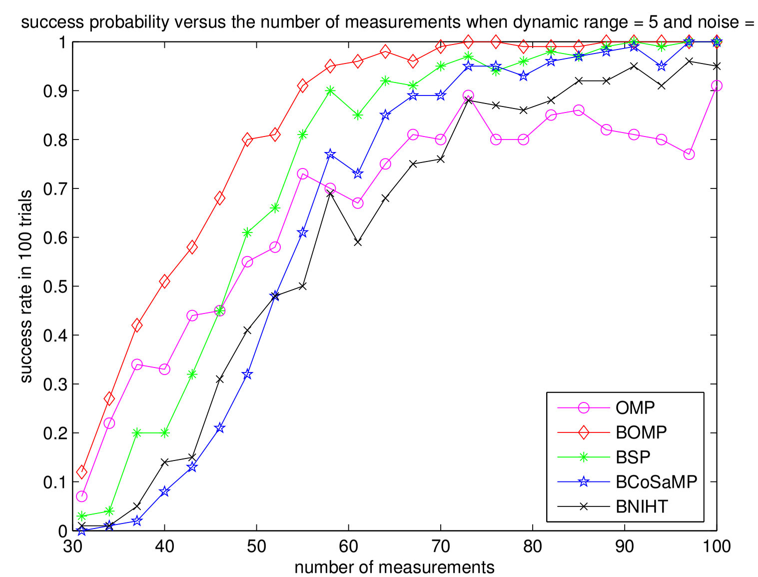
Numerical results

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

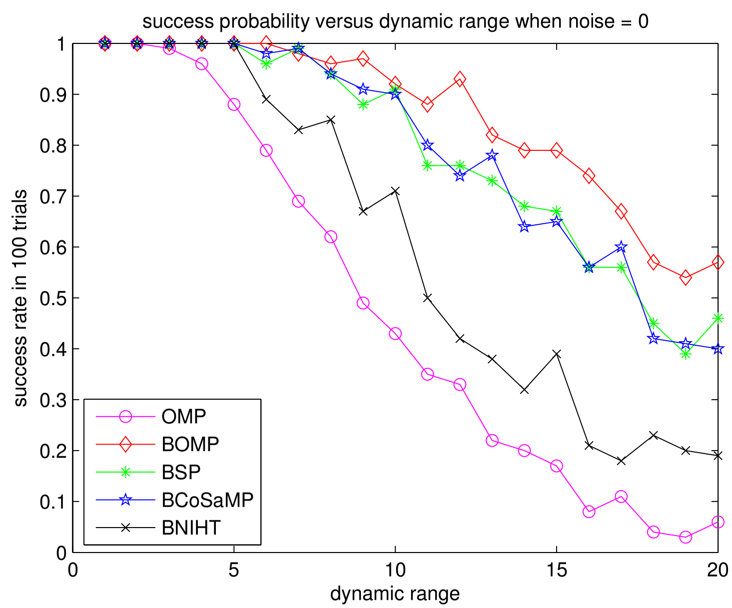
$$d_B(A,B) = \min_{f \in \mathcal{M}} \max_{a \in A} |a - f(a)|$$

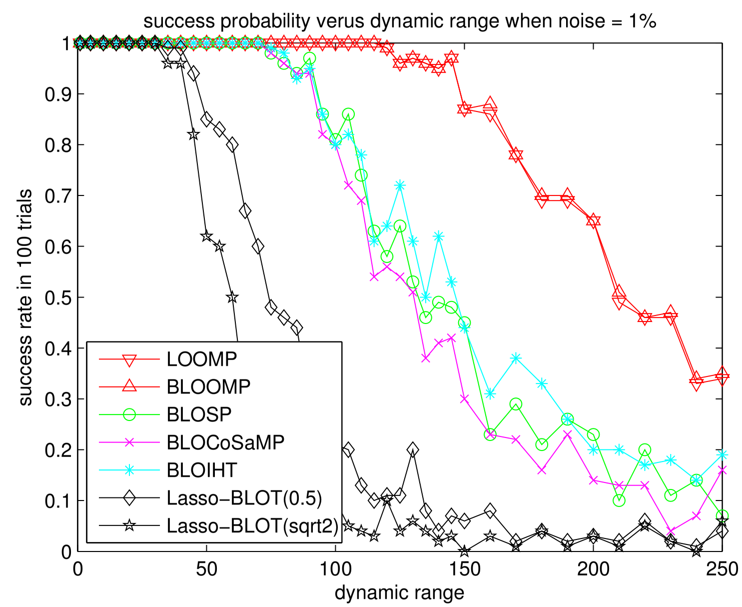
where $\mathcal{M}$ is the collection of all one-to-one mappings from A to B.





For dynamic range greater than 3, BOMP has the best performance.





LO dramatically improves the performance w.r.t. dynamic range

Spectral CS

Duarte-Baraniuk 2010: Spectral Iterated Hard Thresholding (SIHT)

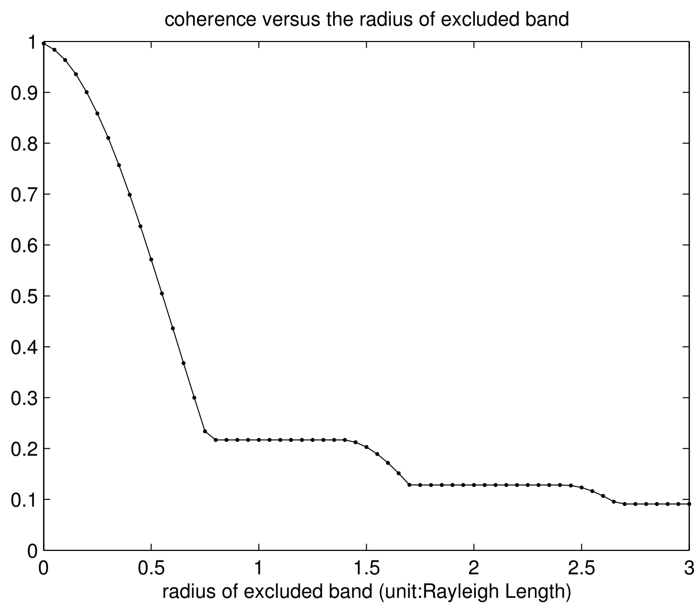
$$\mathbf{y} = \mathbf{\Phi}\mathbf{x} + \mathbf{e} = \mathbf{\Phi}\mathbf{\Psi}\alpha + \mathbf{e}$$

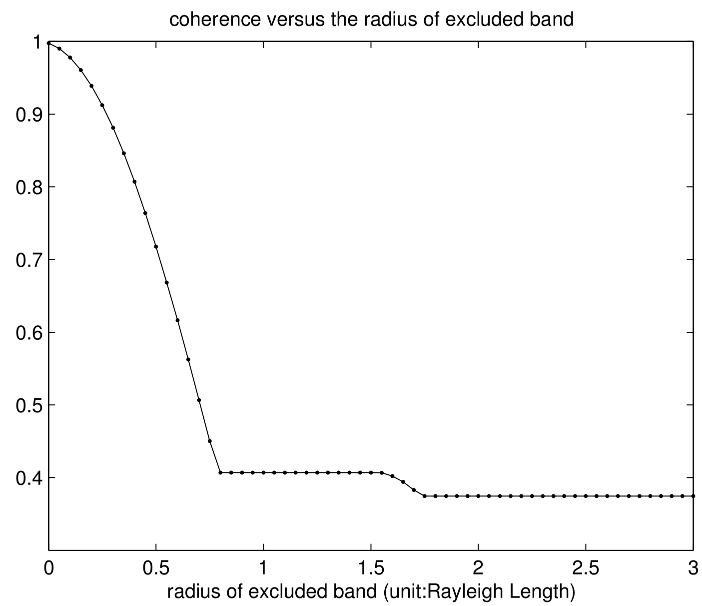
where $\mathbf{\Phi}$ is i.i.d. Gaussian matrix and $\mathbf{\Psi}$ is an oversampled, redundant DFT frame.

Assumption: α is widely separated.

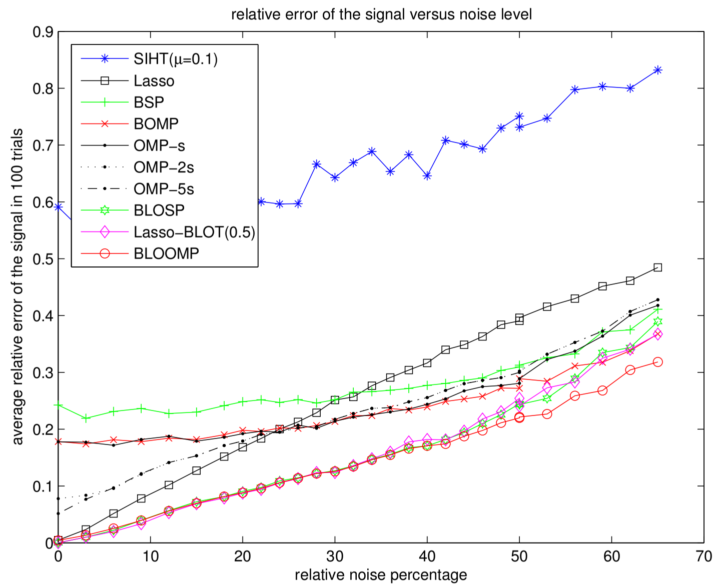
Performance metric:

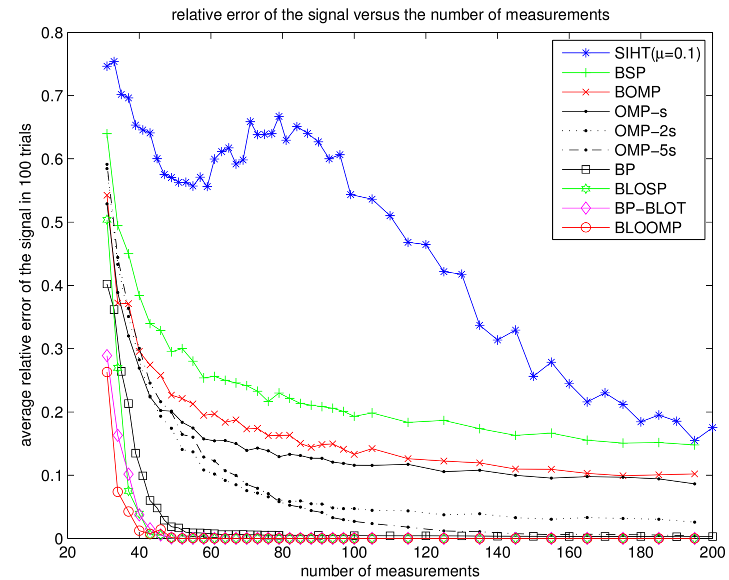
$$\frac{\|\mathbf{\Psi}(\alpha - \hat{\alpha})\|}{\|\mathbf{\Psi}\alpha\|}$$





Coherence bands of the DFT frame $\mathbf{\Psi}$ (left) and $\mathbf{\Phi} = \mathbf{\Phi}\mathbf{\Psi}$ (right).





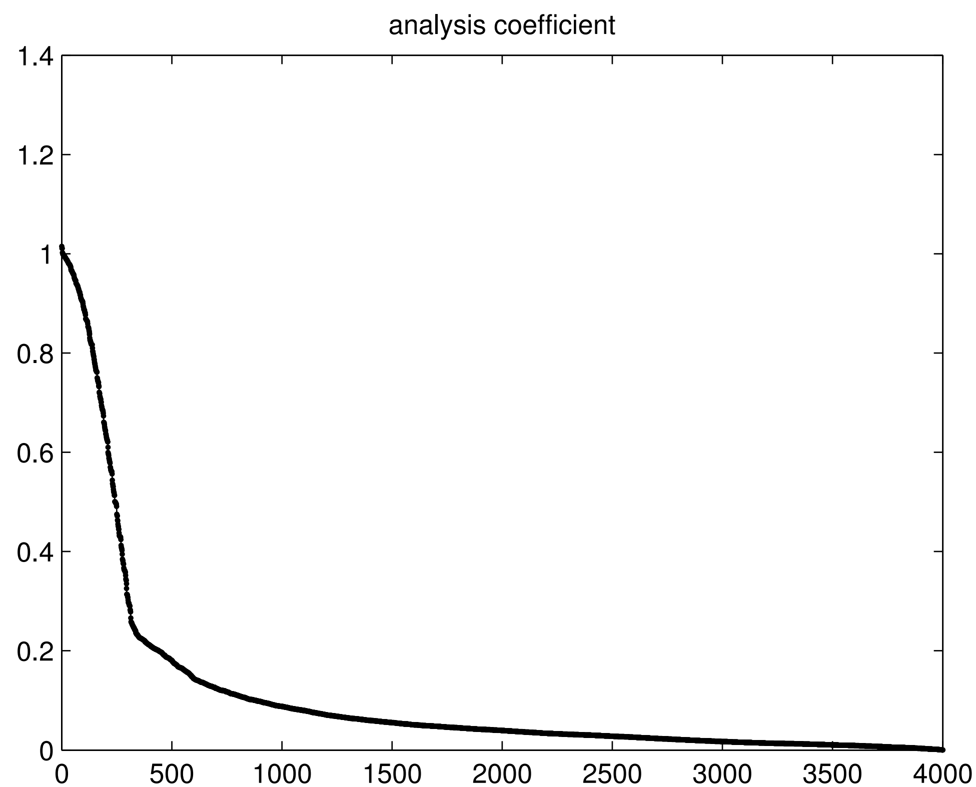
Relative errors versus relative noise (left, dynamic range=1) and number of measurements (right, dynamic range=10)

Frame-adapted BP: synthesis approach

Candes et al 2010:

$$\min_{\mathbf{z}} \|\mathbf{\Psi}^* \mathbf{z}\|_1, \quad \|\mathbf{\Phi} \mathbf{z} - \mathbf{y}\|_2 \leq \varepsilon$$

Assumption: $\mathbf{\Psi}^* \mathbf{z}$ is sparse.



Analysis coefficients $\mathbf{\Psi}^* \mathbf{z}$ reorganized according to magnitudes.

Conclusion