Greedy Sparse Signal Recovery

Heping Song

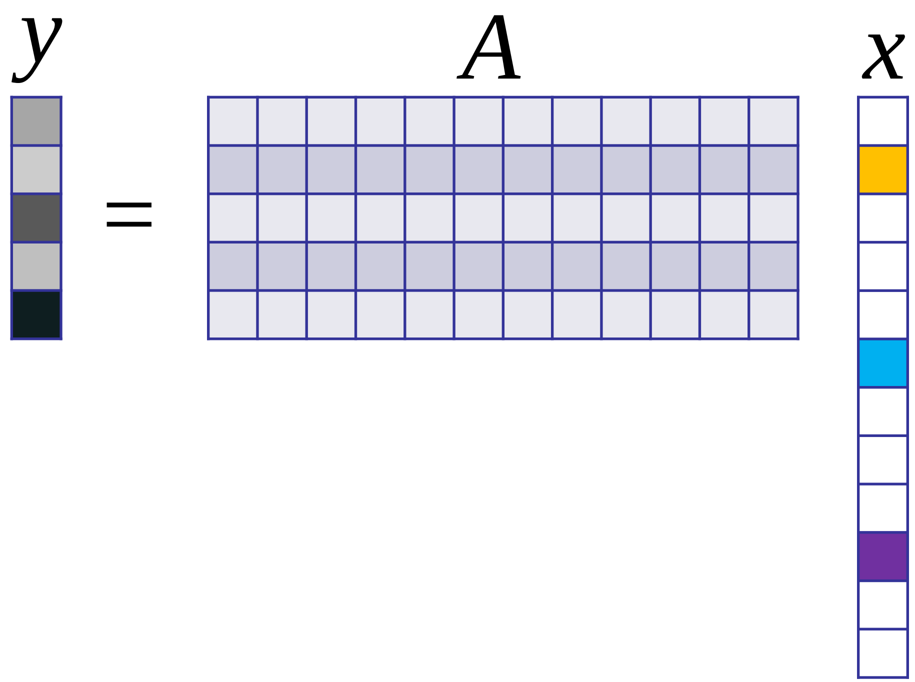
Outline

- Introduction
- RMP
- EMTP
- Demos
- Experiments
- Discussions

Notation

- $x^{(t)}$: the algorithms described in this paper are iterative and the reconstructed signal x in current iteration t is denoted as $x^{(t)}$. The same convention is used for other vectors and matries.
- I, A_I: index set I, the matrix A_I denotes the submatrix of A containing only those columns of A with indexes in I. The same convention is used for vectors.
- $[1, n] \setminus I$: the complement of set I in set $\{1, 2, \cdots, n\}$.
- $\mathrm{supp}(x)$: the support set of a vector x, i.e. the index set corresponding to the nonzeros of x, $\mathrm{supp}(x) = \{i : x_i \neq 0\}$.
- $H_k(x)$: the hard thresholding that sets all but the largest in magnitude k elements of a vector x to zero.
- $|x|$, $\|x\|_{\ell_p}$, x^T: the absolute value, ℓ_p norm and transpose of a vector x, respectively.
- $A^{\dagger}$: the Moore-Penrose pseudoinverse of matrix $A \in \mathbb{R}^{m\times n}$. $A^{\dagger} = A^T(AA^T)^{-1}$ for $m \leq n$; $A^{\dagger} = (A^TA)^{-1}A^T$ for $m \geq n$.

Problem Statement



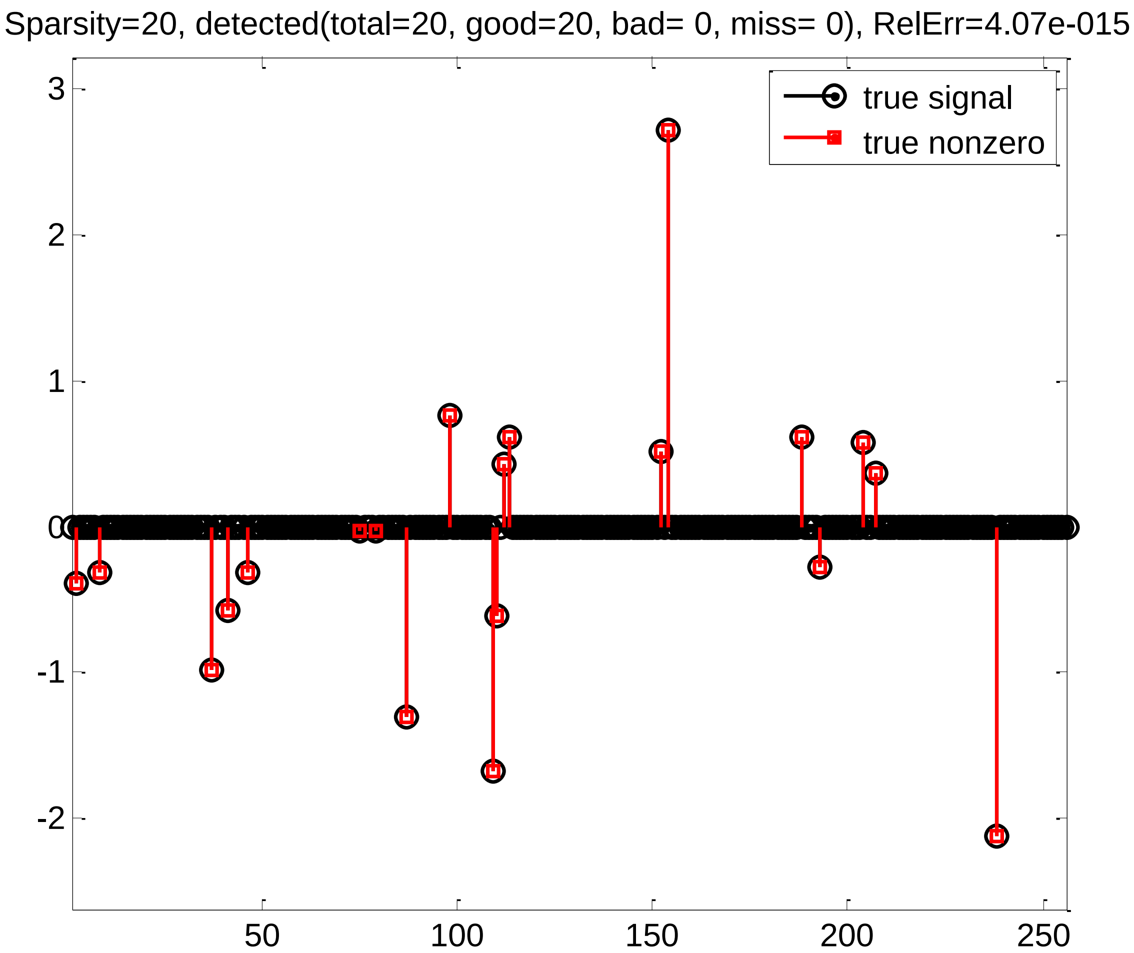
What
Why
How

y: m*1 measurement vector

A: m*n measurement matrix (m<<n)

x: n*1 unknown vector with k nonzeros

Sparse Signal Recovery (SSR)



n=256, *m*=60, *k*=20

Algorithms for SSR

$$(\mathrm{P}_{\ell_0}): \quad \min_{x_0} \|x_0\|_{\ell_0} \quad s.t. \quad Ax = y,$$

$$(\mathrm{BP}): \quad \min_{x_0} \|x_0\|_{\ell_1} \quad s.t. \quad Ax = y.$$

- **fewer measurements**
- **less computation**

Greedy Sparse Recovery

$$y = A_I x_I$$

A, y ⇒ { Support Detection ⇕ Signal Estimation } ⇒ x

$$x_I = A_I^{\dagger} y;$$

$$x_{\bar{I}} = 0$$

Support Detection Strategy

- Select atoms of measurement matrix A to generate y

- Determine active atoms in sparse representation of x

Outline

- Introduction
- RMP
- EMTP
- Demos
- Experiments
- Discussions

Orthogonal Pruning Pursuit (OPP)

- OPP derives a heuristic criterion from preserving a minimum residual by pruning a redundant basis successively
- Support detection strategy minimizes increase of the residual norm at each iteration
- OPP shrinks the support set ***I*** by pruning a basis $\boldsymbol{A_j}$

$$\arg\min_j \frac{(\hat{\mathrm{x}}_j)^2}{||A_j||_2^2}$$

$$\hat{\mathrm{x}} = A_I^{\dagger}\mathrm{y}$$

Residual Minimization Pursuit (RMP)

$$RMP: \quad \arg\max_j \frac{(\{A^T(AA^T)^{-1}r\}_j)^2}{\|A_j\|_2^2}, r = y - Ax_{t-1}$$

$$OMP: \quad \arg\max_j \left|A^T r\right|, r = y - Ax_{t-1}$$

$$1.\text{support detection}: \quad I_t = I_{t-1} \cup \{j\}$$

$$2.\text{signal estimation}: \quad x_I = A_I^{\dagger} y; \quad x_{\bar{I}} = 0$$

Extensions of RMP

- ***k*-RMP**: simultaneously select ***k*** atoms

$$I_t = I_{t-1} \cup \{ \text{indices of } k \text{ largest entries of } \left| A^{\dagger} r_{t-1} \right| \}$$

- ***β*-RMP**: support detection
  by thresholding

$$g = \left| A^{\dagger} r_{t-1} \right|$$

$$I_t = I_{t-1} \cup \{ i : g_i > \beta \max_i g_i \}$$

Why thresholding

- ***k*-RMP**: less iterations $<k$,
  better recovery

- ***β*-RMP**: less iterations $<k$,
  can work with unknown ***k***
  better recovery

Outline

- Introduction
- RMP
- EMTP
- Demos
- Experiments
- Discussions

One Stage Thresholding (OST)

- **ECME**: Expectation conditional maximization either

$$x_t = H_k(x_{t-1} + A^T(AA^T)^{-1} r_{t-1})$$

- **IHT**: Iterative hard thresholding

$$x_t = H_k(x_{t-1} + A^T r_{t-1})$$

slow convergence → accelerated version

Iterative Support Detection (ISD)

- **Refine the failed reconstructions by thresholding the solution of a truncated (BP) problem**

$$T_t = \{1, 2, \cdots, n\} \setminus I_{t-1}$$

$$x_t = \arg\min_x \|x_T\|_1 \ s.t. \ Ax = y$$

$$I_t = \{i : |x_i| > \beta^t \max |x_i|\}$$

$$\beta \in (0, 1)$$

ECME Thresholding Pursuits (EMTP)

- **Combine OST and TST**

1.support detection :

$$x_t = x_{t-1} + A^{\dagger} r_{t-1} \text{ or } x_t = x_{t-1} + A^T r_{t-1}$$

$$\text{strategy 1}: I_t = \text{supp}(H_k(x_t))$$

$$\text{strategy 2}: I_t = \{i : |x_i| > \beta^t \max|x_i|\}$$

$$\text{strategy 3}: I_t = I_{t-1} \cup \{i : |x_i| > \gamma \max|x_i|, i \notin I_{t-1}\}$$

2.signal estimation :

$$x_I = A_I^{\dagger} y; \quad x_{\bar{I}} = 0$$

RMP

Algorithm 1 Residual Minimization Pursuit

Input: Measurement matrix A, measurements y, sparsity level k
Output: The reconstructed signal x

```
Initialization:
t = 1 //iteration number
r_0 = y //initial residual
I_0 = ∅ //initial support set
for t = 1 : k do
    I_t = I_{t-1} ∪ { index of the largest entry of |A^† r_{t-1}| }
    x_{I_t} = A^†_{I_t} y
    x_{[1,n]\I_t} = 0
    r_t = y − A x_t
end for
return x
```

Initialization:
$t = 1$ //iteration number
$\mathrm{r}_0 = y$ //initial residual
$I_0 = \emptyset$ //initial support set
for $t = 1 : k$ **do**
  $I_t = I_{t-1} \cup \{$ index of the largest entry of $|A^{\dagger}\mathrm{r}_{t-1}|$ $\}$
  $\mathrm{x}_{I_t} = A^{\dagger}_{I_t}\mathrm{y}$
  $\mathrm{x}_{[1,n]\backslash I_t} = 0$
  $\mathrm{r}_t = \mathrm{y} - A\mathrm{x}_t$
end for
return x

k-RMP

Algorithm 2 k-Residual Minimization Pursuit

Input: Measurement matrix A, measurements y, sparsity level k

Output: The reconstructed signal x

Initialization:
$t \quad = 1$ //iteration number
$r^0 \quad = y$ //initial residual
$I^0 \quad = \emptyset$ //initial support set
while $halting\ criterion\ false$ **do**
$\quad I^t = I^{t-1} \cup \{\text{indices of } k \text{ largest entries of } |A^\dagger r^{t-1}|\ \}$
$\quad x^t_{I^t} = \arg\min_z \{\ \|y - A_{I^t} z\|_2,\ \text{supp}(z) \subseteq I^t\ \}$
$\quad x^t_{[1,n] \backslash I^t} = 0$
$\quad r^t = y - Ax^t$
$\quad t = t + 1$
end while
return x

β-RMP

Algorithm 3 β-Residual Minimization Pursuit

Input: Measurement matrix A, measurements y, threshold β

Output: The reconstructed signal x

Initialization:

$t \quad = 1$ //iteration number

$r^0 \quad = y$ //initial residual

$I^0 \quad = \emptyset$ //initial support set

while *halting criterion false* **do**

$\quad g^t = |A^\dagger r^{t-1}|$

$\quad I^t = I^{t-1} \cup \{i : g_i^t > \beta \max_i g_i^t\}$

$\quad x_{I^t}^t = \arg\min_z \{ \|y - A_{I^t} z\|_2, \operatorname{supp}(z) \subseteq I^t \}$

$\quad x_{[1,n]\setminus I^t}^t = 0$

$\quad r^t = y - Ax^t$

$\quad t = t + 1$

end while

return x

k-EMTP

Algorithm 4 k-EMTP Algorithm

Input: Measurement matrix A, measurements y, sparsity level k

Output: The reconstructed signal x

```
Initialization:
t = 1 //iteration number
x^(0) = 0 //initial signal
r^(0) = y //initial residual
I^(0) = ∅ //initial support set
while halting criterion false do
    x^(t) = H_k(x^(t-1) + A^† r^(t-1))
    I^(t) = {i : x_i^(t) ≠ 0}
    x^(t)_{I^(t)} = A^†_{I^(t)} y
    x^(t)_{[1,n]\I^(t)} = 0
    r^(t) = y − A x^(t)
    t = t + 1
end while
return x
```

β-EMTP

Algorithm 5 β-EMTP Algorithm

Input: Measurement matrix A, measurements y, thresholding parameter β

Output: The reconstructed signal x

Initialization:
$t = 1$ //iteration number
$x^{(0)} = 0$ //initial signal
$r^{(0)} = y$ //initial residual
$I^{(0)} = \emptyset$ //initial support set
while $halting\ criterion\ false$ **do**
  $x^{(t)} = x^{(t-1)} + A^{\dagger} r^{(t-1)}$
  $I^{(t)} = \{i : |x_i^{(t)}| > \beta^t \max_i |x_i^{(t)}|\}$
  $x_{I^{(t)}}^{(t)} = A_{I^{(t)}}^{\dagger} y$
  $x_{[1,n] \setminus I^{(t)}}^{(t)} = 0$
  $r^{(t)} = y - Ax^{(t)}$
  $t = t + 1$
end while
return x

γ-EMTP

Algorithm 6 γ-EMTP Algorithm

Input: Measurement matrix A, measurements y, thresholding parameter γ

Output: The reconstructed signal x

```
Initialization:
  t = 1 //iteration number
  x^(0) = 0 //initial signal
  r^(0) = y //initial residual
  I^(0) = ∅ //initial support set
while halting criterion false do
    x^(t) = x^(t-1) + A^† r^(t-1)
    I^(t) = I^(t-1) ∪ {i : |x_i^(t)| > γ max_{i∉I^(t-1)} |x_i^(t)|}
    x^(t)_{I^(t)} = A^†_{I^(t)} y
    x^(t)_{[1,n]\I^(t)} = 0
    r^(t) = y - A x^(t)
    t = t + 1
end while
return x
```

Initialization:

$t = 1$ //iteration number

$x^{(0)} = 0$ //initial signal

$r^{(0)} = y$ //initial residual

$I^{(0)} = \emptyset$ //initial support set

while *halting criterion false* **do**

$x^{(t)} = x^{(t-1)} + A^{\dagger} r^{(t-1)}$

$I^{(t)} = I^{(t-1)} \cup \{i : |x_i^{(t)}| > \gamma \max_{i \notin I^{(t-1)}} |x_i^{(t)}|\}$

$x^{(t)}_{I^{(t)}} = A^{\dagger}_{I^{(t)}} y$

$x^{(t)}_{[1,n] \setminus I^{(t)}} = 0$

$r^{(t)} = y - Ax^{(t)}$

$t = t + 1$

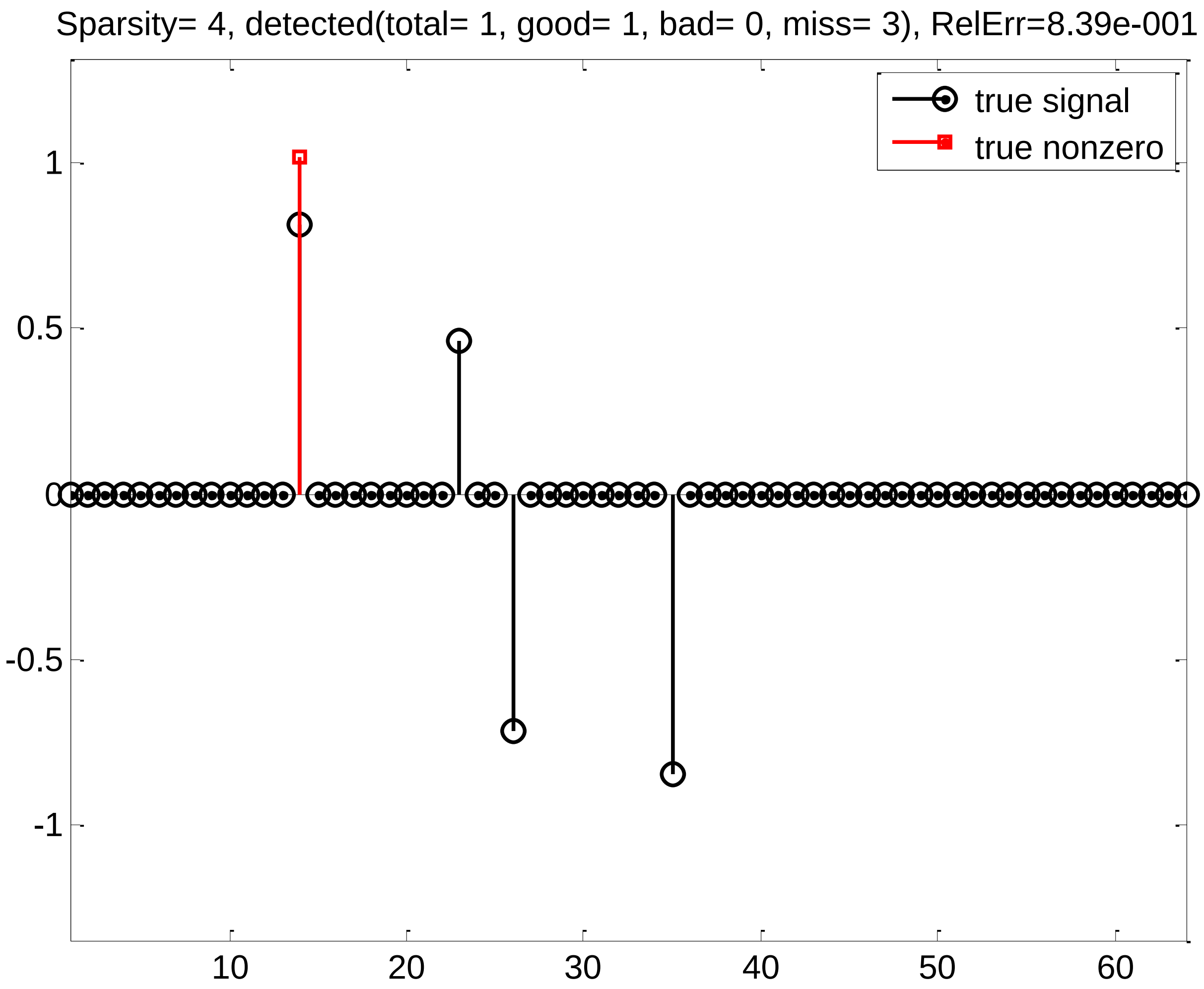
end while

return x

Outline

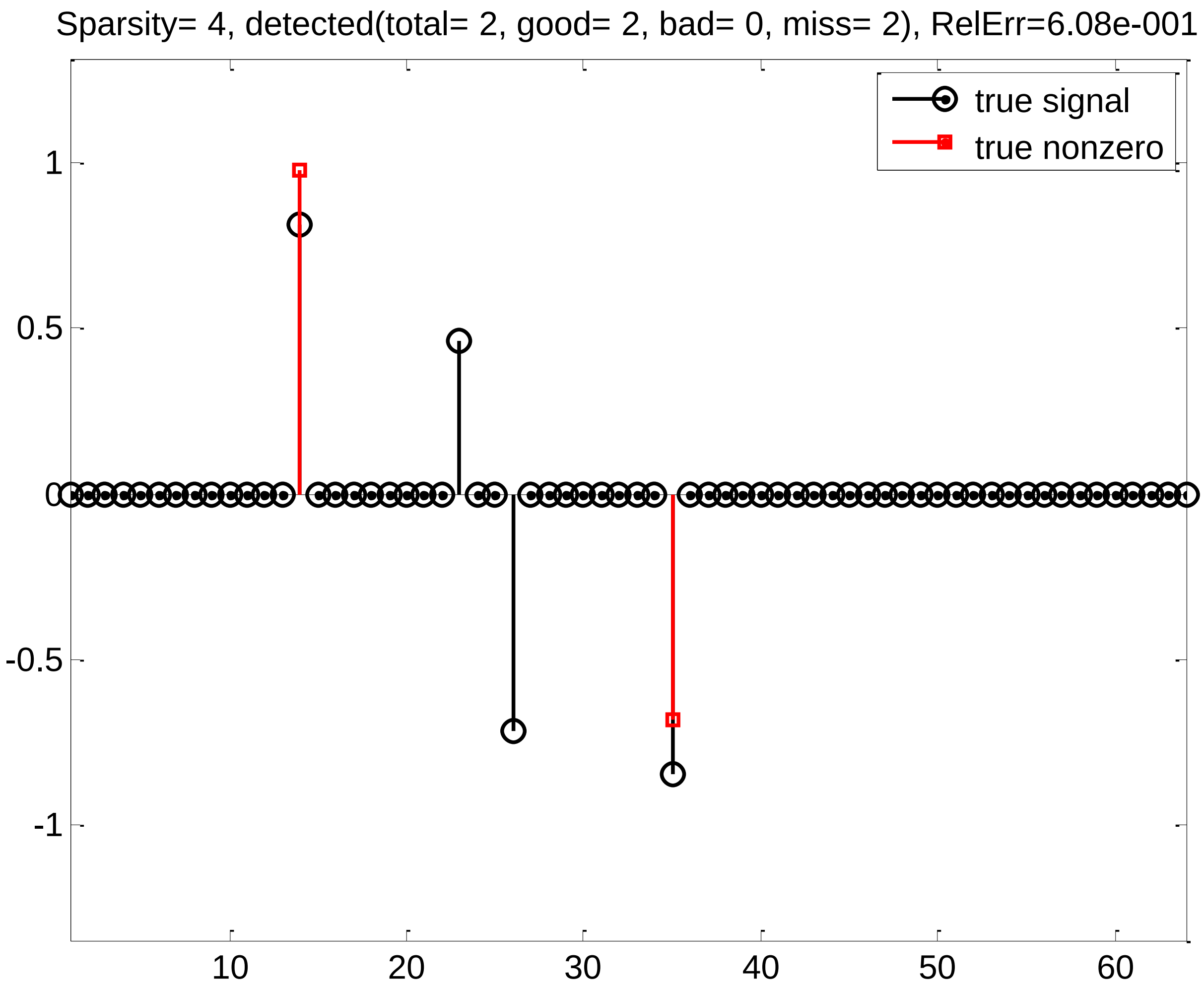
- Introduction
- RMP
- EMTP
- Demos
- Experiments
- Discussions

Demo - RMP(1)



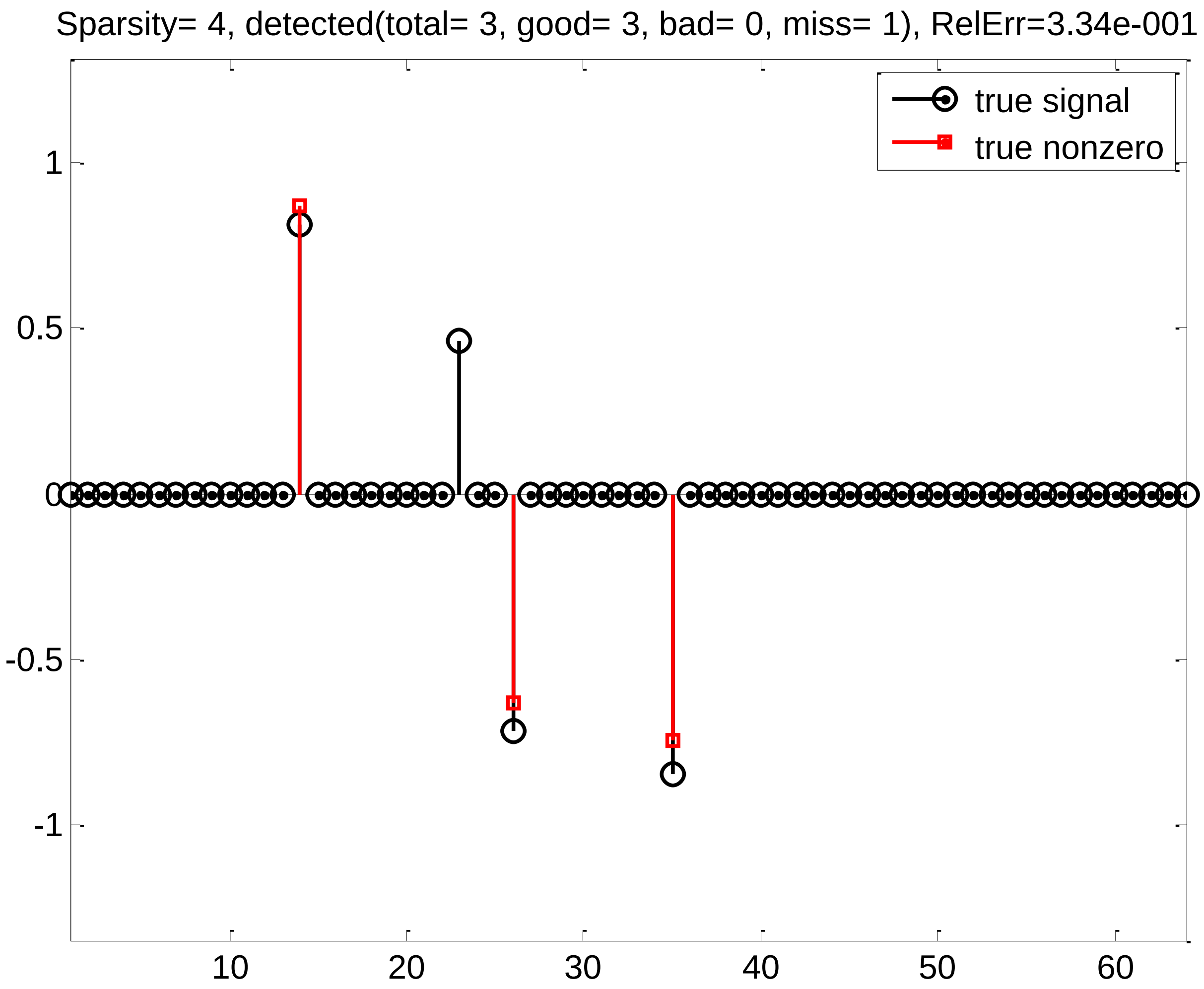
Sparsity= 4, detected(total= 1, good= 1, bad= 0, miss= 3), RelErr=8.39e-001
true signal
true nonzero
1
0.5
0
-0.5
-1
10
20
30
40
50
60

Demo - RMP(2)

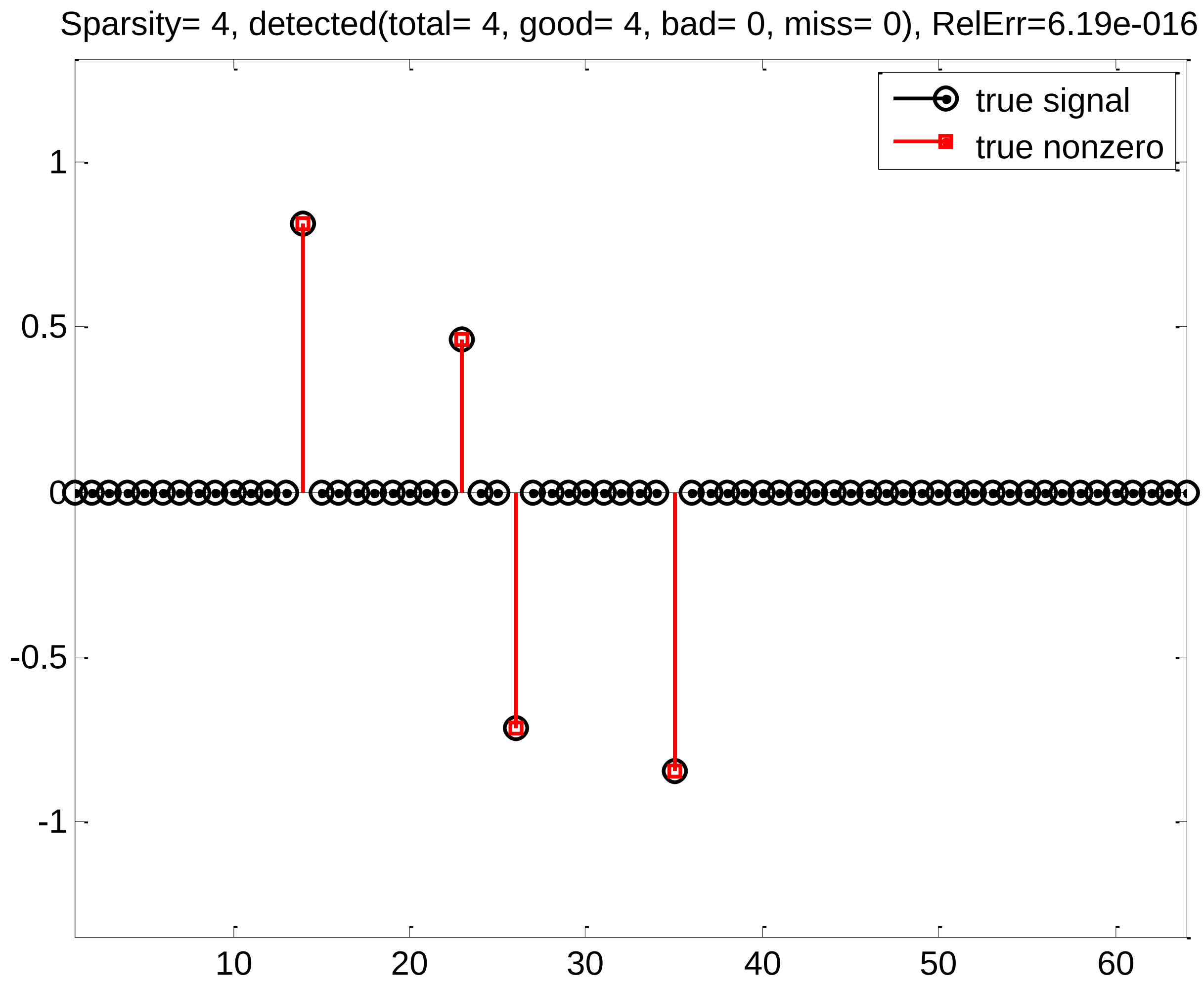


Sparsity= 4, detected(total= 2, good= 2, bad= 0, miss= 2), RelErr=6.08e-001
true signal
true nonzero
1
0.5
0
-0.5
-1
10
20
30
40
50
60

Demo - RMP(3)

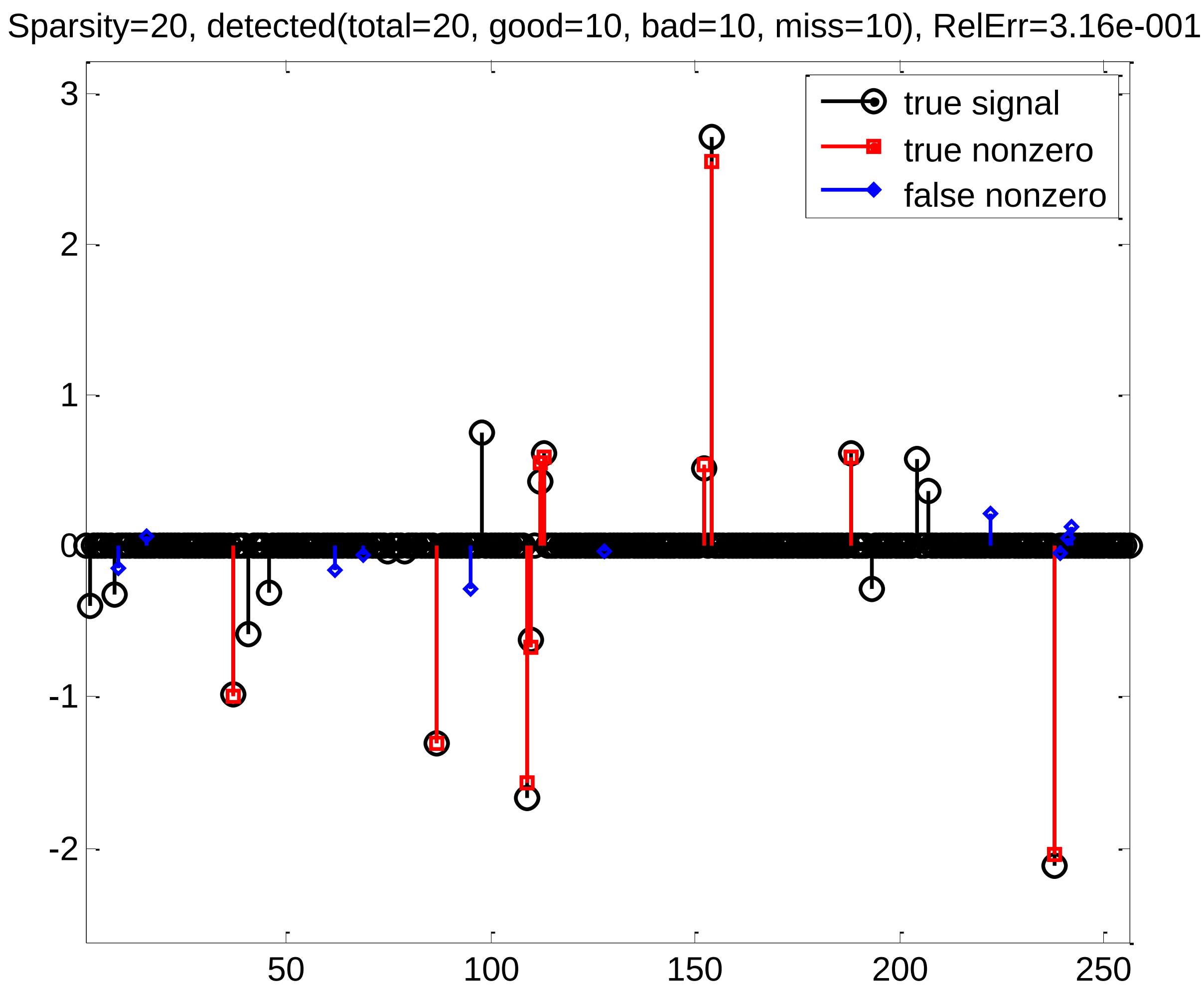

Sparsity= 4, detected(total= 3, good= 3, bad= 0, miss= 1), RelErr=3.34e-001
true signal
true nonzero
1
0.5
0
-0.5
-1
10
20
30
40
50
60

Demo - RMP(4)

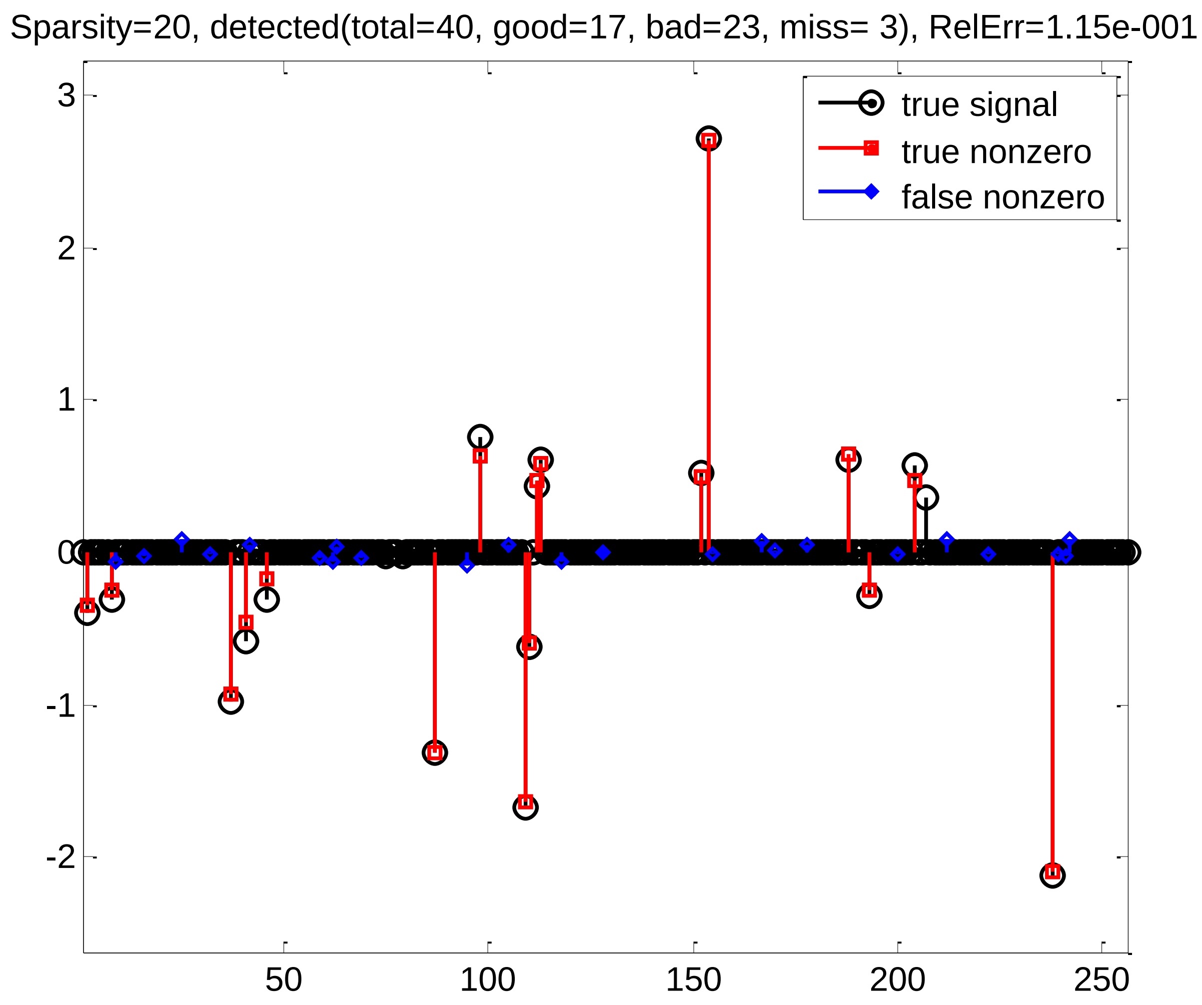


Sparsity= 4, detected(total= 4, good= 4, bad= 0, miss= 0), RelErr=6.19e-016
true signal
true nonzero
1
0.5
0
-0.5
-1
10
20
30
40
50
60

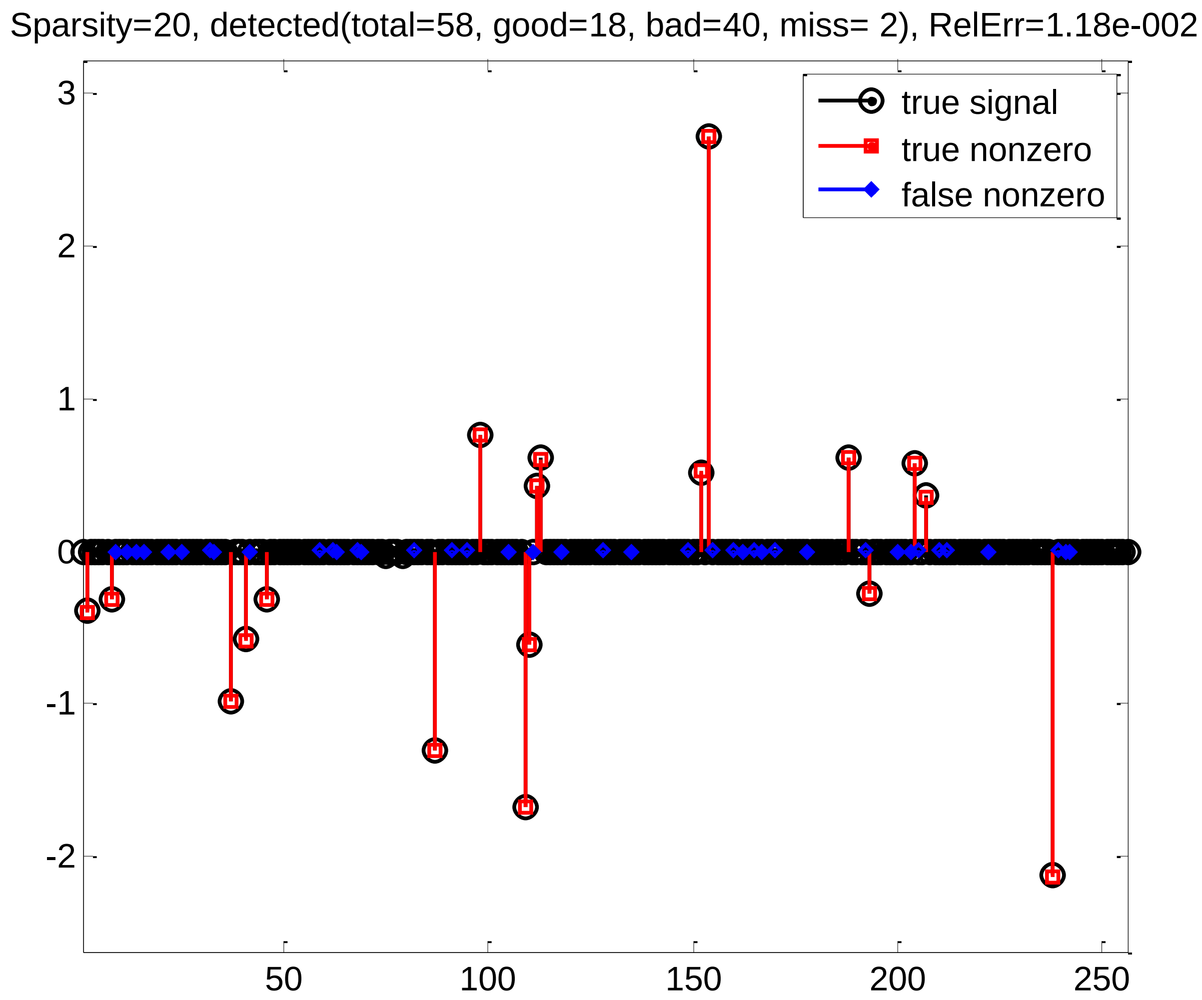
Demo - *k*RMP(1)


Sparsity=20, detected(total=20, good=10, bad=10, miss=10), RelErr=3.16e-001
true signal
true nonzero
false nonzero
3
2
1
0
-1
-2
50
100
150
200
250

Demo - *k*RMP(2)

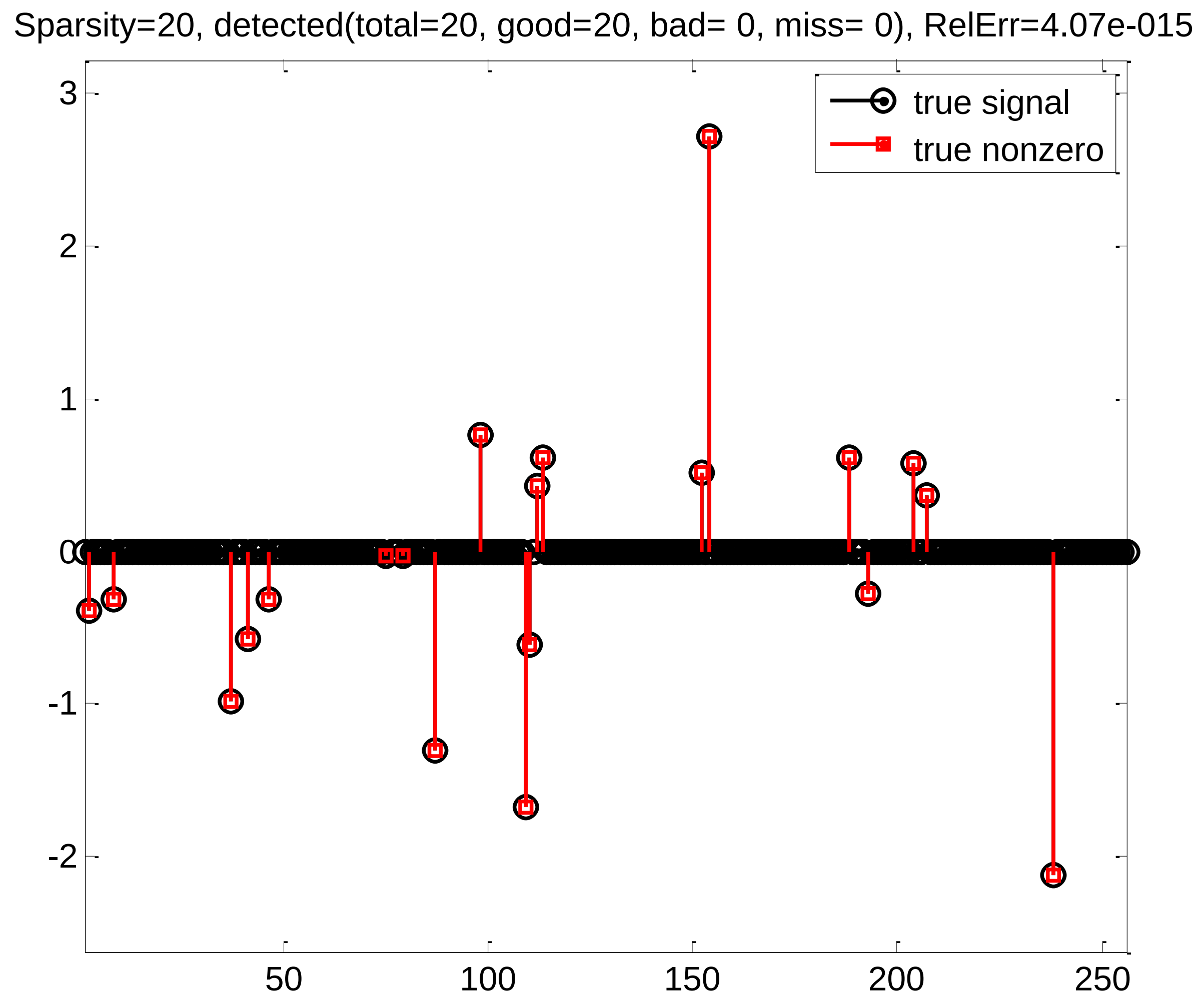

Sparsity=20, detected(total=40, good=17, bad=23, miss= 3), RelErr=1.15e-001
true signal
true nonzero
false nonzero
3
2
1
0
-1
-2
50
100
150
200
250

Demo - *k*RMP(3)



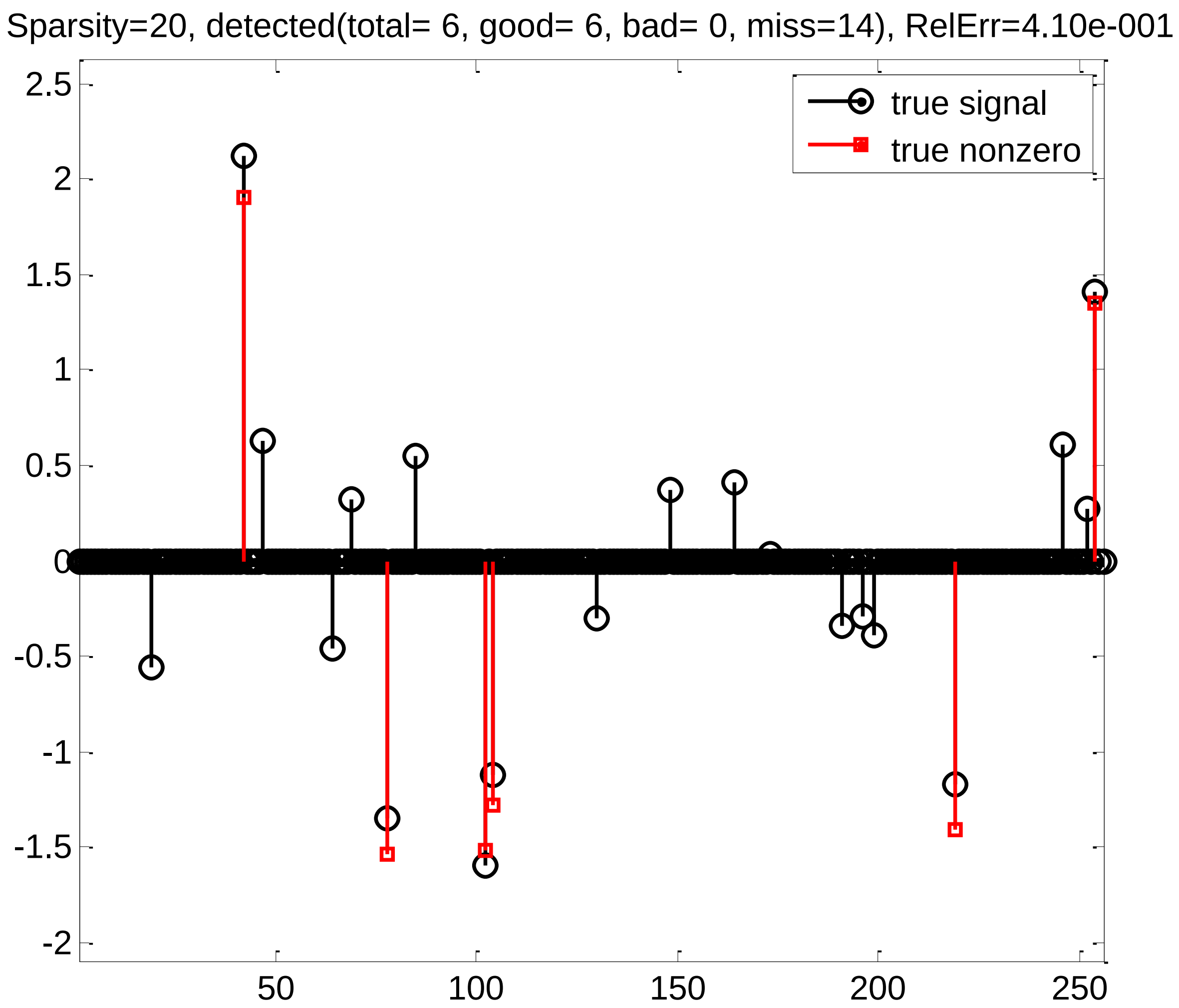
Sparsity=20, detected(total=58, good=18, bad=40, miss= 2), RelErr=1.18e-002
true signal
true nonzero
false nonzero
3
2
1
0
-1
-2
50
100
150
200
250

Demo - *k*RMP(4)

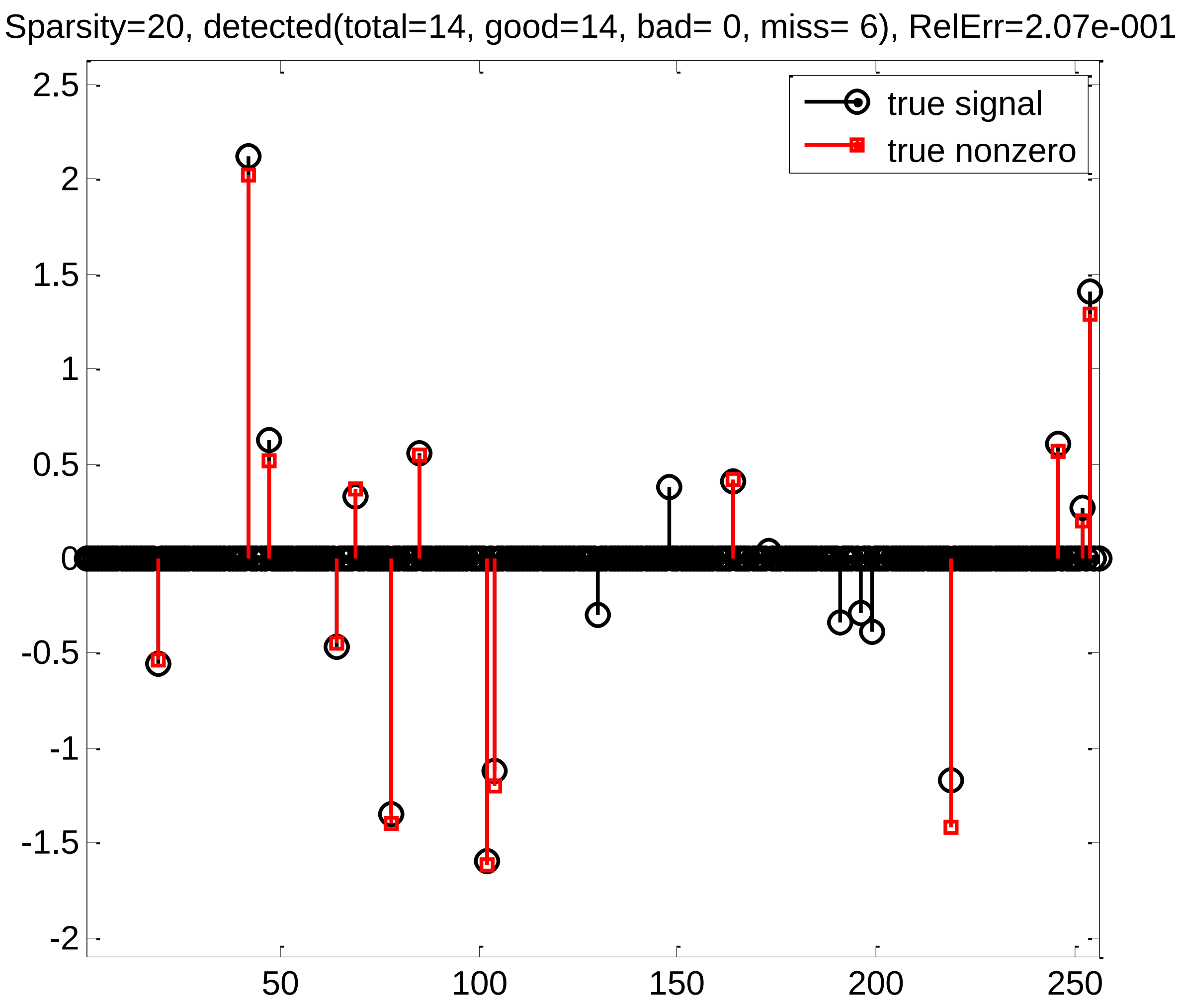


Sparsity=20, detected(total=20, good=20, bad= 0, miss= 0), RelErr=4.07e-015
true signal
true nonzero
3
2
1
0
-1
-2
50
100
150
200
250

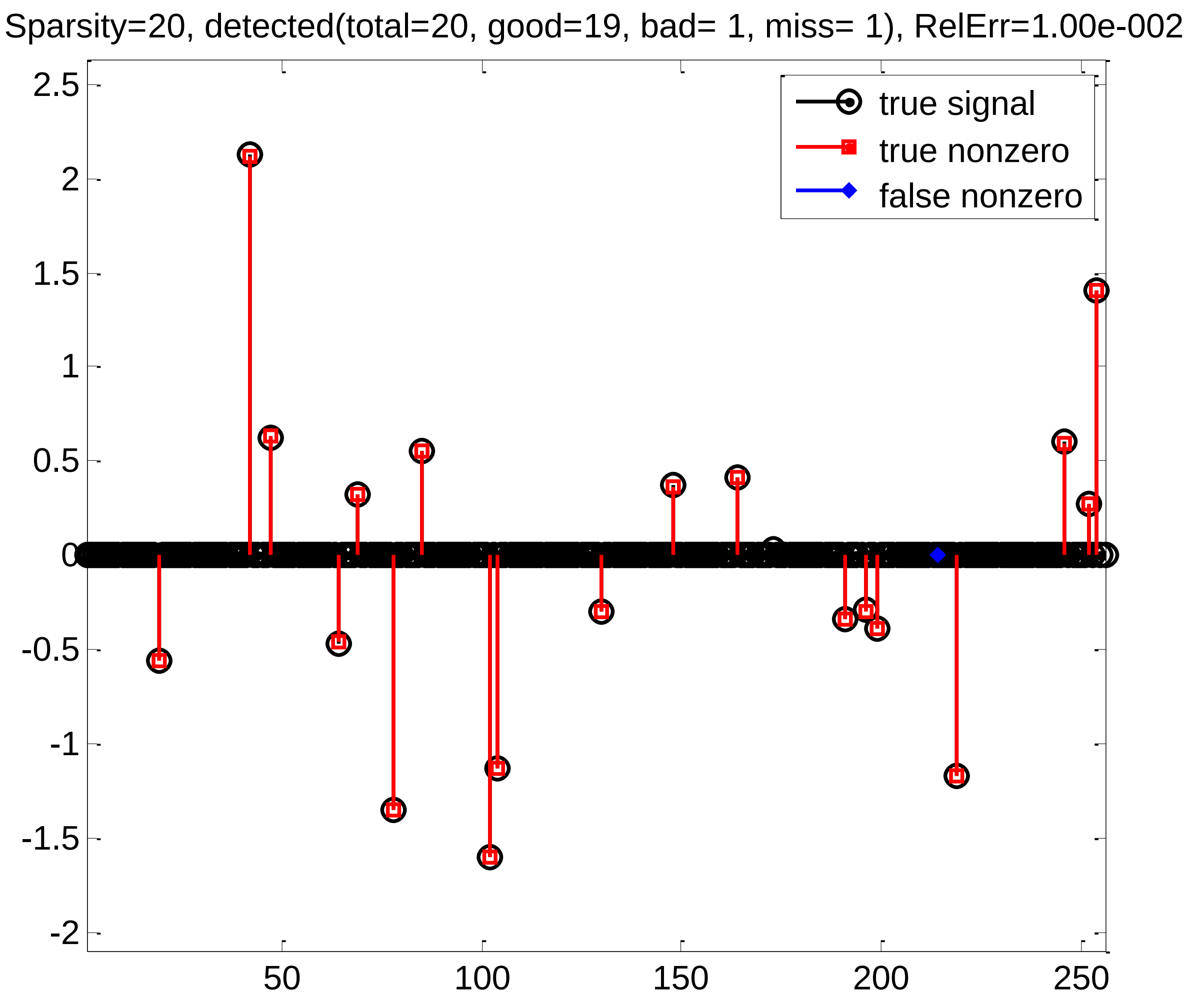
Demo - βRMP(1)



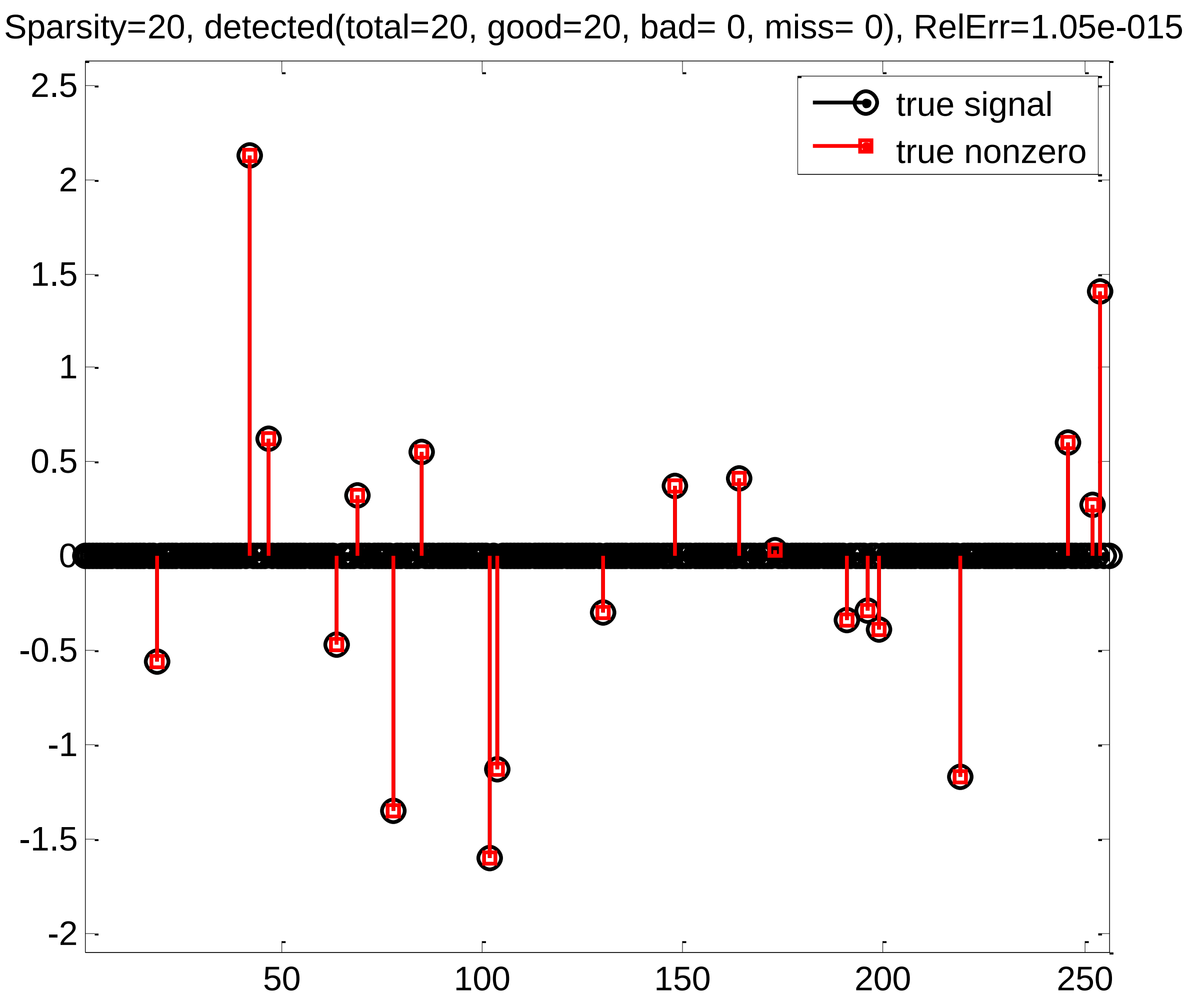
Demo - βRMP(2)



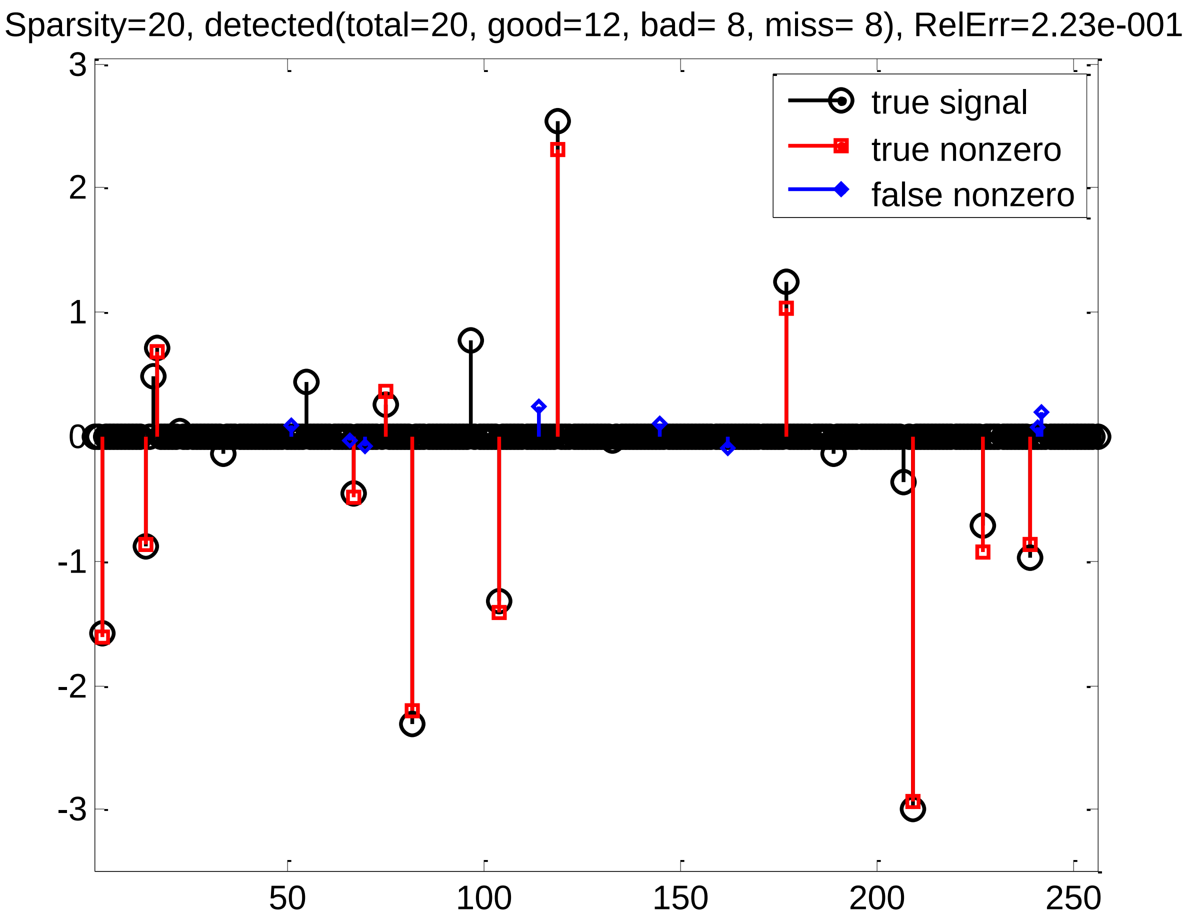
Demo - βRMP(3)



Demo - βRMP(4)

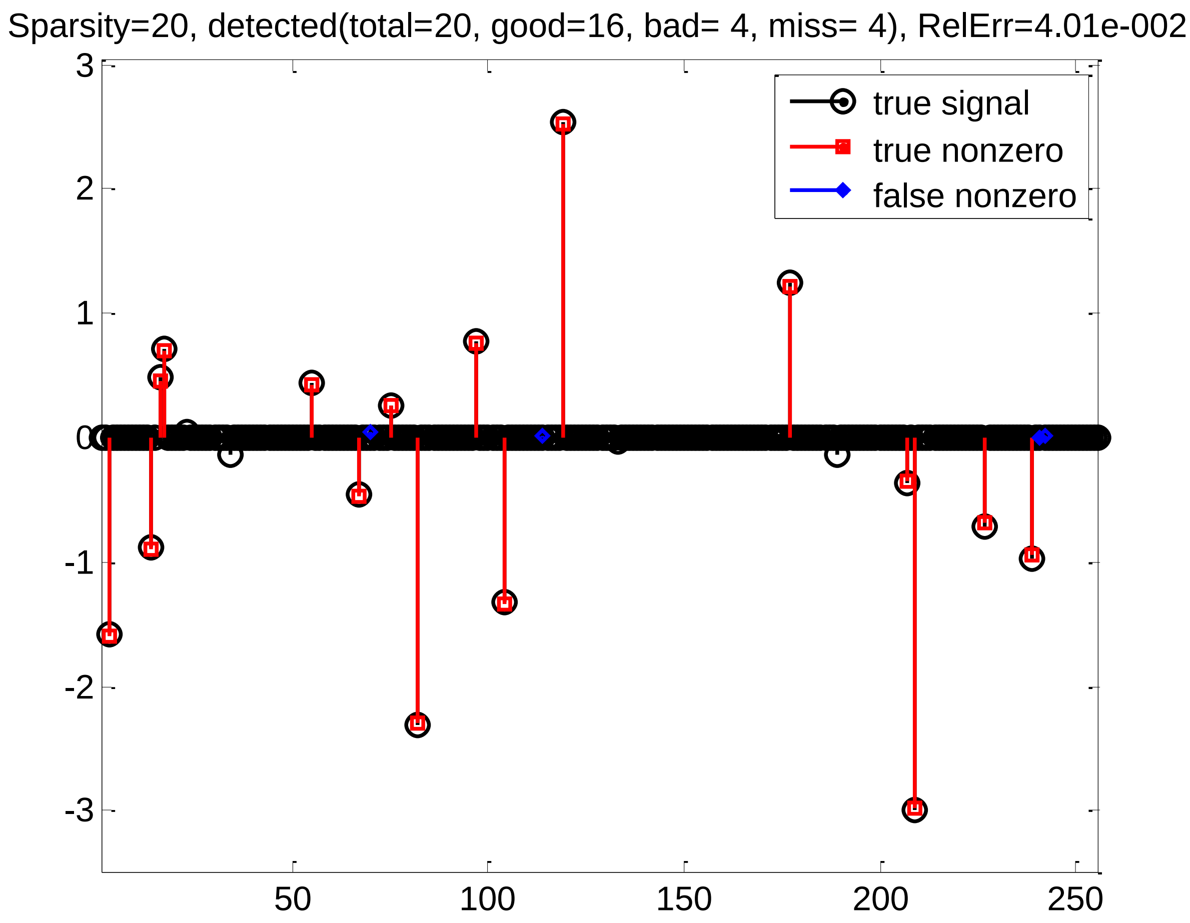


Demo - *k*EMTP(1)



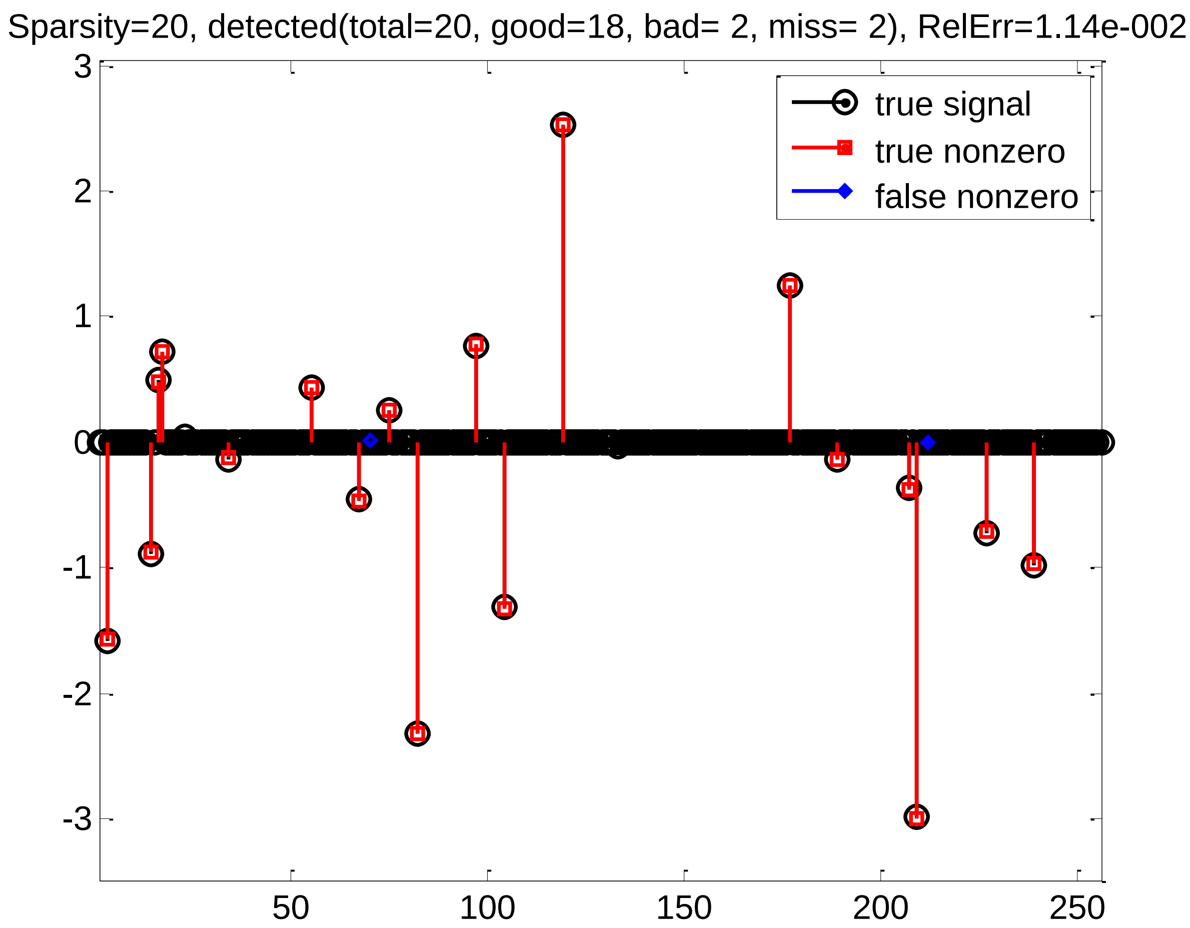
Sparsity=20, detected(total=20, good=12, bad= 8, miss= 8), RelErr=2.23e-001
true signal
true nonzero
false nonzero
3
2
1
0
-1
-2
-3
50
100
150
200
250

Demo - *k*EMTP(2)

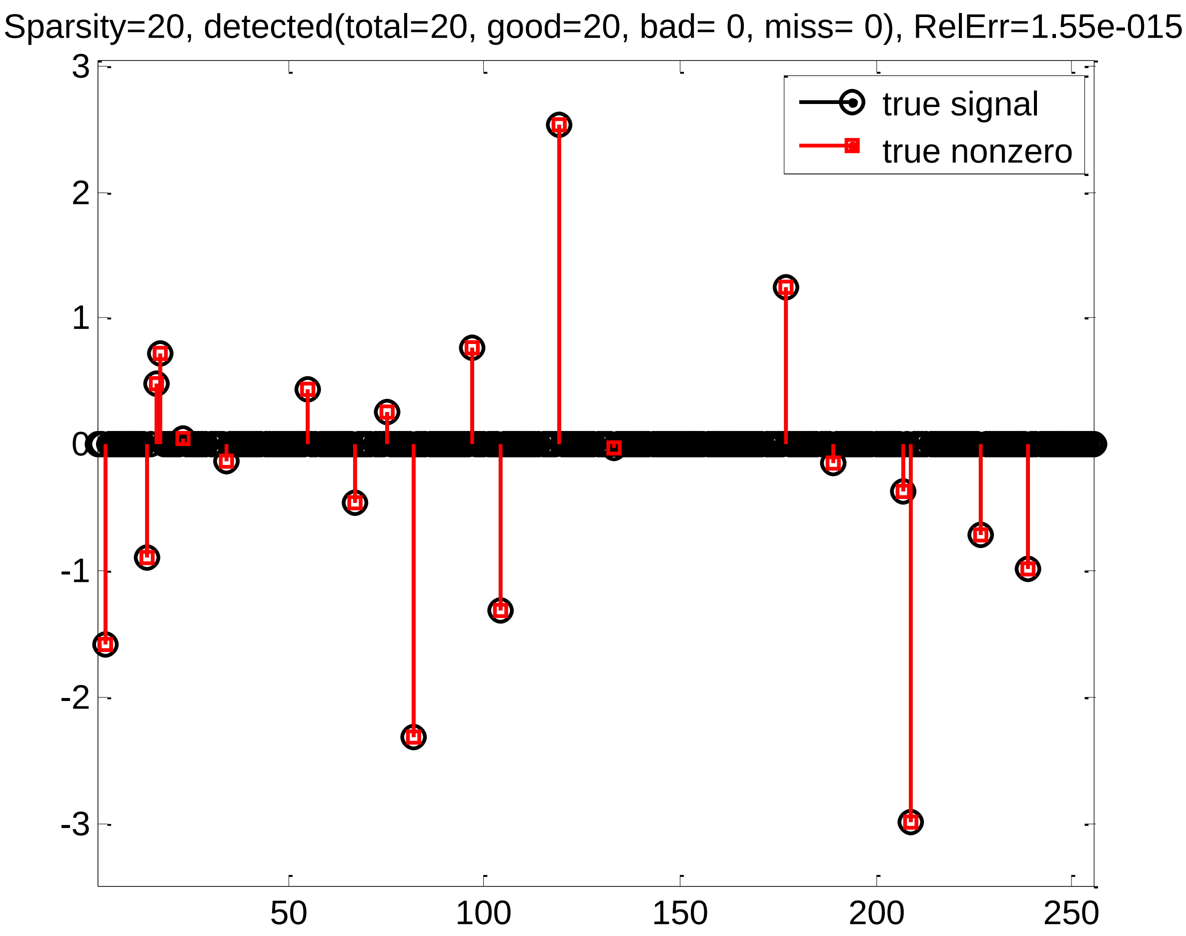


Sparsity=20, detected(total=20, good=16, bad= 4, miss= 4), RelErr=4.01e-002
true signal
true nonzero
false nonzero
3
2
1
0
-1
-2
-3
50
100
150
200
250

Demo - *k*EMTP(3)

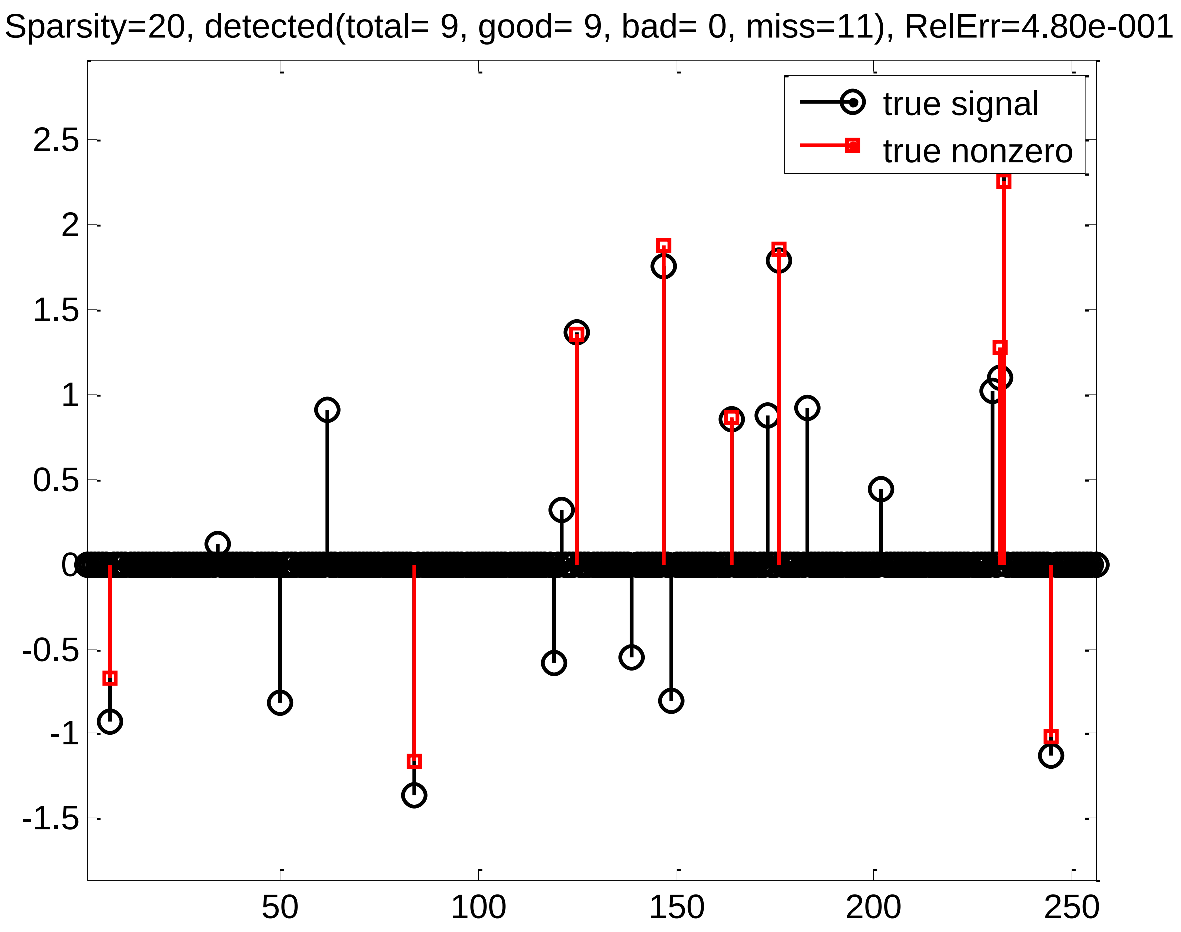

Sparsity=20, detected(total=20, good=18, bad= 2, miss= 2), RelErr=1.14e-002
true signal
true nonzero
false nonzero
3
2
1
0
-1
-2
-3
50
100
150
200
250

Demo - *k*EMTP(4)

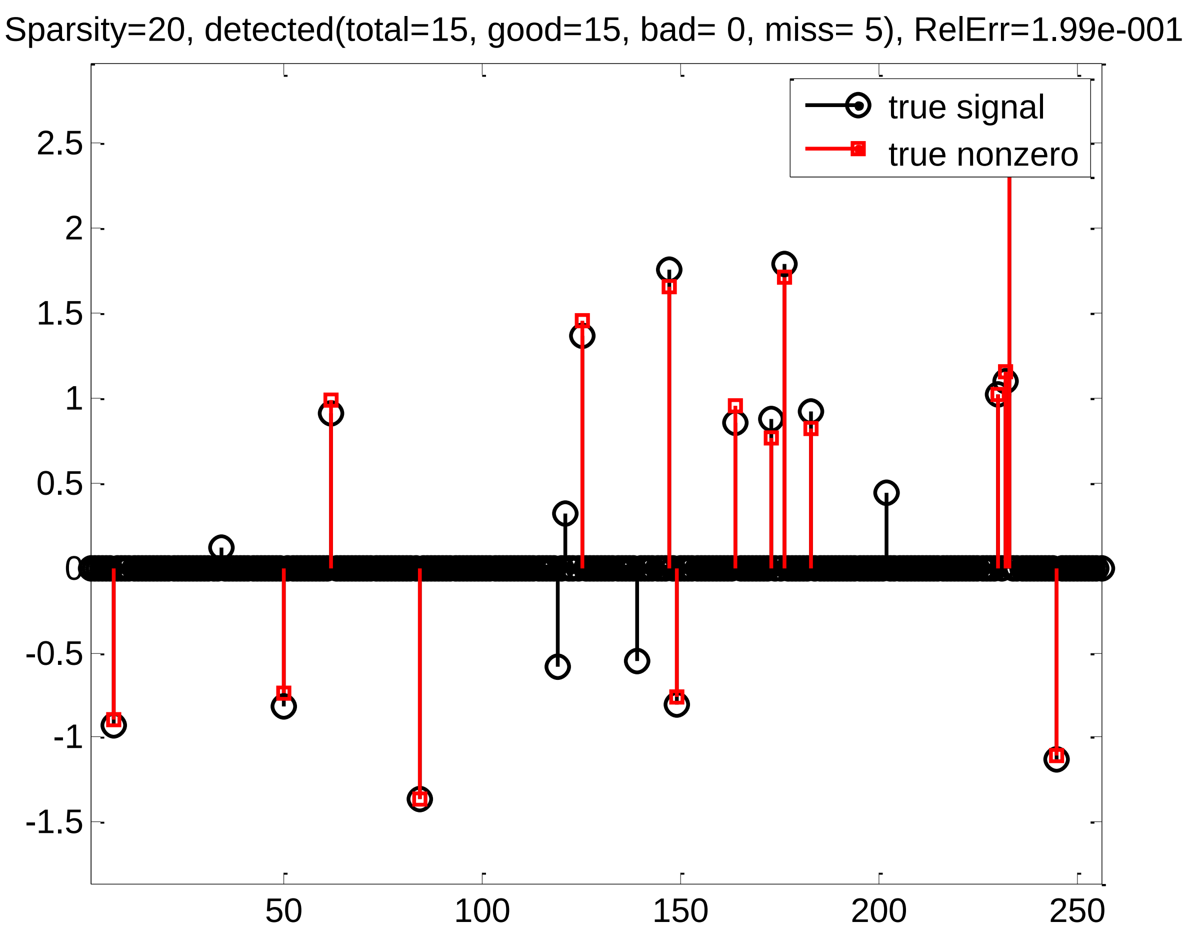


Sparsity=20, detected(total=20, good=20, bad= 0, miss= 0), RelErr=1.55e-015
true signal
true nonzero
3
2
1
0
-1
-2
-3
50
100
150
200
250

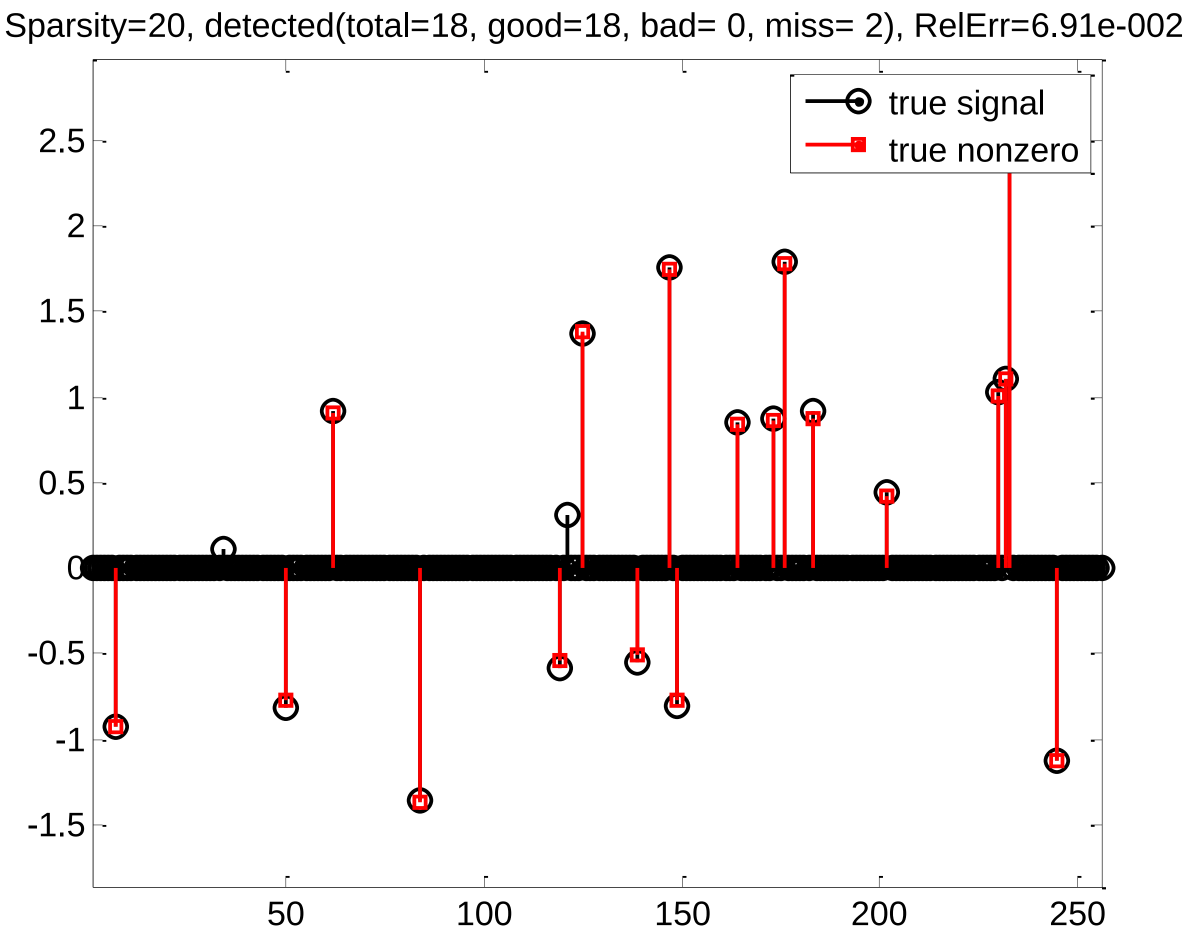
Demo - βEMTP(1)



Demo - βEMTP(2)

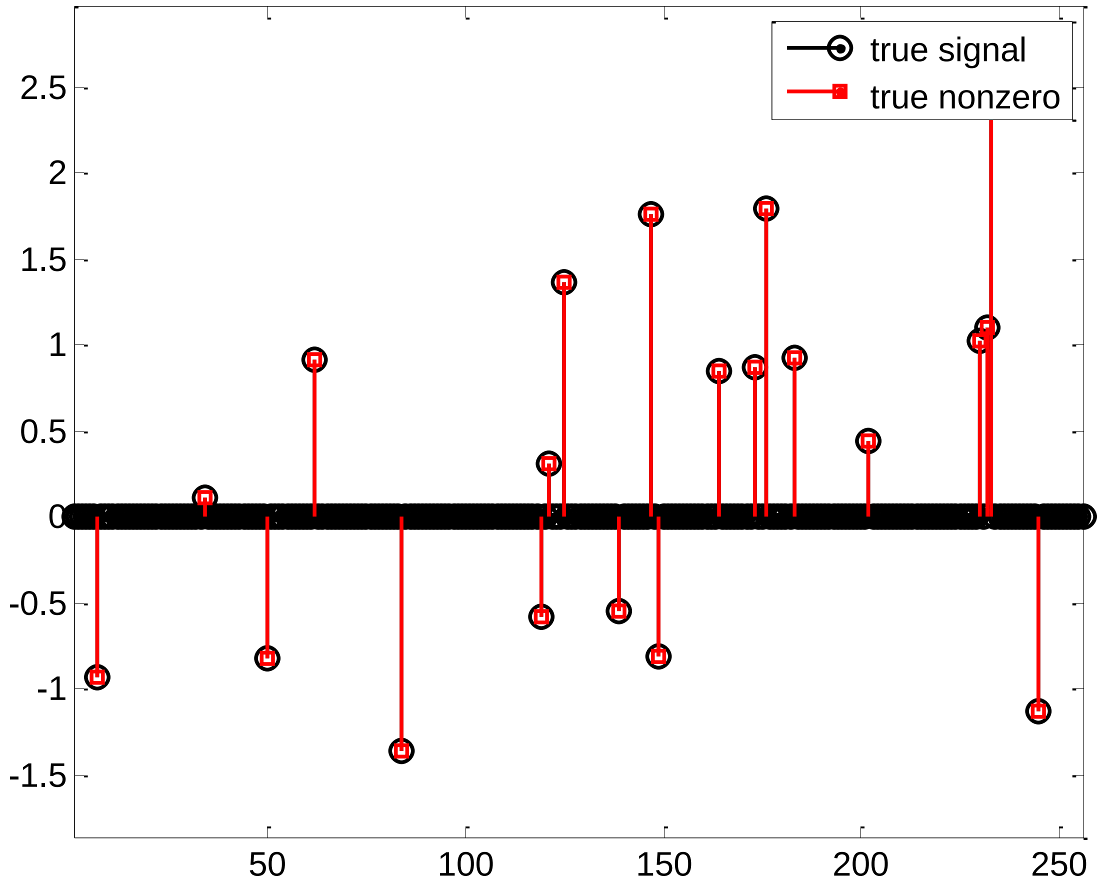


Demo - βEMTP(3)

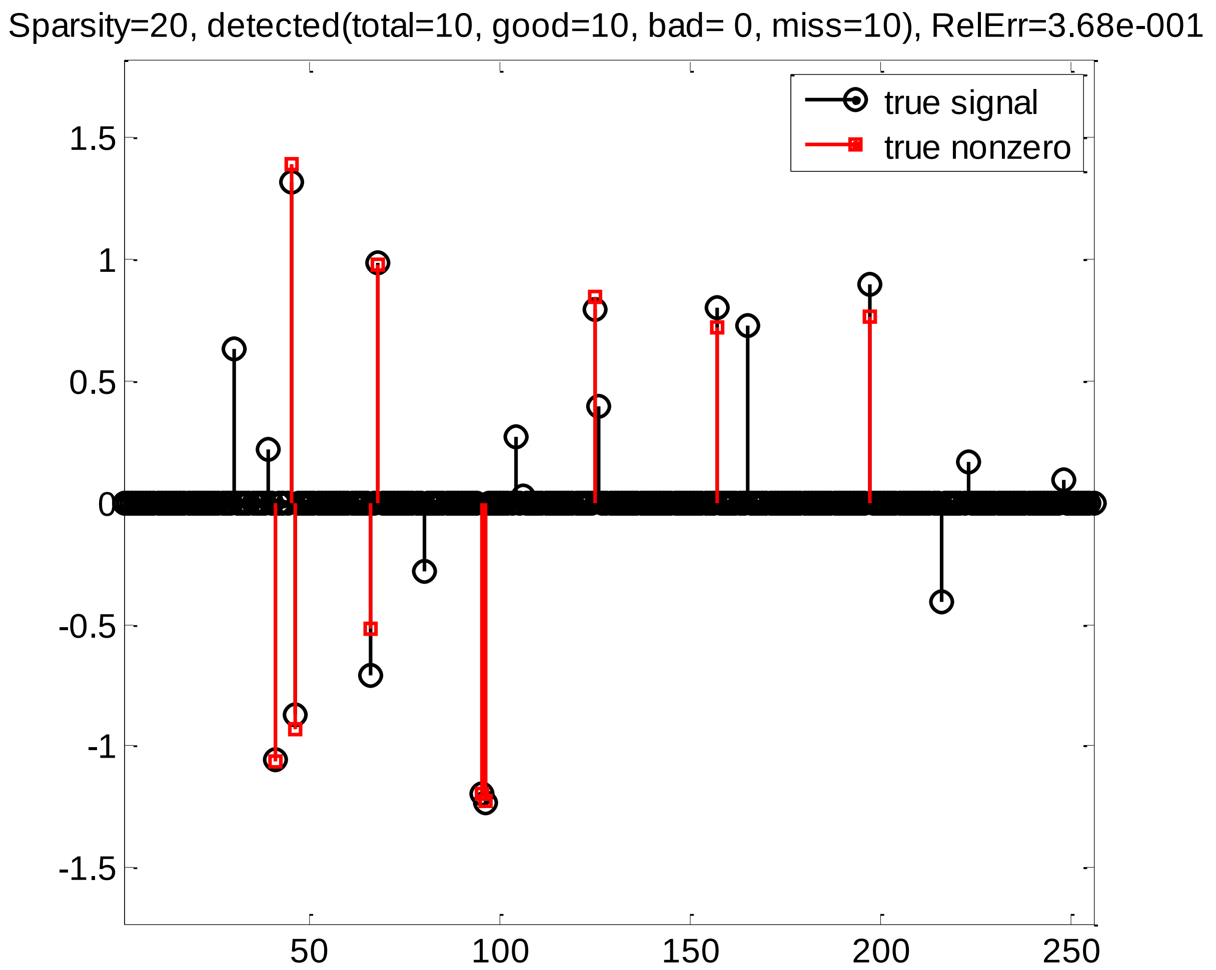


Demo - βEMTP(4)

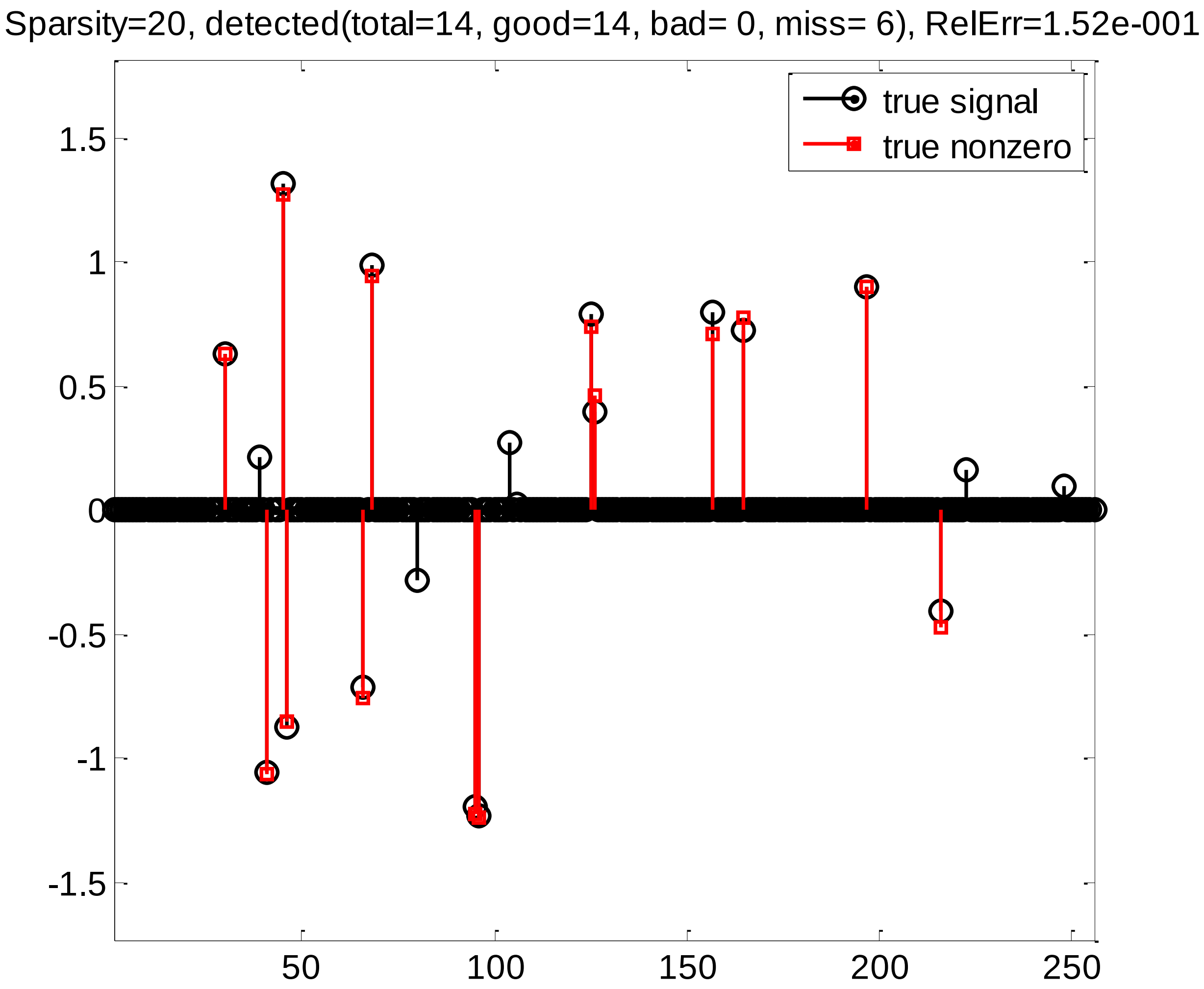
Sparsity=20, detected(total=20, good=20, bad= 0, miss= 0), RelErr=1.50e-015



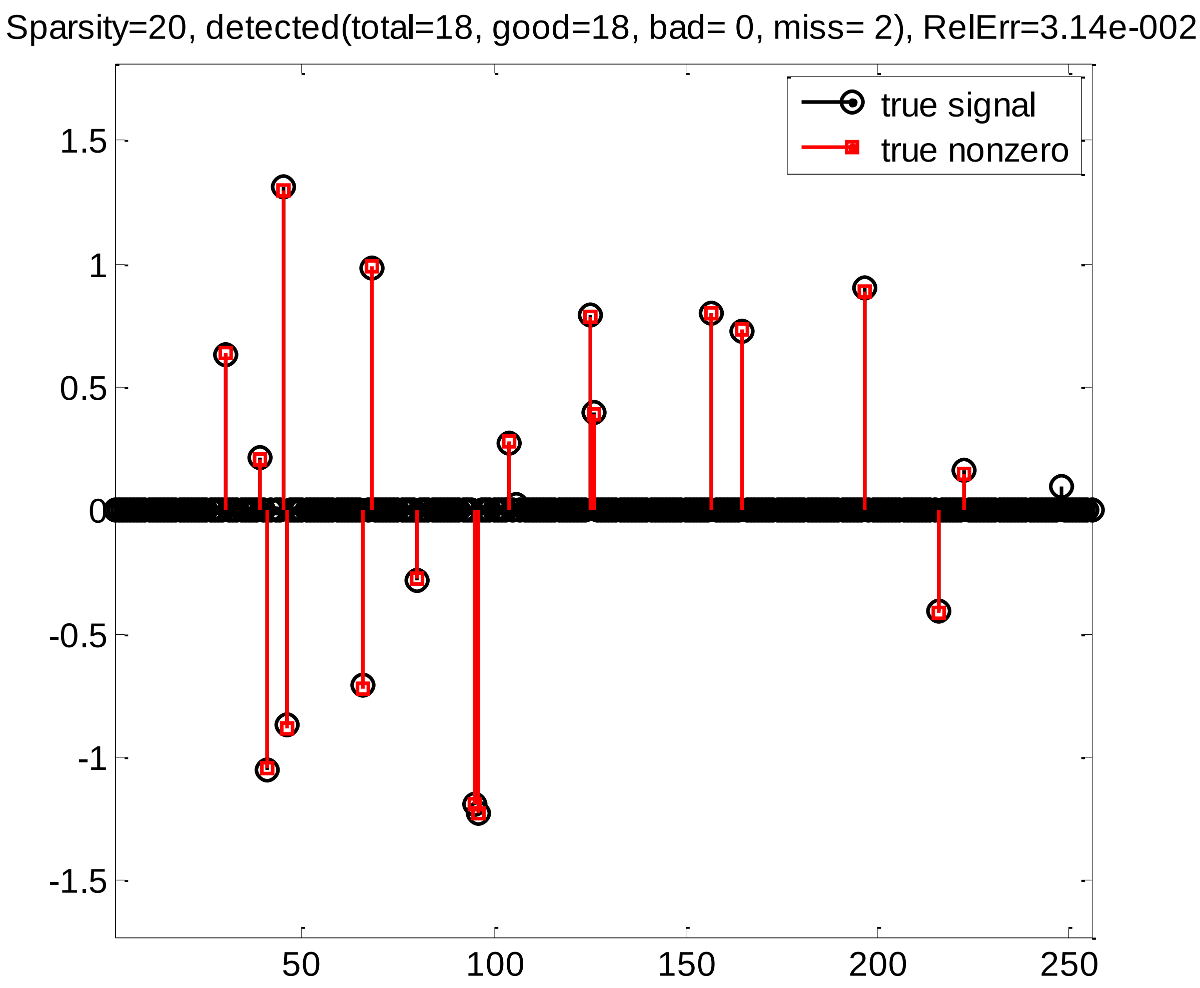
Demo – γEMTP(1)



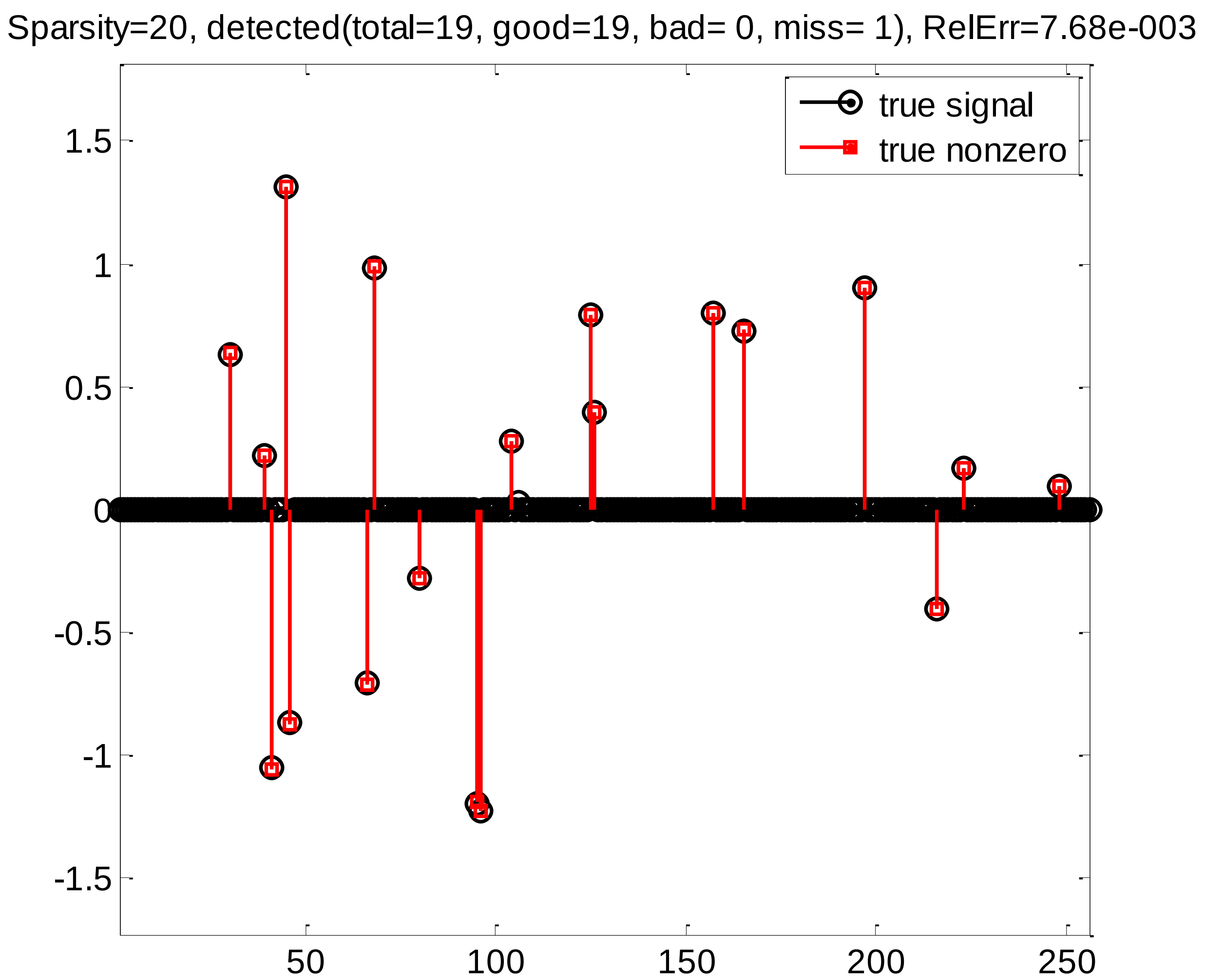
Demo – γEMTP(2)



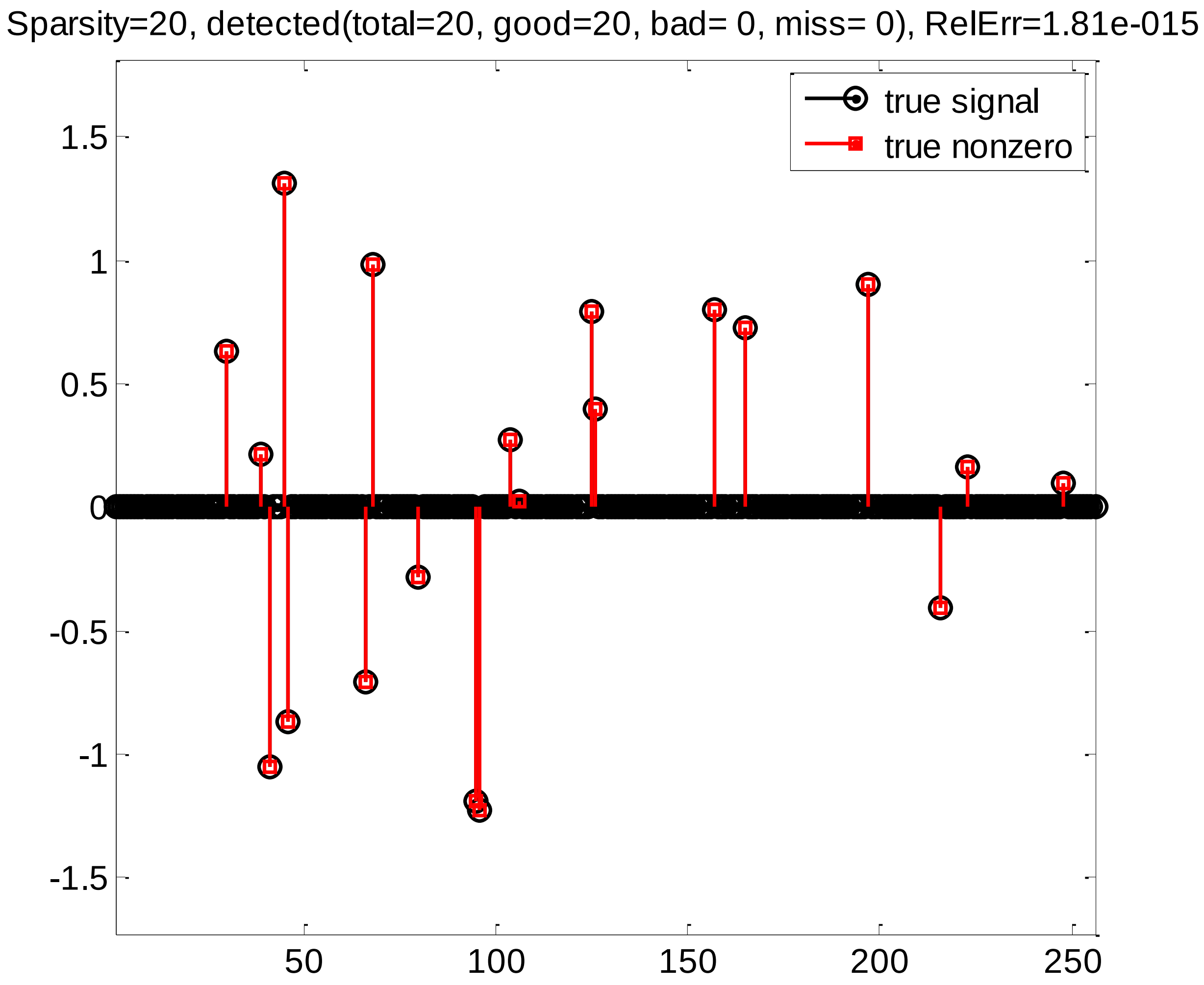
Demo – γEMTP(3)



Demo – γEMTP(4)

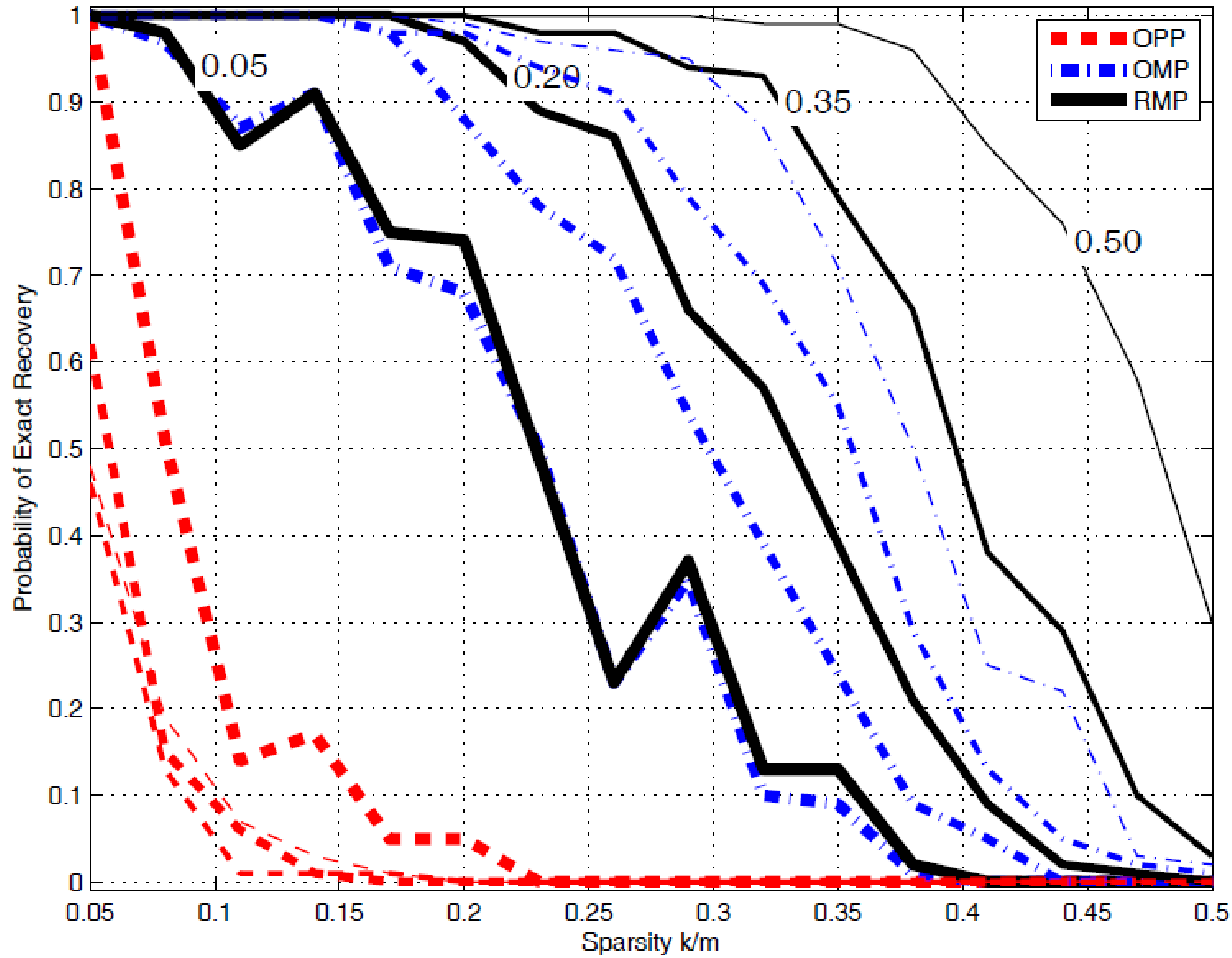


Demo – γEMTP(5)



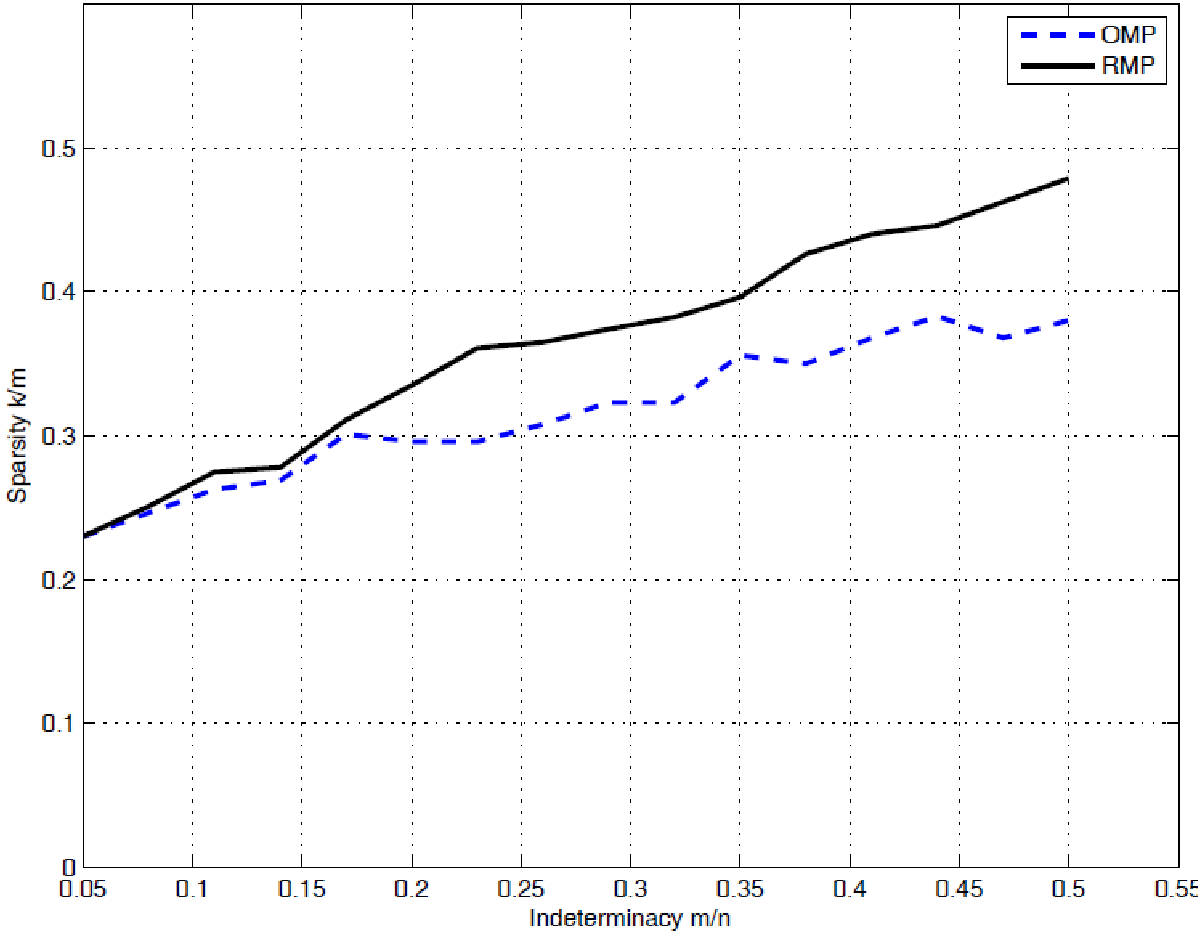
Outline

Comparisons-RMP(1)



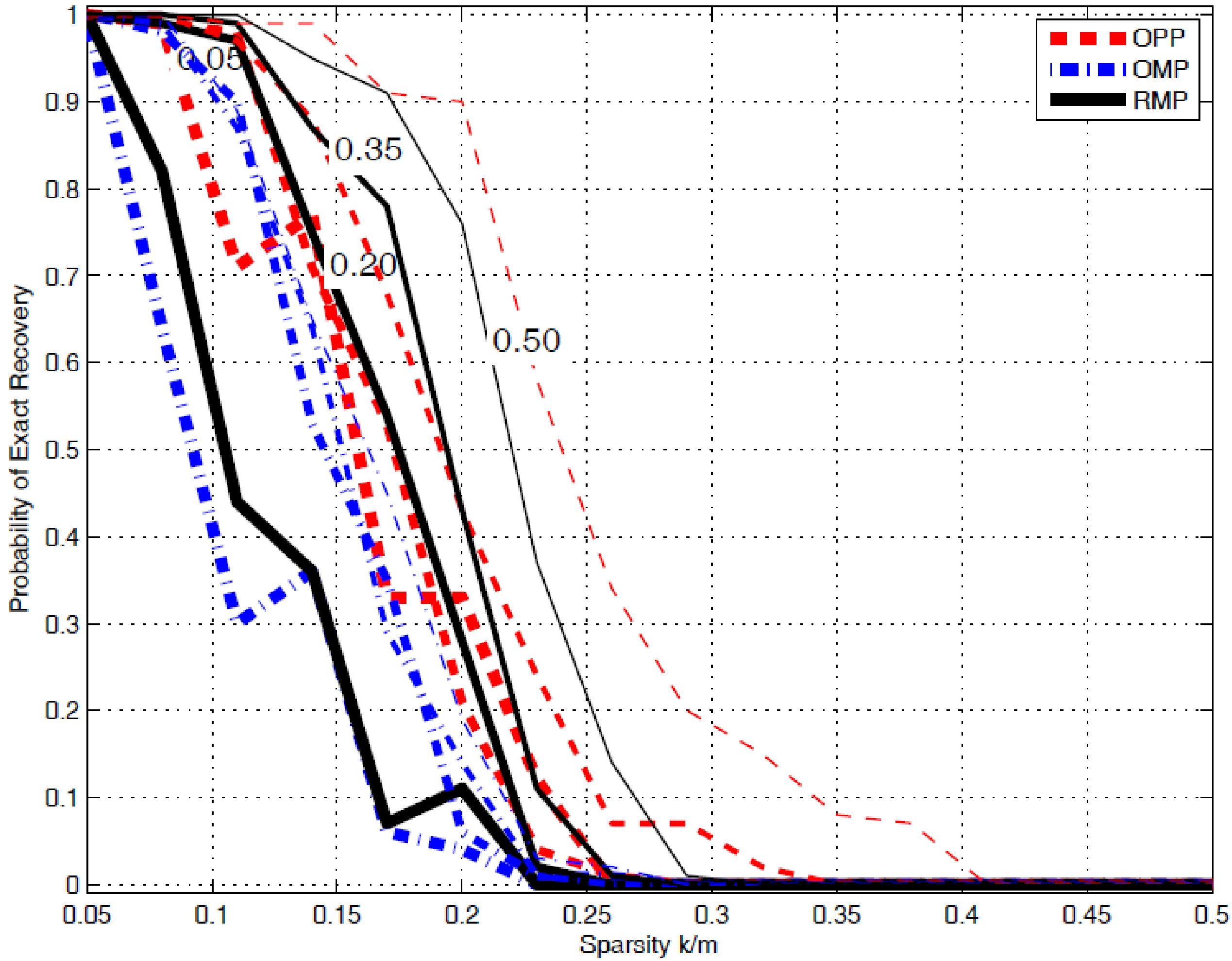
Sparse Gaussian signal, n=400

Comparisons-RMP(2)



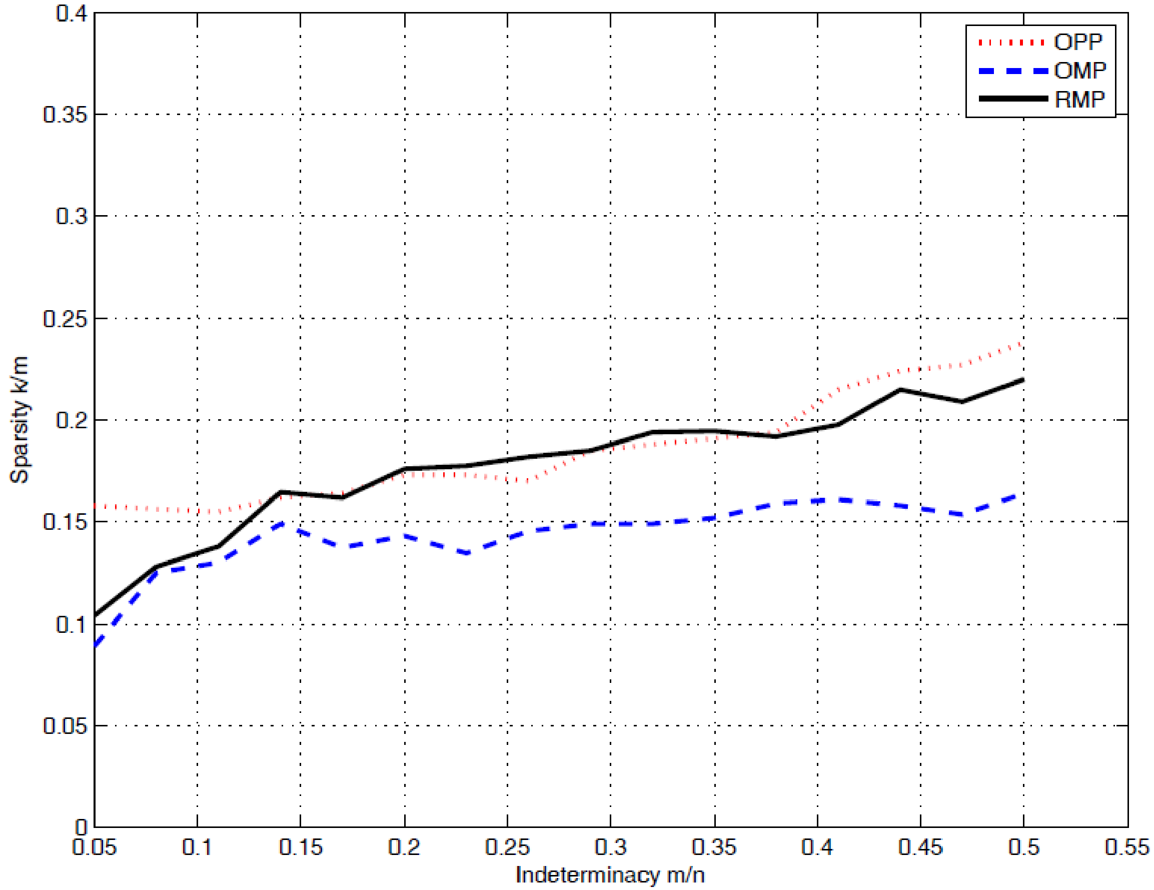
Phase transition,
Sparse Gaussian signal, n=400

Comparisons-RMP(3)



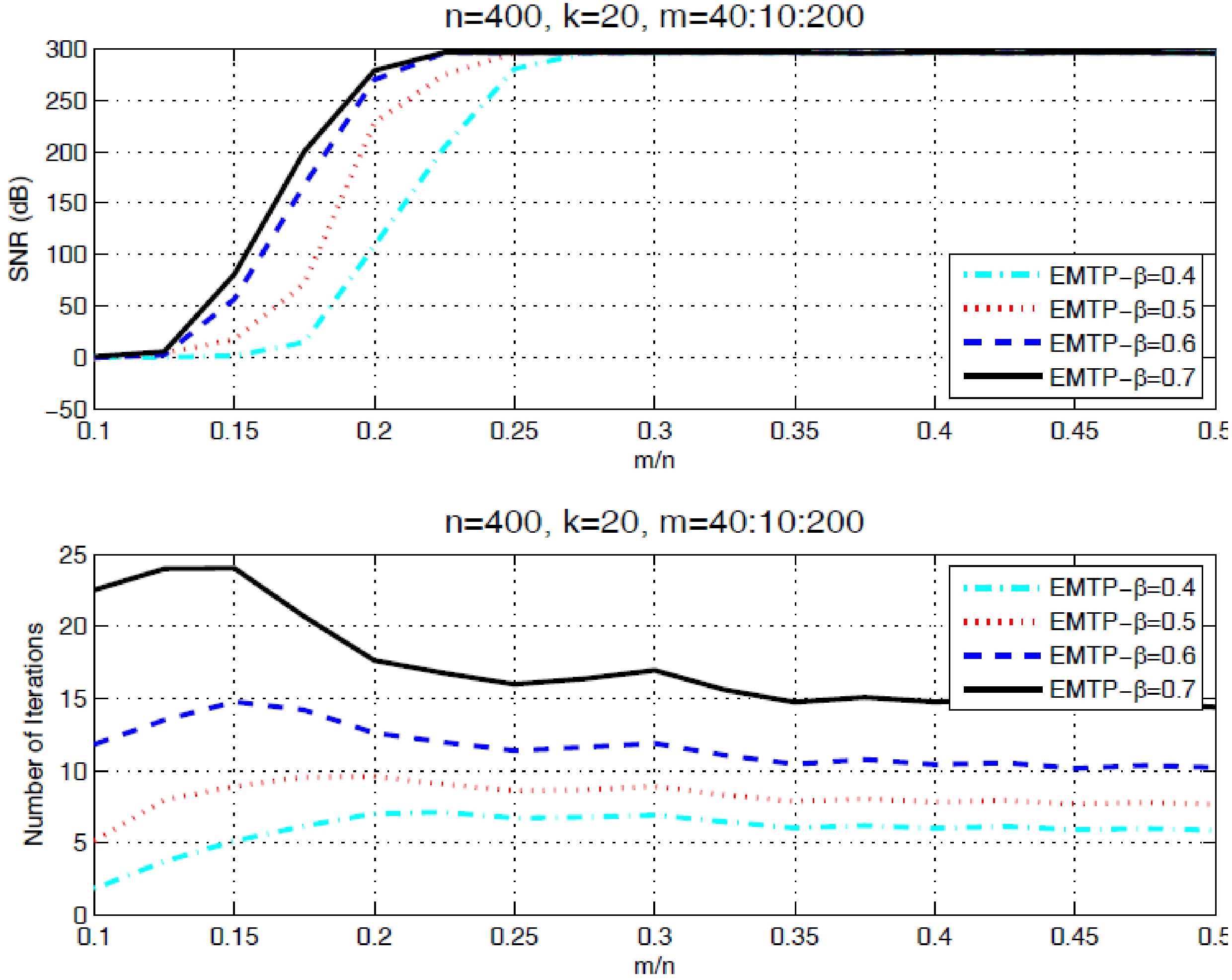
Sparse Bernoulli signal, n=400

Comparisons-RMP(4)



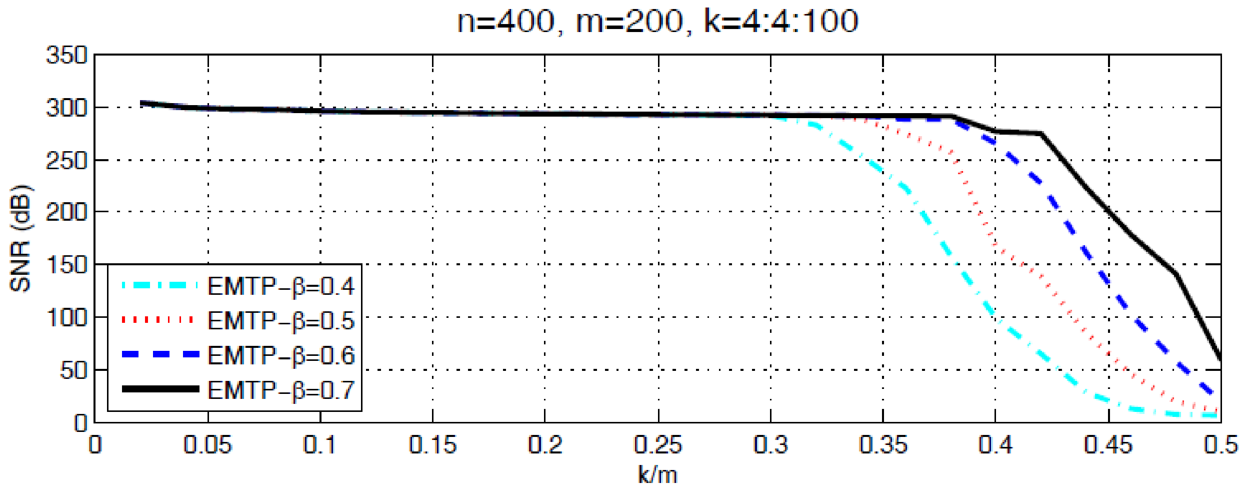
Phase transition,
Sparse Bernoulli signal, n=400

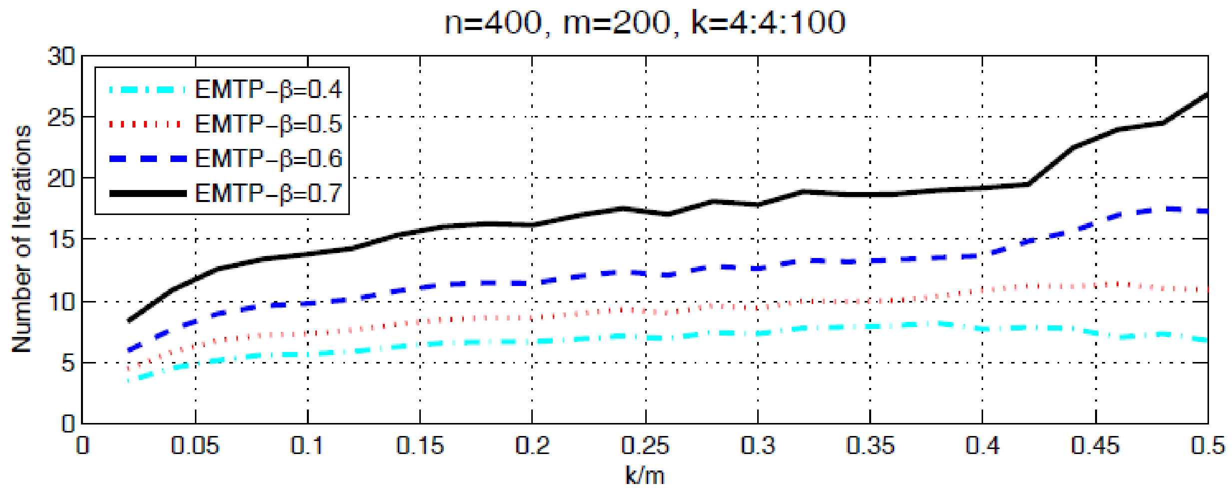
Comparisons-EMTP(1)



Sparse Gaussian signal

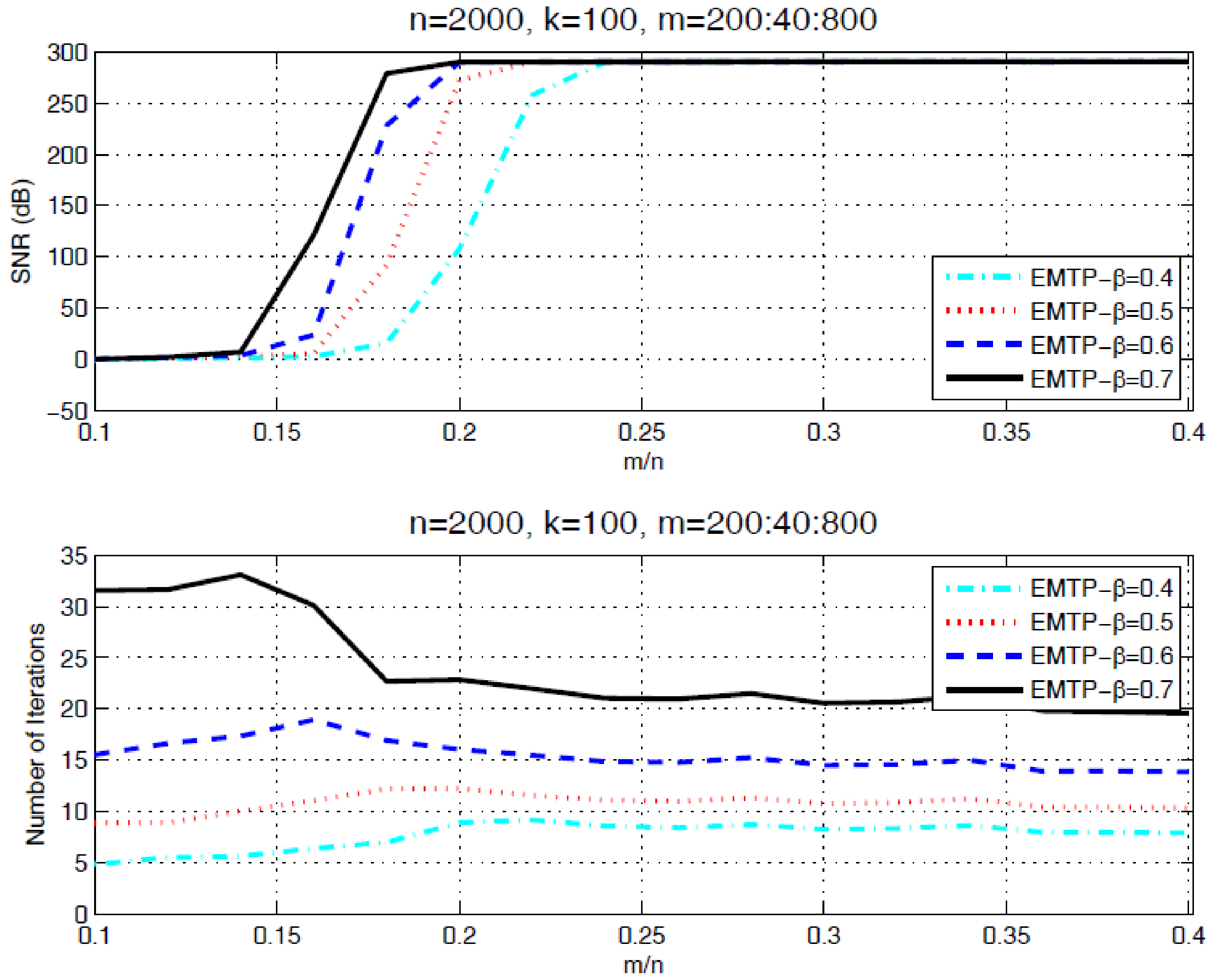
Comparisons-EMTP(2)





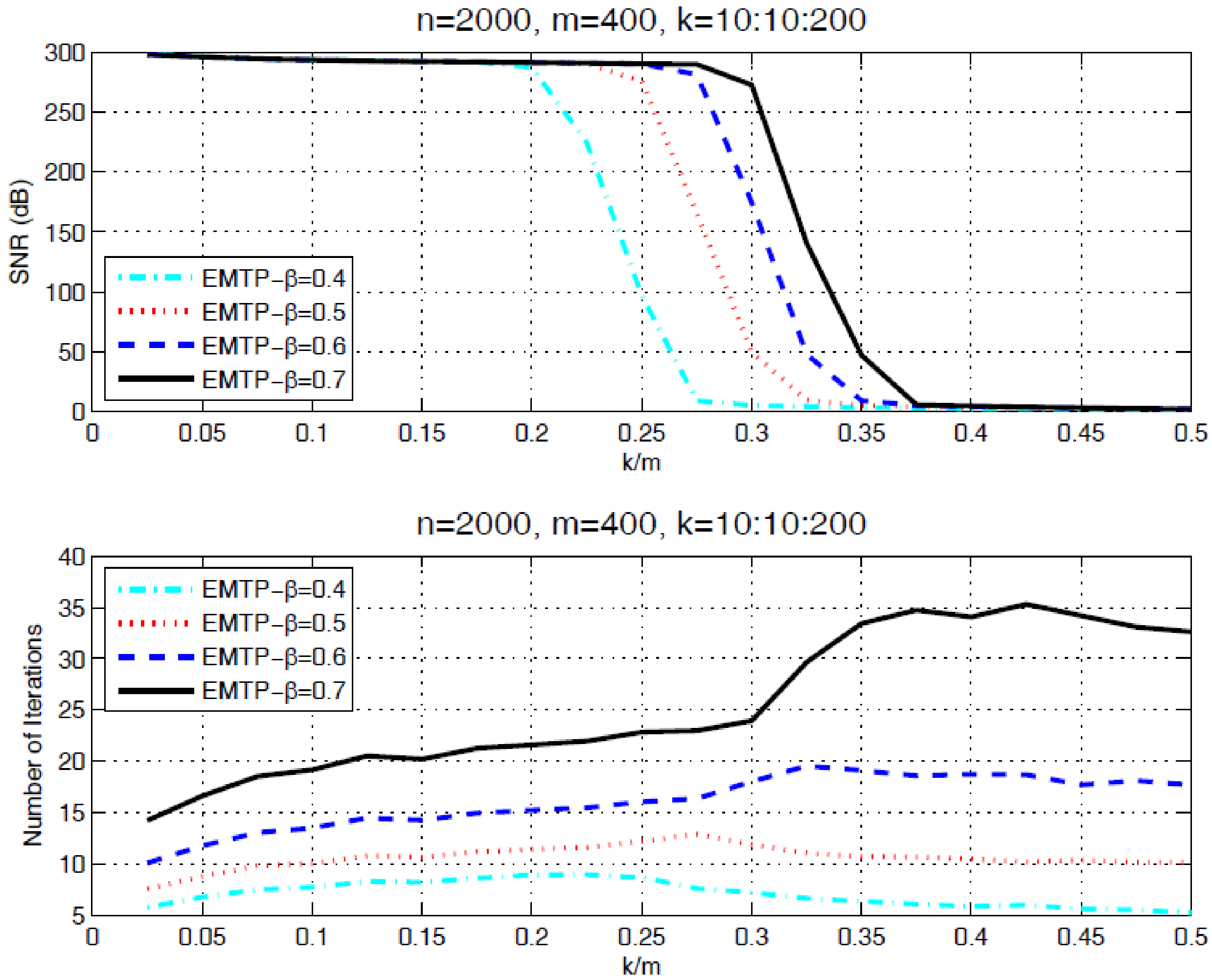
Sparse Gaussian signal

Comparisons-EMTP(3)



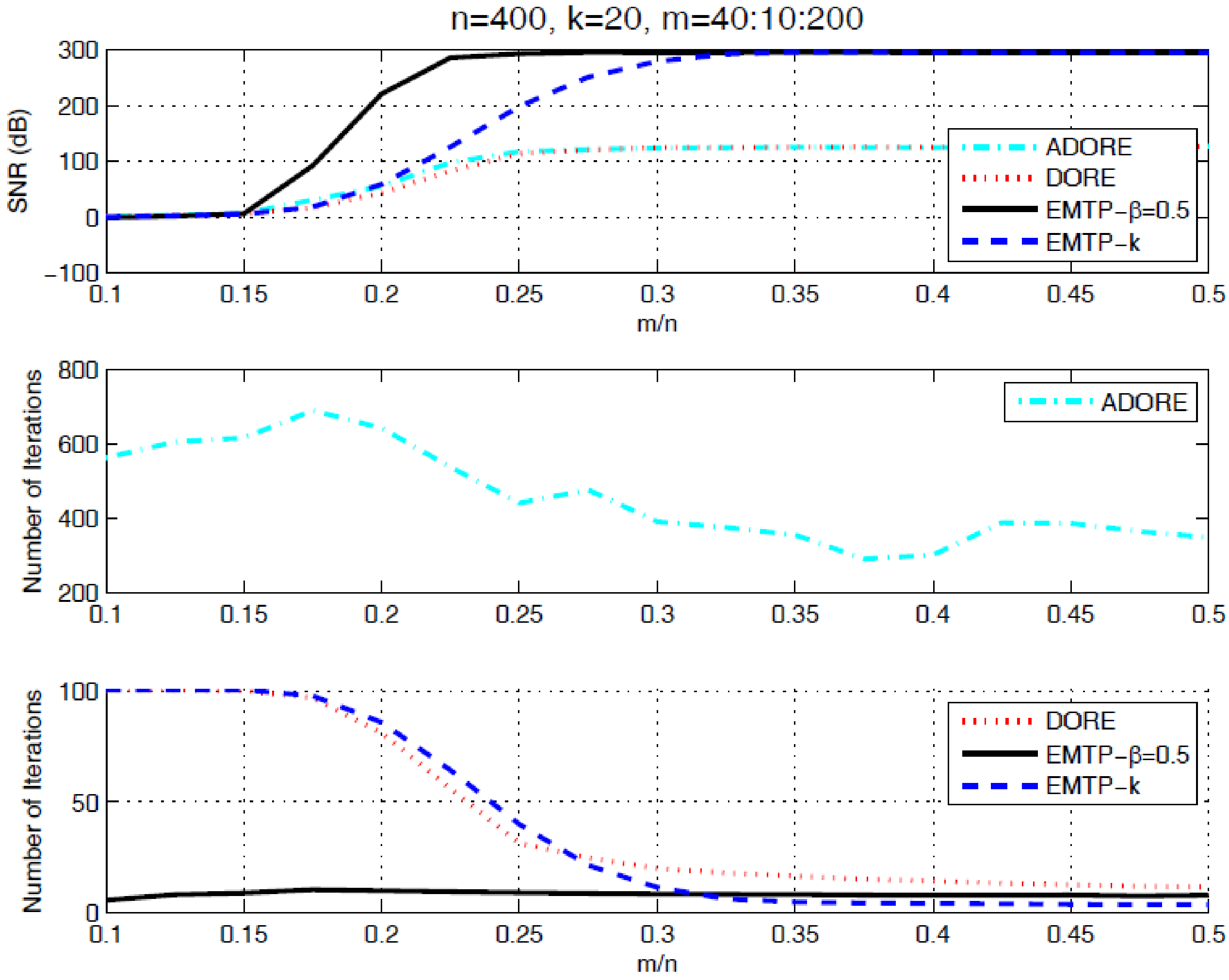
Sparse Gaussian signal

Comparisons-EMTP(4)



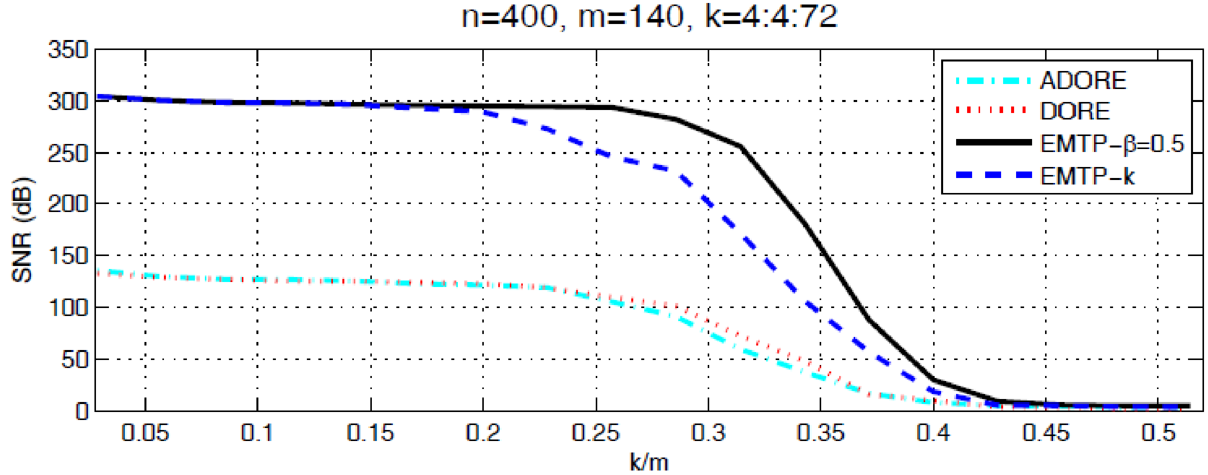
Sparse Gaussian signal

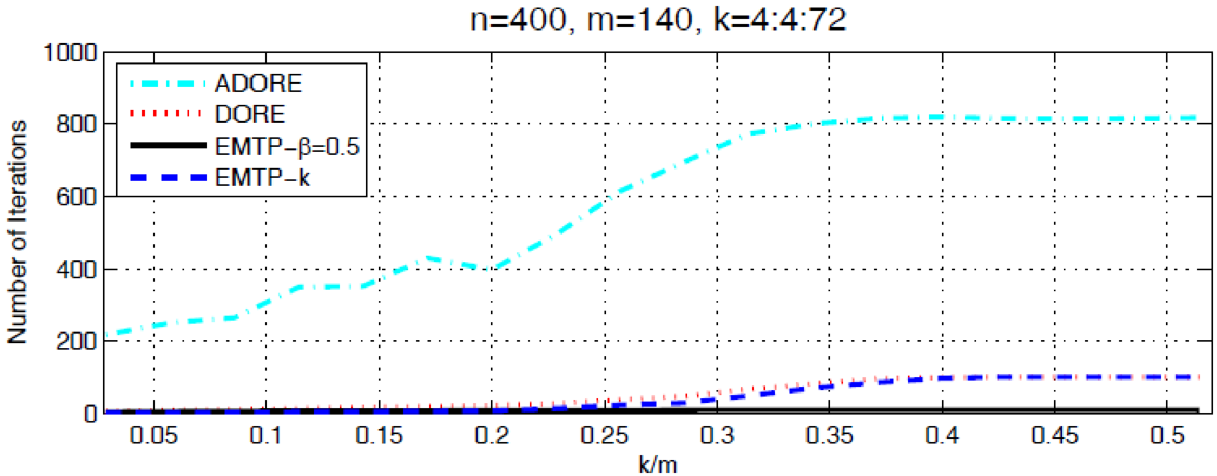
Comparisons-EMTP(5)



Sparse Gaussian signal

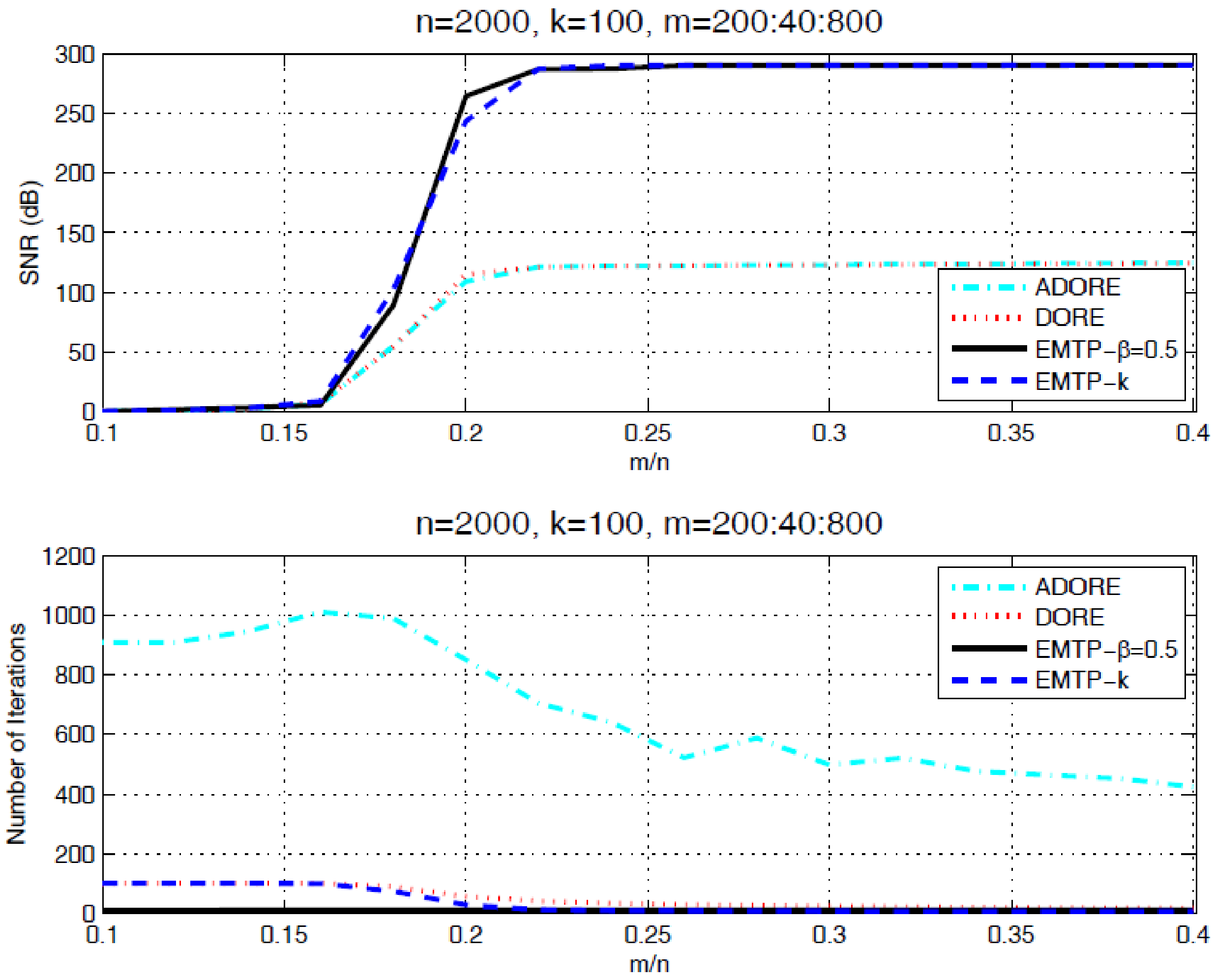
Comparisons-EMTP(6)





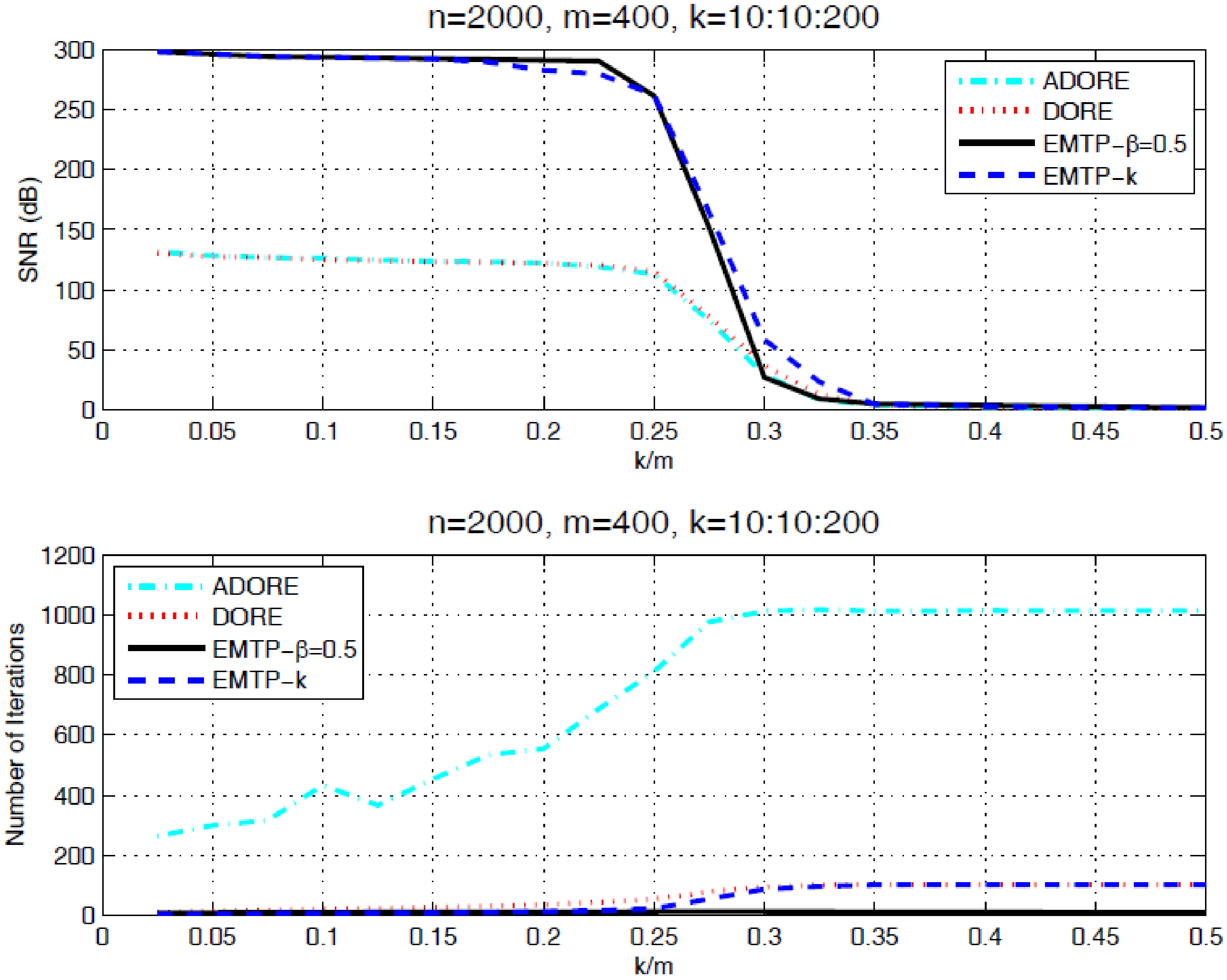
Sparse Gaussian signal

Comparisons-EMTP(7)



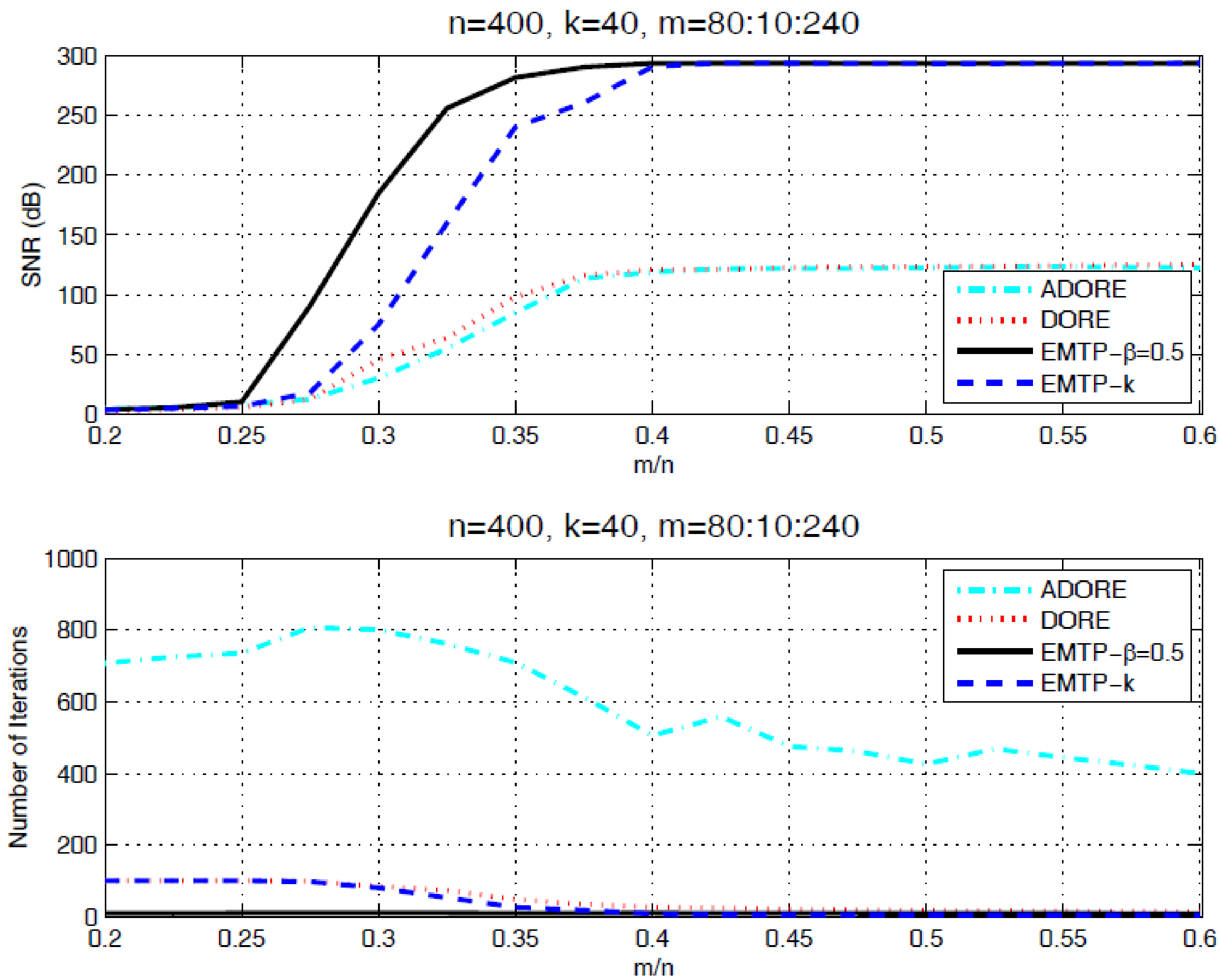
Sparse Gaussian signal

Comparisons-EMTP(8)



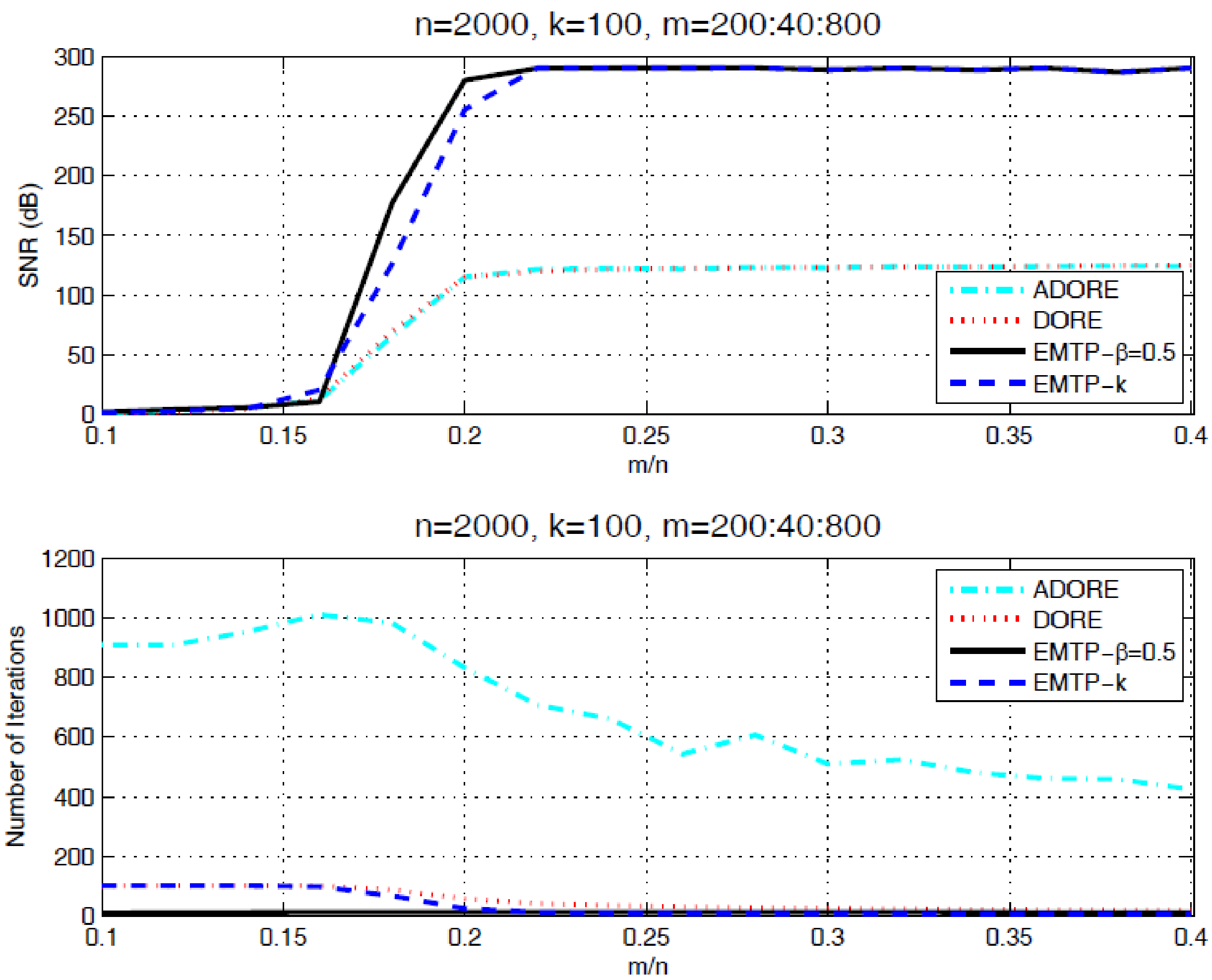
Sparse Gaussian signal

Comparisons-EMTP(9)



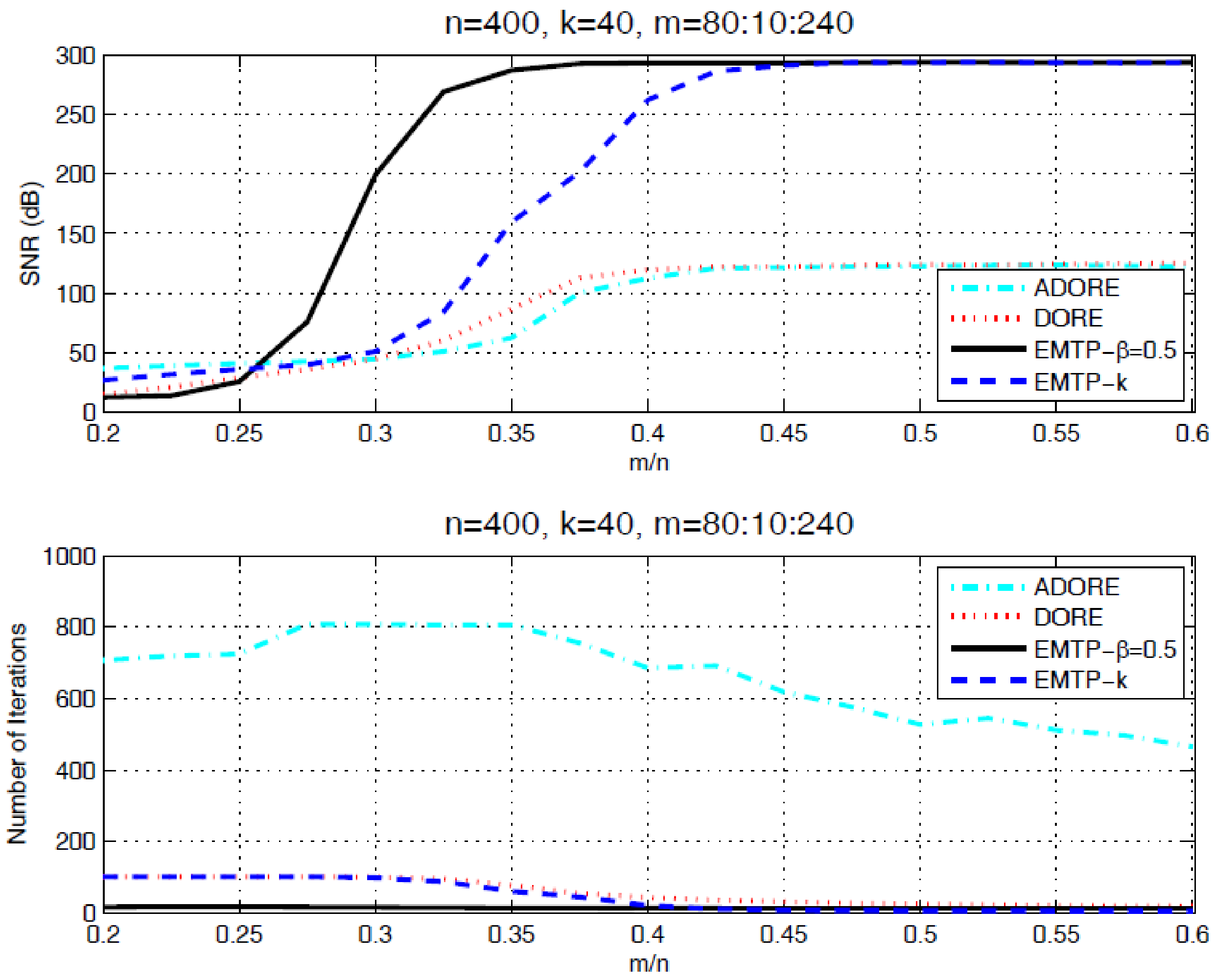
Sparse Laplacian signal

Comparisons-EMTP(10)



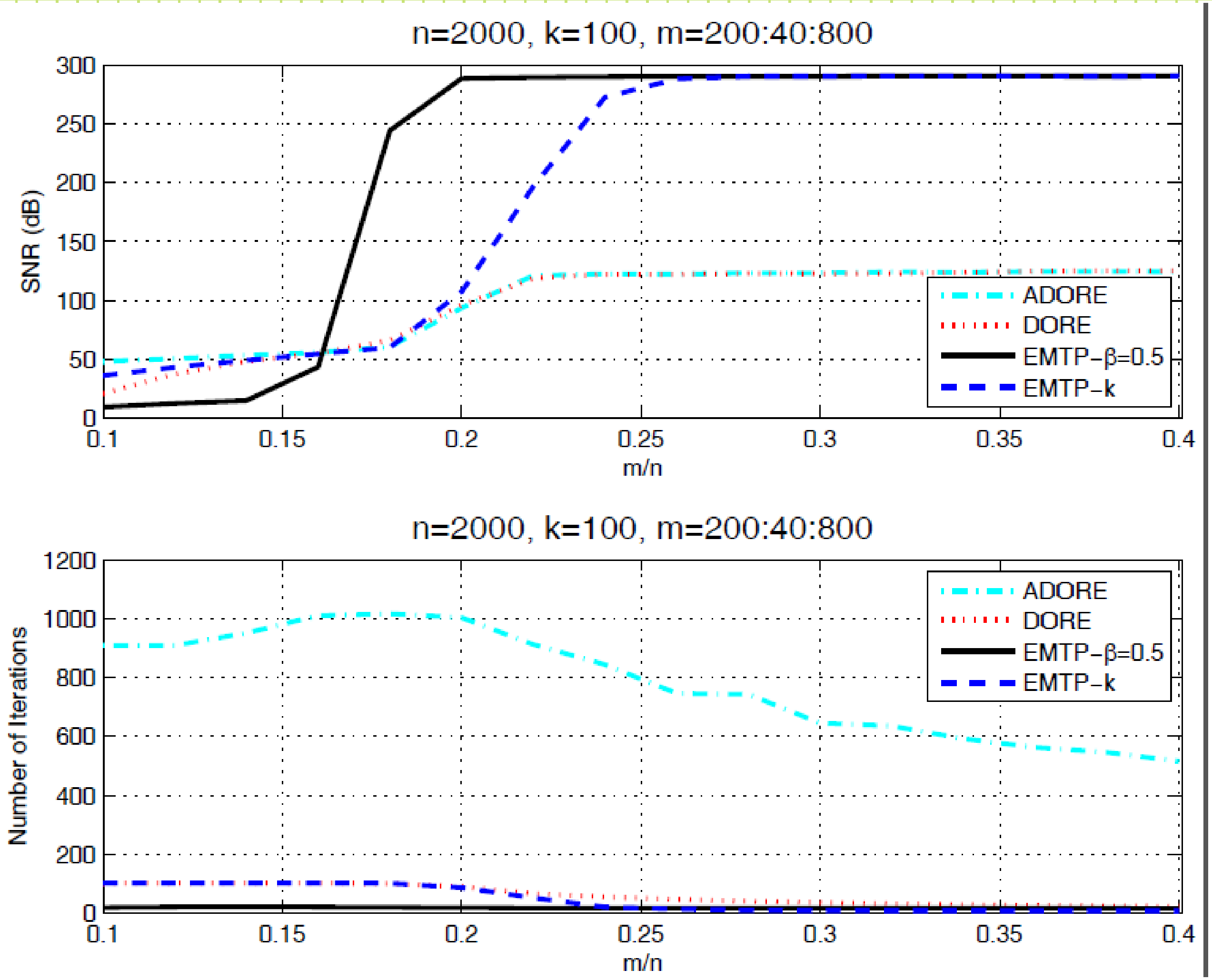
Sparse Laplacian signal

Comparisons-EMTP(11)



Power-law decaying signal

Comparisons-EMTP(12)



Power-law decaying signal

Outline

- Introduction
- RMP
- EMTP
- Demos
- Experiments
- Discussions

Discussions(1)

- Better recovery

  --Computational cost
- True signals

  --Fast decaying signals
- Thresholding methods

  --Ranking methods
- Why greedy

  --Statistics, communications, coding

Discussions(2)

- More numerical experiments

  --noisy dataset, images …
- Phase transition

  --IHT/ECME
- Orthogonal projection

  --Gradient pursuits
- Theoretical analysis

  --Truncated null space property

Q&A

Thank You!