Joint Channel Identification and Estimation in Wireless Network: Sparsity and Optimization

Thang Van Nguyen, Tony Q. S. Quek, and Hyundong Shin

IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, VOL. 17, NO. 5, PP. 3141-3153, MAY 2018

- To acquire UL CSI, the length of pilots is no less than number of users
- Sparse network: a huge number of users are in idle
- Using ADMM for CS-based problems
- Contributions:
  - Blind CS algorithm, joint user identification and channel estimation (unknown sparsity level)
  - Cluster sparsity and ADMM based algorithm (known sparsity)

- Blind Signal Recovery

$$\boldsymbol{y} = \sum_{k=1}^{N} x_k \boldsymbol{\phi}_k + \boldsymbol{w} = \boldsymbol{\Phi}\boldsymbol{x} + \boldsymbol{w}$$

$$\mathbf{x}^{\mathrm{CS}} = \underset{\mathbf{z}\in\mathbb{C}^N,\ \|\mathbf{z}\|_{\ell_0}\leq k}{\arg\min} \|\mathbf{y} - \boldsymbol{\Phi}\mathbf{z}\|_{\ell_2}$$

- $m = \mathcal{O}(k \log \frac{N}{k})$

- Sequential sampling technique without sparsity knowledge

$$\hat{\mathbf{x}}^{[m]} = \underset{\mathbf{z}\in\mathbb{C}^N}{\arg\min} \|\mathbf{z}\|_{\ell_1} \quad \text{s.t. } \|\mathbf{y}^{[m]} - \boldsymbol{\Phi}^{[m]}\mathbf{z}\|_{\ell_2} \leq \epsilon.$$

Algorithm 1 Blind CS Signal Recovery

Input: Measurement matrix $\boldsymbol{\Phi}$, size of signal vector N, measurement error ϵ, stopping threshold ν, and constant C.

0. Initialize $k = 1$, $m = \lceil C \log N \rceil$, and $\eta > \nu$. Solve (7) to obtain $\hat{\mathbf{x}}^{[m]}$.

While $\eta > \nu$ and $m < N$:

1. Update $m \leftarrow m + 1$.
2. Collect another measurement sample to update $\mathbf{y}^{[m]}$ and $\boldsymbol{\Phi}^{[m]}$. Compute $\hat{\mathbf{x}}^{[m]}$ using (7).
3. Update $\eta = \|\hat{\mathbf{x}}^{[m]} - \hat{\mathbf{x}}^{[m-1]}\|_{\ell_2}$.

end

Output: $\hat{\mathbf{x}} = \hat{\mathbf{x}}^{[m]}$.

D. M. Malioutov, S. R. Sanghavi, and A. S. Willsky, “Sequential compressed sensing,” IEEE J. Sel. Topics Signal Process., vol. 4, no. 2, pp. 435–444, Apr. 2010.

- Blind Signal Recovery

Proposition 1: Let $\{\hat{\mathbf{x}}^{[m]}\}_{m=1,2,\ldots}$ be a sequence of solutions of (7) in Algorithm 1. Then, the sequence $\{\|\hat{\mathbf{x}}^{[m]}\|_{\ell_1}\}_{m=1,2,\ldots}$ is a non-decreasing sequence and upper bounded by $\|\mathbf{x}\|_{\ell_1}$.

Proof: For all m, it follows from (7) that $\|\hat{\mathbf{x}}^{[m]}\|_{\ell_1} \leq \|\mathbf{x}\|_{\ell_1}$ since $\|\mathbf{e}\|_{\ell_2} \leq \epsilon$. Furthermore, we have

$$\epsilon \geq \|\mathbf{y}^{[m+1]} - \mathbf{\Phi}^{[m+1]}\hat{\mathbf{x}}^{[m+1]}\|_{\ell_2} \tag{8}$$

$$\geq \|\mathbf{y}^{[m]} - \mathbf{\Phi}^{[m]}\hat{\mathbf{x}}^{[m+1]}\|_{\ell_2} \tag{9}$$

- Guarantee the convergence
- Drawbacks: (1) unknown sparsity (2) high computation

- Blind Signal Recovery

Algorithm 2 Refined Blind CS Signal Recovery

Input: Measurement matrix $\mathbf{\Phi}$, size of signal vector N, measurement error ϵ, stopping threshold ν, and constant C.

0. Initialize $k = 1$, $m_k = \lceil C \log N \rceil$. Solve (7) to obtain $\hat{\mathbf{x}}^{[m_k]}$.

While $\|\mathbf{y}^{[m_k]} - \mathbf{\Phi}^{[m_k]}\mathcal{H}_k\{\hat{\mathbf{x}}^{[m_k]}\}\|_{\ell_2} > \nu$ and $m_k < N$:

1. Update $k \leftarrow k + 1$. Compute $m_k = \lceil Ck \log \frac{N}{k} \rceil$.
2. Collect $m_k - m_{k-1}$ measurement samples to update $\mathbf{y}^{[m_k]}$ and $\mathbf{\Phi}^{[m_k]}$. Compute $\hat{\mathbf{x}}^{[m_k]}$ using (7).

end

Output: k, $\hat{\mathbf{x}} = \mathcal{H}_k\{\hat{\mathbf{x}}^{[m_k]}\}$.

- $\mathcal{H}_k(\boldsymbol{a})$: take k entries with largest amplitudes of vector $\boldsymbol{a}$, remaining entries are set to zero

- An CS algorithm is consistent if it can recover exactly the positions of non-zero elements in the signal vector.

- Blind Signal Recovery

Proposition 2: Let $\mathbf{x}$ be a s-sparse signal vector in Algorithm 2. For each sparsity level $k \in \{1, 2, \ldots, s\}$, we assume that the measurement matrix $\mathbf{\Phi}^{[m_k]}$ satisfies the RIP condition for all $(k+s)$-sparse vectors with $\delta_{s+k} \in (0, 1)$. If $\delta_{2s} < \sqrt{2} - 1$, $\nu \geq \frac{5+(3-\sqrt{2})\delta_{2s}}{1-(1+\sqrt{2})\delta_{2s}}\epsilon$, and

$$\min_{|\mathbf{x}[i]|>0,\ i=1,2,\ldots,N} |\mathbf{x}[i]| > \frac{\epsilon + \nu}{\sqrt{1 - \max_{k=1,2,\ldots,s} \delta_{k+s}}},$$

then Algorithm 2 is consistent.

- RIP

$$(1 - \delta_{2k}) \|\mathbf{z}\|_{\ell_2}^2 \leq \|\mathbf{\Phi z}\|_{\ell_2}^2 \leq (1 + \delta_{2k}) \|\mathbf{z}\|_{\ell_2}^2$$

Blind Signal Recovery

Proof: We first show that Algorithm 2 must stop after s iterations. Indeed, for $k = s$, it follows from [12, Theorem 1.9] that

$$\|\hat{\mathbf{x}}^{[m_s]} - \mathbf{x}\|_{\ell_2} \leq 2\frac{1-(1-\sqrt{2})\,\delta_{2s}}{1-(1+\sqrt{2})\,\delta_{2s}}\frac{\sigma_s(\mathbf{x})}{\sqrt{s}} + 4\frac{\sqrt{1+\delta_{2s}}}{1-(1+\sqrt{2})\,\delta_{2s}}\epsilon \quad (10)$$

where $\sigma_s(\mathbf{x}) \triangleq \min_{\mathbf{u}\in\mathbb{C}^N,\ \|\mathbf{u}\|_{\ell_0}=s} \|\mathbf{u}-\mathbf{x}\|_{\ell_1} = 0$. Using (10), we have

$$\begin{aligned}
&\|\mathbf{y}^{[m_k]} - \mathbf{\Phi}^{[m_k]}\mathcal{H}_k\left\{\hat{\mathbf{x}}^{[m_k]}\right\}\|_{\ell_2} \\
&\leq \|\mathbf{y}^{[m_k]} - \mathbf{\Phi}^{[m_k]}\mathbf{x}\|_{\ell_2} + \|\mathbf{\Phi}^{[m_k]}\left(\mathcal{H}_k\left\{\hat{\mathbf{x}}^{[m_k]}\right\} - \mathbf{x}\right)\|_{\ell_2} && (11)\\
&\leq \epsilon + \sqrt{1+\delta_{2s}}\|\mathcal{H}_k\left\{\hat{\mathbf{x}}^{[m_k]}\right\} - \mathbf{x}\|_{\ell_2} && (12)\\
&\leq \epsilon + \sqrt{1+\delta_{2s}}\|\hat{\mathbf{x}}^{[m_k]} - \mathbf{x}\|_{\ell_2} && (13)\\
&\leq \nu && (14)
\end{aligned}$$

leading to the stopping of Algorithm 2. Now, we assume that Algorithm 2 stops after $k \leq s$ iterations, then using the RIP, we have

$$\begin{aligned}
&\sqrt{1-\delta_{s+k}}\|\mathcal{H}_k\left\{\hat{\mathbf{x}}^{[m_k]}\right\} - \mathbf{x}\|_{\ell_2} \\
&\leq \|\mathbf{\Phi}^{[m_k]}\left(\mathcal{H}_k\left\{\hat{\mathbf{x}}^{[m_k]}\right\} - \mathbf{x}\right)\|_{\ell_2} && (15)\\
&\leq \|\mathbf{\Phi}^{[m_k]}\mathcal{H}_k\left\{\hat{\mathbf{x}}^{[m_k]}\right\} - \mathbf{y}^{[m_k]}\|_{\ell_2} + \|\mathbf{\Phi}^{[m_k]}\mathbf{x} - \mathbf{y}^{[m_k]}\|_{\ell_2} && (16)\\
&\leq \epsilon + \nu. && (17)
\end{aligned}$$

Therefore, $\mathcal{S}\{\mathcal{H}_k\{\hat{\mathbf{x}}\}\} = \mathcal{S}\{\mathbf{x}\}$. From which, we complete the proof. □

Y. C. Eldar and G. Kutyniok, Compressed Sensing: Theory and Applications. Cambridge, U.K.: Cambridge Univ. Press, 2012.

Cluster Sparsity

- Signal vector can be partitioned into sub-vectors in which we can have prior information of the sparsity level in each cluster
- enhance the accuracy of signal recovery, reduce the number of measurement samples

$$\mathbf{x}^{\mathrm{C}} = \underset{\mathbf{z}\in\mathbb{C}^N,\ \|\mathbf{z}_i\|_{\ell_0}\le s_i,\ i=1,2,\ldots,L}{\arg\min} f_{\mathrm{C}}(\mathbf{z})$$

where

$$f_{\mathrm{C}}(\mathbf{z}) = \|\mathbf{y} - \mathbf{\Phi}\mathbf{z}\|_{\ell_2}^2$$

$$\mathbf{z} = \begin{bmatrix}\mathbf{z}_1^{\mathrm{T}} & \mathbf{z}_2^{\mathrm{T}} & \cdots & \mathbf{z}_L^{\mathrm{T}}\end{bmatrix}^{\mathrm{T}}$$

$$\mathbf{x}^{\mathrm{D}} = \underset{\mathbf{z}\in\mathbb{C}^N,\ \|\mathbf{z}_i\|_{\ell_0}\le s_i,\ i=1,2,\ldots,L}{\arg\min} \times \left\{ \min_{\mathbf{y}_i\in\mathbb{C}^m,\ i=1,2,\ldots,L} f_{\mathrm{D}}(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_L, \mathbf{z}) \right\}$$

$$\text{s.t.} \sum_{i=1}^{L} \mathbf{y}_i = \mathbf{y}$$

$$f_{\mathrm{D}}(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_L, \mathbf{z}) = \sum_{i=1}^{L} \|\mathbf{y}_i - \mathbf{\Phi}_i\mathbf{z}_i\|_{\ell_2}^2$$

Cluster Sparsity

Proposition 3: Let $\mathbf{x}^{\mathrm{C}}$ and $\mathbf{x}^{\mathrm{D}}$ be the optimal solutions of problems in (18) and (20), respectively. Let $\mathbf{y}_i^\star(\mathbf{z})$, $i = 1, 2, \ldots, L$, be the minimum of $f_{\mathrm{D}}(\mathbf{z}, \mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_L)$ for given $\mathbf{z}$, then we have $f_{\mathrm{C}}(\mathbf{x}^{\mathrm{C}}) = L f_{\mathrm{D}}(\mathbf{x}^{\mathrm{D}}, \mathbf{y}_1^\star(\mathbf{x}^{\mathrm{D}}), \mathbf{y}_2^\star(\mathbf{x}^{\mathrm{D}}), \ldots, \mathbf{y}_L^\star(\mathbf{x}^{\mathrm{D}}))$. Furthermore, if $\mathbf{x}^{\mathrm{C}}$ is the unique solution of (18), then $\mathbf{x}^{\mathrm{C}} = \mathbf{x}^{\mathrm{D}}$.

Proof: Using the Lagrangian approach, we compute

$$\mathcal{L} = \sum_{i=1}^{L} \|\mathbf{y}_i - \mathbf{\Phi}_i \mathbf{z}_i\|_{\ell_2}^2 + \Re\left\{\boldsymbol{\lambda}^{\mathrm{H}}\left(\sum_{i=1}^{L} \mathbf{y}_i - \mathbf{y}\right)\right\} \tag{22}$$

where $\boldsymbol{\lambda} \in \mathbb{C}^m$ is a Lagrange multiplier vector. By taking the partial differential of $\mathcal{L}$ with respect to $\mathbf{y}_i$, we obtain

$$\mathbf{y}_i^\star(\mathbf{z}) = \mathbf{\Phi}_i \mathbf{z}_i + \frac{1}{L}(\mathbf{y} - \mathbf{\Phi}\mathbf{z}). \tag{23}$$

From which, we can conclude that both problems in (18) and (20) are equivalent. □

Algorithm 3 Cluster Sparse Algorithm

Input: Measurement matrix $\mathbf{\Phi}$, received signal vector $\mathbf{y}$.
0. Initialize $\boldsymbol{z} = \boldsymbol{z}^0$, $l = 0$.
Repeat
1. Update $\boldsymbol{y}_i^{l+1}$ using the formula in (23).
2. Update $\boldsymbol{z}_i^{l+1}$ by solving the sub-problem:

$$\boldsymbol{z}_i^{l+1} = \underset{\boldsymbol{z} \in \mathbb{C}^{n_i}, \|\boldsymbol{z}\|_{\ell_0} \le s_i}{\arg\min} \|\boldsymbol{y}_i^{l+1} - \mathbf{\Phi}_i \boldsymbol{z}\|_{\ell_2}^2.$$

4. Update $l = l + 1$.
until satisfying the stopping criterion.
Output: $\mathbf{x}^{\mathrm{D}} = \boldsymbol{z}^l$.

ADMM with Cluster Sparsity

$$\boldsymbol{t} = \left[\mathbf{t}_1^{\mathrm{T}} \quad \mathbf{t}_2^{\mathrm{T}} \quad \cdots \quad \mathbf{t}_L^{\mathrm{T}}\right]^{\mathrm{T}} \quad \mathbf{t}_i = \mathbf{y}_i - \boldsymbol{\Phi}_i \mathbf{z}_i \quad f(\boldsymbol{z}) = 0 \quad g(\boldsymbol{t}) = \|\boldsymbol{t}\|_{\ell_2}^2 \quad \mathbf{Q} = \mathbf{1}_L^{\mathrm{T}} \otimes \mathbf{I}_m$$

- Optimization problem

$$\min_{\substack{\boldsymbol{z}\in\mathbb{C}^N, \boldsymbol{t}\in\mathbb{C}^{mL} \\ \|\mathbf{z}_i\|_{\ell_0} \le s_i,\ i=1,2,\ldots,L}} f(\boldsymbol{z}) + g(\boldsymbol{t})$$

$$\text{s.t. } \boldsymbol{\Phi}\boldsymbol{z} + \mathbf{Q}\boldsymbol{t} = \mathbf{y}$$

- Lagrangian function

$$\begin{aligned} \mathcal{L}_\rho(\boldsymbol{z}, \boldsymbol{t}, \boldsymbol{\lambda}) = f(\boldsymbol{z}) + g(\boldsymbol{t}) + \Re\mathfrak{e}\left\{\boldsymbol{\lambda}^{\mathrm{H}}\left(\boldsymbol{\Phi}\boldsymbol{z} + \mathbf{Q}\boldsymbol{t} - \mathbf{y}\right)\right\} \\ + \frac{\rho}{2}\|\boldsymbol{\Phi}\boldsymbol{z} + \mathbf{Q}\boldsymbol{t} - \mathbf{y}\|_{\ell_2}^2 \end{aligned}$$

- replacing

$$f(\boldsymbol{z}) = \sum_{i=1}^{L} f_i\left(\|\boldsymbol{z}_i\|_{\ell_0}\right) \qquad f_i(x) = \begin{cases} 0, & 0 \le x \le s_i \\ \infty, & x > s_i. \end{cases}$$

ADMM with Cluster Sparsity

Theorem 1: Let $f(\boldsymbol{z})$ and $g(\boldsymbol{t})$ be convex functions. Assume that $(\boldsymbol{z}^\star, \boldsymbol{t}^\star)$ is the the solution of the following problem

$$\min_{\boldsymbol{z}\in\mathbb{C}^N, \boldsymbol{t}\in\mathbb{C}^{mL}} \left\{ f(\boldsymbol{z}) + g(\boldsymbol{t}) \right\}$$
$$\text{s.t. } \boldsymbol{\Phi}\boldsymbol{z} + \mathbf{Q}\boldsymbol{t} = \mathbf{y} \qquad (31)$$

and $\tau \|\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{\Phi}\|_\infty < 1$. Then, if $f(\boldsymbol{z})$ and $g(\boldsymbol{t})$ have subdifferential and continuous differential for all $\boldsymbol{z} \in \mathbb{C}^N$ and $\boldsymbol{t} \in \mathbb{C}^{mL}$, respectively, such that $\partial g(\boldsymbol{t}) + \rho \mathbf{Q}^{\mathrm{H}}\mathbf{Q}\boldsymbol{t}$ is one-to-one mapping. Then, Algorithm 4 always converges and yields a solution $(\boldsymbol{z}^l, \boldsymbol{t}^l)$ such that $f(\boldsymbol{z}^l) + g(\boldsymbol{t}^l)$ approaches to $f(\boldsymbol{z}^\star) + g(\boldsymbol{t}^\star)$.

Algorithm 4 Sparse ADMM Algorithm

Input: Measurement matrix $\boldsymbol{\Phi}$, received signal vector $\mathbf{y}$.

0. Initialize ρ, $\boldsymbol{t} = \boldsymbol{t}^0$, $\boldsymbol{\lambda} = \boldsymbol{\lambda}^0$, $l = 0$.

Repeat

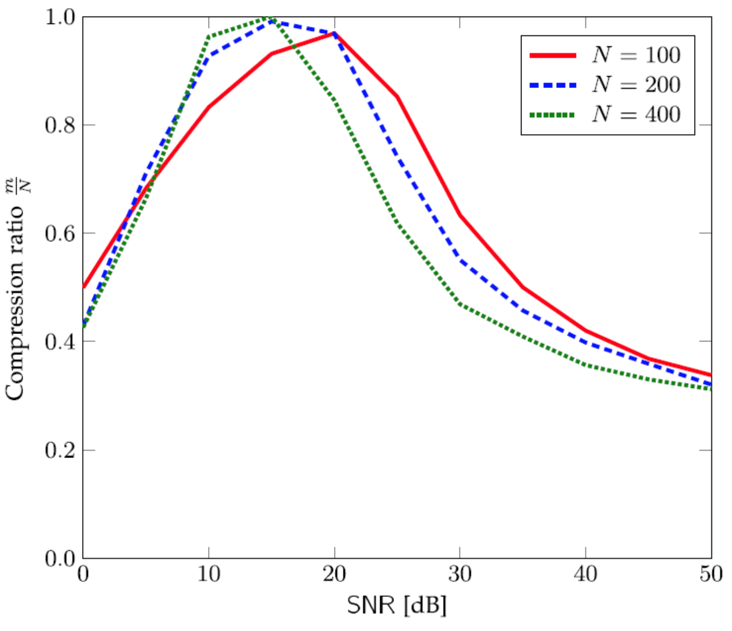
1. Update $\boldsymbol{z}^{l+1} = \arg\min_{\boldsymbol{z}\in\mathbb{C}^N} \mathcal{L}_\rho(\boldsymbol{z}, \boldsymbol{t}^l, \boldsymbol{\lambda}^l)$.
2. Update $\boldsymbol{t}^{l+1} = \arg\min_{\boldsymbol{t}\in\mathbb{C}^{mL}} \mathcal{L}_\rho(\boldsymbol{z}^{l+1}, \boldsymbol{t}, \boldsymbol{\lambda}^l)$.
3. Update $\boldsymbol{\lambda}^{l+1} = \boldsymbol{\lambda}^l + \rho\left(\boldsymbol{\Phi}\boldsymbol{z}^{l+1} + \mathbf{Q}\boldsymbol{t}^{l+1} - \mathbf{y}\right)$.
4. Update $l = l + 1$.

until satisfying the stopping criterion.

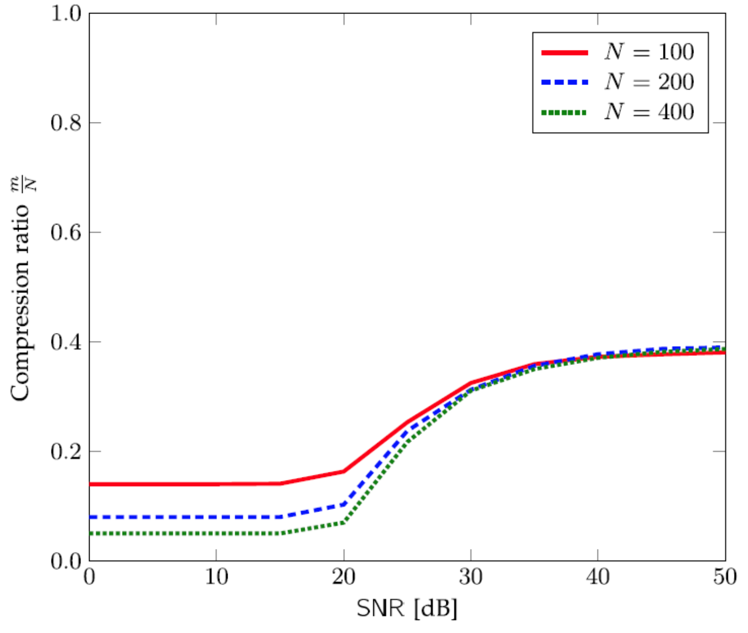
Output: $\mathbf{x}^{\mathrm{D}} = \boldsymbol{z}^l$.

J. Yang and Y. Zhang, "Primal and dual alternating direction algorithms for l1-l1-norm minimization problems in compressive sensing," Comput. Optim. Appl., vol. 54, no. 2, pp. 441–459, 2013.

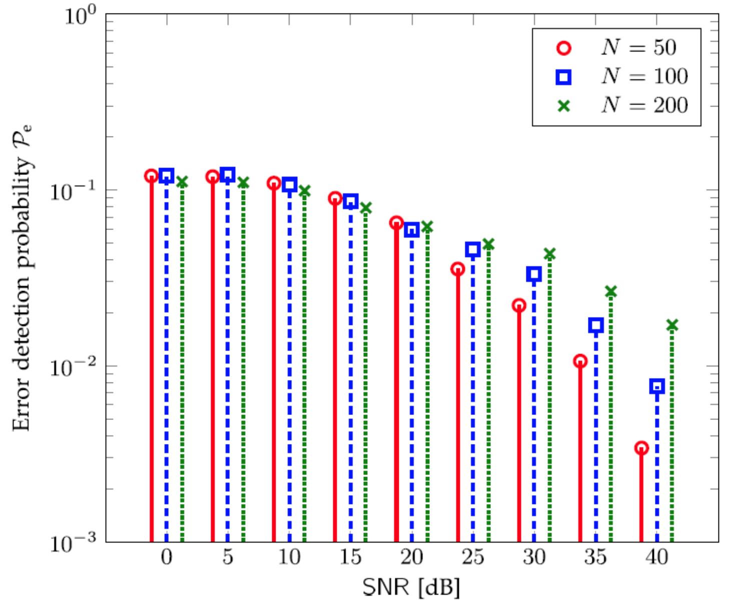
Simulations



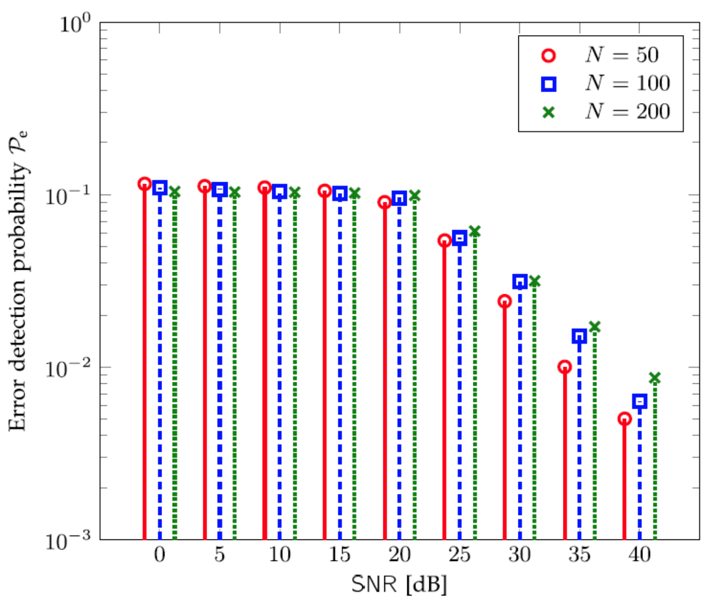
(a) Blind CS signal recovery algorithm



(b) Refined blind CS signal recovery algorithm



(a) Blind CS signal recovery algorithm



(b) Refined blind CS signal recovery algorithm

Simulations