RECOVERY OF SPARSE SIGNALS VIA BRANCH AND BOUND LEAST-SQUARES

TEXAS
The University of Texas at Austin

Abolfazl Hashemi and Haris Vikalo
Department of Electrical and Computer Engineering, University of Texas at Austin

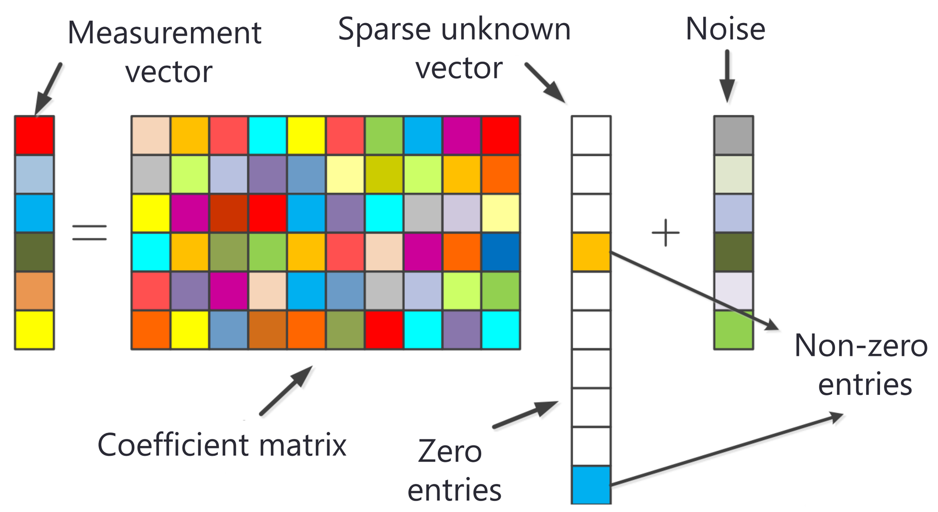
The University of Texas at Austin
Electrical and Computer Engineering
Cockrell School of Engineering

BACKGROUND

- Sparse Linear Regression:

$$\mathbf{y} = \mathbf{A}\boldsymbol{x} + \boldsymbol{\nu}$$

$$\underset{\boldsymbol{x}}{\text{minimize}} \quad \|\mathbf{y} - \mathbf{A}\boldsymbol{x}\|_2^2 \qquad \text{subject to} \quad \|\boldsymbol{x}\|_0 \le k$$



Existing Methods

- Convex Relaxation (e.g., LASSO)
- Tree Based (e.g., MMP)
- Greedy (e.g., OMP, OLS)

Tradeoff between Speed and accuracy

- OLS selection Criterion and update rule:

$$j_s = \arg\max_{j \in \mathcal{I}\backslash\mathcal{S}_{i-1}} \left| \mathbf{r}_{\mathcal{S}_{i-1}}^\top \frac{\mathbf{P}^\perp_{\mathcal{S}_{i-1}}\mathbf{a}_j}{\|\mathbf{P}^\perp_{\mathcal{S}_{i-1}}\mathbf{a}_j\|_2} \right| \qquad \mathbf{P}^\perp_{\mathcal{S}_i} = \mathbf{P}^\perp_{\mathcal{S}_{i-1}} - \frac{\mathbf{P}^\perp_{\mathcal{S}_{i-1}}\mathbf{a}_{j_s}\mathbf{a}_{j_s}^\top\mathbf{P}^\perp_{\mathcal{S}_{i-1}}}{\|\mathbf{P}^\perp_{\mathcal{S}_{i-1}}\mathbf{a}_{j_s}\|_2^2}$$

- Accelerated OLS (AOLS):

$$\mathbf{a}_j \leftarrow \mathbf{a}_j - \sum_{l=1}^{i} \frac{\mathbf{a}_j^\top\mathbf{u}_l}{\|\mathbf{u}_l\|_2^2}\mathbf{u}_l$$

$$\mathbf{q}_j = (\mathbf{a}_j^\top\mathbf{r}_i/\|\mathbf{a}_j\|_2^2)\,\mathbf{a}_j$$

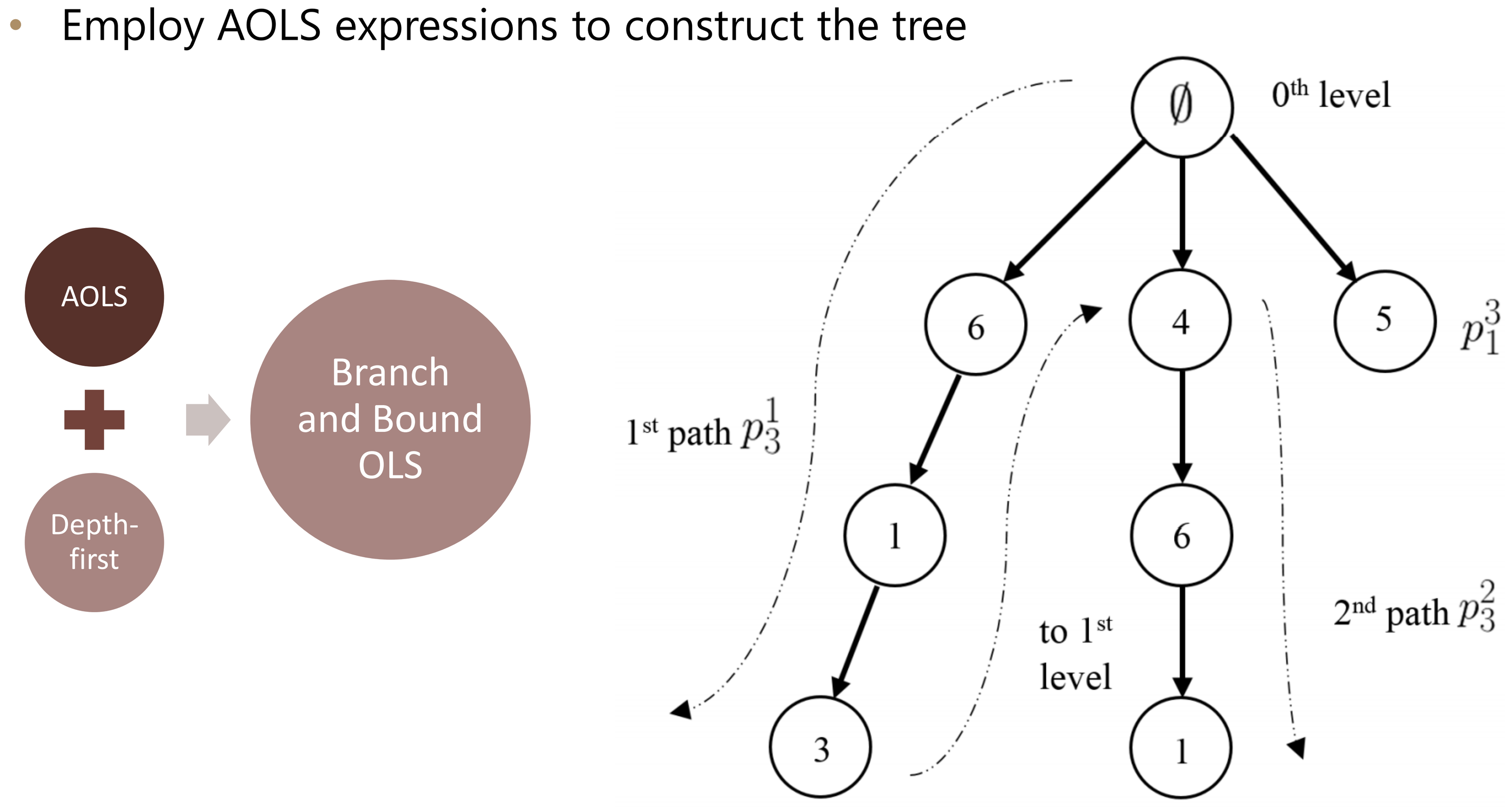
$$j_s = \text{argmax}_{j\in\mathcal{I}\backslash\mathcal{S}_i} \|\mathbf{q}_j\|_2$$

$$\mathbf{r}_{\mathcal{S}_i} = \mathbf{r}_{\mathcal{S}_{i-1}} - \mathbf{u}_{i+1}$$

$$\mathbf{u}_{i+1} = \mathbf{q}_{j_s}$$

BRANCH AND BOUND LEAST-SQUARES

- Construct a tree whose nodes represent columns of coefficient matrix
- A branch-and-bound search to traverse the tree in a depth-first manner
- Use a schedule $\mathbf{L} = [L_1, \ldots, L_k]$ to control the size of the search space
- Employ AOLS expressions to construct the tree



BRANCH AND BOUND LEAST-SQUARES

Algorithm 1 Branch and Bound Least-Squares (BBLS)

Input: $\mathbf{y}$, $\mathbf{A}$, sparsity level k, threshold ϵ, schedule $\mathbf{L}$, max number of paths N_p

Output: recovered support $\hat{\mathcal{S}}$, estimated signal $\hat{\boldsymbol{x}}$

1. (Initialize) $\mathcal{S} = \emptyset$, $\mathbf{r}_{p_0^\ell} = \mathbf{y}$, $r_{\ell_2} = \|\mathbf{y}\|_2$, $i = 1$, $\ell = 1$.
2. (Bounding) Let $\mathbf{S}_i = [\,]$ and $l_i = 0$,

for $j \in \mathcal{I}\backslash\mathcal{S}$ **do**

$$\mathbf{a}_j \leftarrow \mathbf{a}_j - \mathbb{I}(i>2)\frac{\mathbf{a}_j^\top\mathbf{u}_l}{\|\mathbf{u}_l\|_2^2}\mathbf{u}_l, \quad \mathbf{q}_j = \frac{\mathbf{a}_j^\top\mathbf{r}_{i-1}}{\|\mathbf{a}_j\|_2^2}\mathbf{a}_j$$

end for

Select $\mathbf{S}_i = [j_{s_1}, \ldots, j_{s_{L_i}}]$ corresponding to L_i largest terms $\|\mathbf{q}_j\|_2$

3. (Branching) $l_i = l_i + 1$. If $l_i > L_i$ go to 4, else $\mathcal{S} = \mathcal{S} \cup \{\mathbf{S}_i(l_i)\}$, $\mathbf{u}_i = \mathbf{q}_{j_{s_{l_i}}}$, $\mathbf{r}_{p_i^\ell} = \mathbf{r}_{p_{i-1}^\ell} - \mathbf{u}_i$, go to 5.
4. (Decrease i) If $i = 1$ go to 7, else $\mathcal{S} = \mathcal{S}\backslash\{\mathbf{S}_i(l_i)\}$, $i = i-1$, and go to 2.
5. (Increase i) If $i = k$ go to 6, else $i = i+1$ and go to 2.
6. (Solution found) Save the ℓ^{th} path $p_k^\ell = \mathcal{S}$ and its objective value $\left\|\mathbf{r}_{p_k^\ell}\right\|_2$. If $\left\|\mathbf{r}_{p_k^\ell}\right\|_2 < r_{\ell_2}$ update $r_{\ell_2} = \left\|\mathbf{r}_{p_k^\ell}\right\|_2$. $\ell = \ell+1$, if $\ell > N_p$ or $r_{\ell_2} < \epsilon$ go to 7, else go to 3.
7. Terminate the algorithm. Return the path $p_k^{\ell^*}$ with minimum residual norm as $\hat{\mathcal{S}}$, and the estimate $\hat{\boldsymbol{x}} = \mathbf{A}_{\hat{\mathcal{S}}}^\dagger\mathbf{y}$.

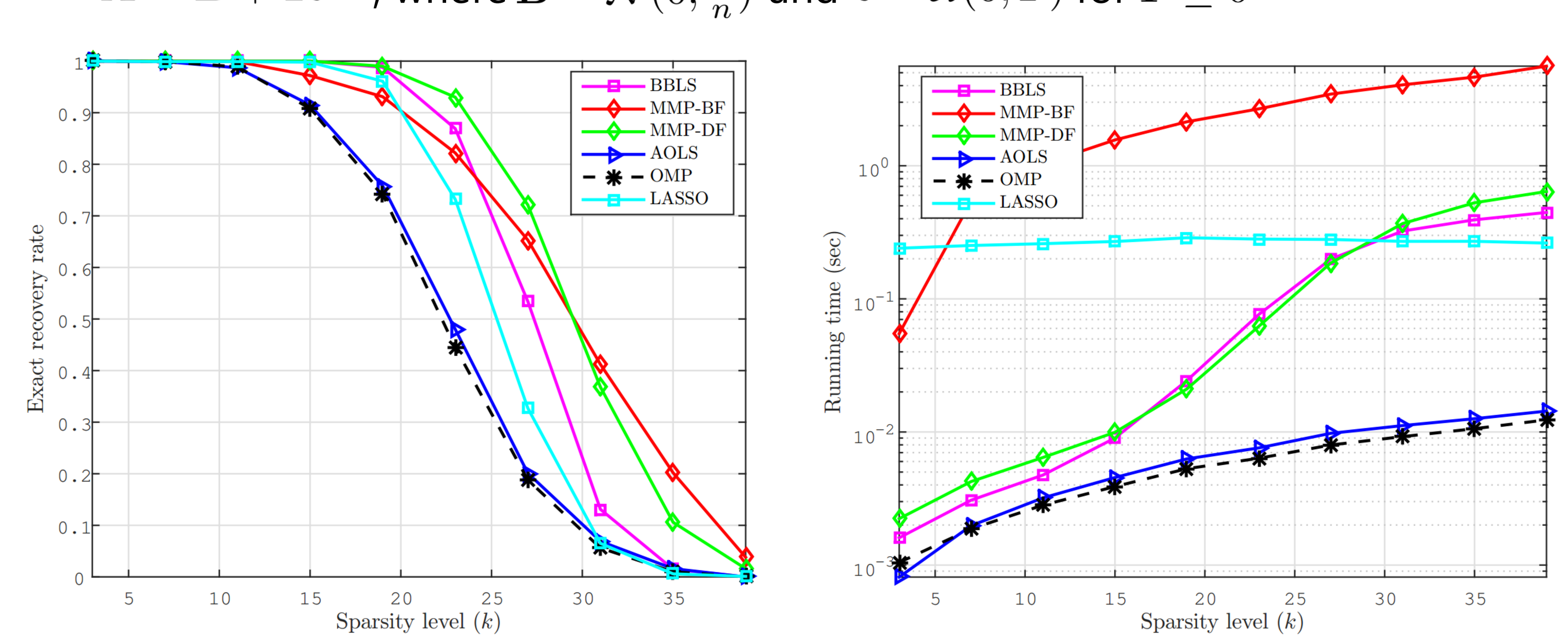
PERFORMANCE ANALYSIS

- Let $0 < \epsilon < 1$ and $0 < \delta < 1$ be universal constants, and $c_0(\epsilon) = \frac{\epsilon^2}{6}(1-\epsilon)$
- Assume $\mathbf{A} \sim \mathcal{N}(0, 1/n)$ or $\mathbf{A} \sim \mathcal{B}(\frac{1}{2}, \pm\frac{1}{\sqrt{n}})$, and noiseless measurements
- If $p_i^\ell = \{s_1^\ell, \ldots, s_i^\ell\} \subset \mathcal{S}_{true}$, then, at least one among L_{i+1} children of s_i^ℓ is in $\mathcal{S}_{true}$ with probability

$$p \ge \left(1 - 2e^{-(n-i)c_0(\epsilon)}\right)^2 \left(1 - 2(\frac{12}{\delta})^k e^{-nc_0(\frac{\delta}{2})}\right)\left(1 - 2e^{-\frac{n}{k-i}\frac{1-\epsilon}{1+\epsilon}(1-\delta)^2}\right)^{m-k-L_{i+1}+1}$$

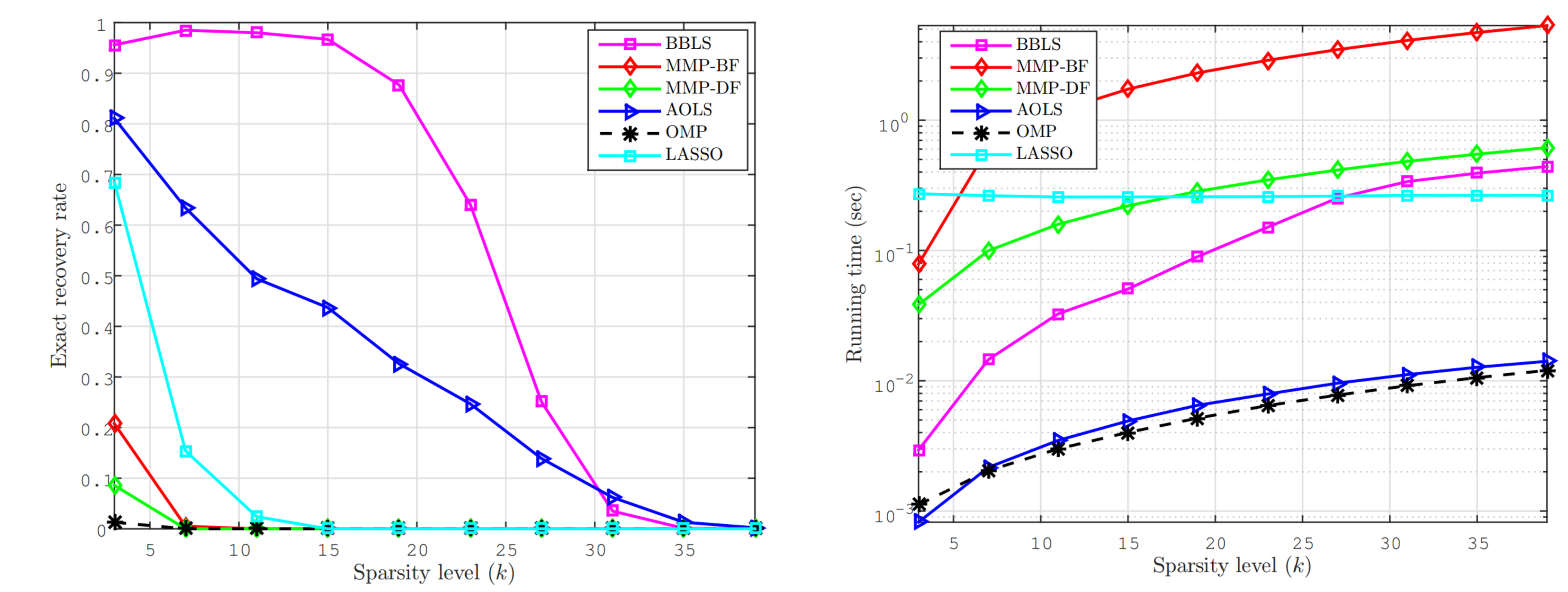
SIMULATION RESULTS

- n = 64, m = 128, $\boldsymbol{x} \sim \mathcal{N}(0,1)$, vary k for 1000 independent instances
- $\mathbf{A} = \mathbf{B} + \mathbf{1}\mathbf{t}^\top$, where $\mathbf{B} \sim \mathcal{N}(0, \frac{1}{n})$ and $\mathbf{t} \sim \mathcal{U}(0, T)$ for $T \ge 0$

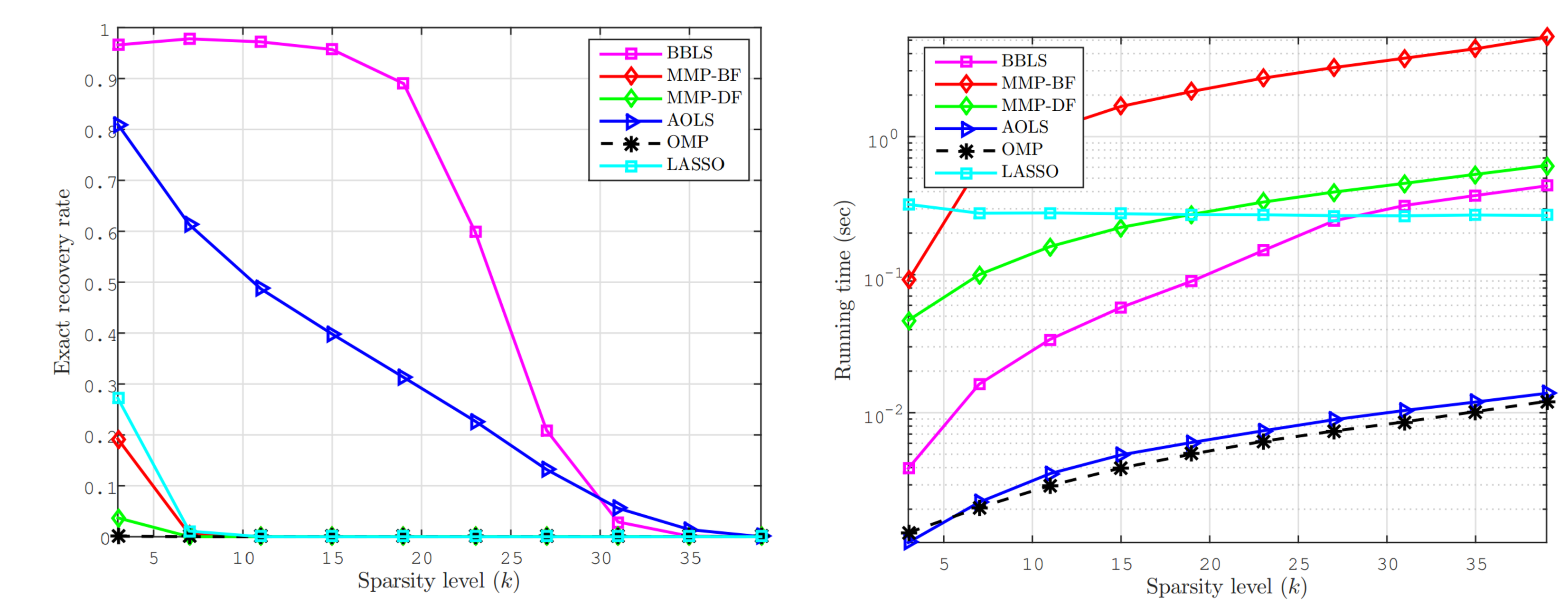


Comparison results for incoherent matrix (T=0)

SIMULATION RESULTS



Comparison results for coherent matrix (T=10)



Comparison results for highly coherent matrix (T=100)

CONCLUSIONS

The proposed algorithm, BBLS:

- is a depth-first search scheme for sparse reconstruction.
- selects different number of indices in each level according to a schedule.
- has guarantees for its achievable reconstruction probability.
- is capable of highly accurate recovery even for correlated dictionaries.

Future work: performance analysis under hybrid dictionaries.

REFERENCES

[1] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans.* Inf. Theory, vol. 53, no. 12, pp. 4655–4666, 2007

[2] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans.* Inf. Theory, vol. 51, no. 12, pp. 4203–4215, 2005

[3] Abolfazl Hashemi and Haris Vikalo, "Sparse linear regression via generalized orthogonal least-squares," *arXiv preprint, arXiv:1602.06916*, 2016.

[4] S. Kwon, et al. "Multipath matching pursuit," *IEEE Trans.* Inf. Theory, vol. 60, no. 5, pp. 2986–3001, 2014.