Support Recovery for Orthogonal Matching Pursuit: 
Upper and Lower bounds

Raghav Somani ¹, Chirag Gupta ², Prateek Jain ¹ & Praneeth Netrapalli ¹

¹Microsoft Research Lab - India

²Machine Learning Department, Carnegie Mellon University

December 6, 2018
Sparse Linear Regression (SLR)

\[ \bar{x} = \arg \min_{\|x\|_0 \leq s^*} \|Ax - y\|_2 \]

- Unconditionally, NP hard.
- Tractable under the assumption of Restricted Strong Convexity (RSC).
- Fundamental quantity capturing hardness:
  - Standard optimization: Condition number
    \[ \kappa = \text{smoothness} / \text{strong convexity} \]
  - Sparse optimization: Restricted Condition number
    \[ \tilde{\kappa} = \text{restricted smoothness} / \text{restricted strong convexity} \]
Sparse Linear Regression (SLR)

\[ \bar{x} = \arg \min_{\|x\|_0 \leq s^*} \|Ax - y\|_2^2 \]

- **Unconditionally, NP hard. 😞**
- **Tractable under the assumption of Restricted Strong Convexity (RSC). 😊**
- **Fundamental quantity capturing hardness:**
  - Standard optimization: Condition number
    - Smoothness
    - Strong convexity
  - Sparse optimization: Restricted Condition number
    - Restricted smoothness
    - Restricted strong convexity
Sparse Linear Regression (SLR)

\[ \bar{x} = \arg \min_{\|x\|_0 \leq s^*} \| A x - y \|_2^2 \]

- Unconditionally, NP hard. 😞
- Tractable under the assumption of Restricted Strong Convexity (RSC). 😊
- Fundamental quantity capturing hardness:
  - Standard optimization: Condition number
  - Sparse optimization: Restricted Condition number

R. Somani, C. Gupta, P. Jain & P. Netrapalli
Sparse Linear Regression (SLR)

Unconditionally, NP hard. 😞

Tractable under the assumption of Restricted Strong Convexity (RSC). 😊

Fundamental quantity capturing hardness:
- Standard optimization: Condition number

$$\kappa = \frac{\text{smoothness}}{\text{strong convexity}}$$

- Sparse optimization: Restricted Condition number

$$\tilde{\kappa} = \frac{\text{restricted smoothness}}{\text{restricted strong convexity}}$$
Setup and Goals

We work under the model where

\[ y = A \bar{x} + \eta \]

- Observations
- Measurement matrix
- \( s^* \)-sparse vector
- Noise

Goals of SLR

1. Bounding Generalization error/Excess Risk: \( G(x) = \frac{1}{n} \| A (x - \bar{x}) \|_2 \)
2. Support Recovery: Recover the support of \( \bar{x} \)

We study SLR under RSC assumption for OMP.
**Setup and Goals**

We work under the model where

\[ y = A \bar{x} + \eta \]

- **Observations**
- Measurement matrix
- \( s^* \)-sparse vector
- Noise

**Goals of SLR**

1. **Bounding Generalization error/Excess Risk**
   
   \[ G(x) = \frac{1}{n} \| A(x - \bar{x}) \|_2 \]

2. **Support Recovery**
   
   Recover the support of \( \bar{x} \)

We study SLR under RSC assumption for OMP.
Setup and Goals

We work under the model where

- Observations
- Measurement matrix
- $s^*$-sparse vector
- Noise

\[ y = A \bar{x} + \eta \]

Goals of SLR

1. Bounding Generalization error/Excess Risk - $\mathcal{O}(n) = \frac{1}{n} \| A (x - \bar{x}) \|_2$
2. Support Recovery - Recover the support of $\bar{x}$

We study SLR under RSC assumption for OMP.
Setup and Goals

We work under the model where

- Observations
- Measurement matrix
- $s^*$-sparse vector
- Noise

\[ y = A \bar{x} + \eta \]

Goals of SLR

1. Bounding Generalization error/Excess Risk
   \[ G(x) = \frac{1}{n} \|A(x - \bar{x})\|_2 \]

2. Support Recovery
   - Recover the support of $\bar{x}$

We study SLR under RSC assumption for OMP.
Setup and Goals

We work under the model where

\[ y = A \bar{x} + \eta \]

**Observations** \( y \)  
**Measurement matrix** \( A \)  
**\( s^* \)-sparse vector** \( \bar{x} \)  
**Noise** \( \eta \)

**Goals of SLR**

1. **Bounding Generalization error/Excess Risk** - \( G(x) := \frac{1}{n} \| A(x - \bar{x}) \|_2^2 \).
2. **Support Recovery** - Recover the support of \( \bar{x} \).

We study SLR under RSC assumption for OMP.
Setup and Goals

We work under the model where

\[ y = A \bar{x} + \eta \]

- Observations
- Measurement matrix
- \( s^* \)-sparse vector
- Noise

Goals of SLR

1. **Bounding Generalization error/Excess Risk** - \( G(x) := \frac{1}{n} \| A(x - \bar{x}) \|_2^2 \).

2. **Support Recovery** - Recover the support of \( \bar{x} \).

We study SLR under RSC assumption for OMP.
Setup and Goals

We work under the model where

- Observations
- Measurement matrix
- $s^*$-sparse vector
- Noise

\[ y = A \bar{x} + \eta \]

Goals of SLR

1. **Bounding Generalization error/Excess Risk** - $G(x) := \frac{1}{n} \left\| A(x - \bar{x}) \right\|_2^2$.

2. **Support Recovery** - Recover the support of $\bar{x}$.

We study SLR under RSC assumption for OMP.
Orthogonal Matching Pursuit

- Incremental Greedy algorithm
- Popular and easy to implement
- Widely studied in literature

\[ \hat{x}_{k+1} \text{ iteratively selected greedily} \]
## Known results and our contribution

### Upper bound

<table>
<thead>
<tr>
<th>Known Generalization bound</th>
<th>( \propto \frac{1}{n} \sigma^2 s^* \tilde{\kappa}^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Our Generalization bound</td>
<td>( \propto \frac{1}{n} \sigma^2 s^* \tilde{\kappa} \log \tilde{\kappa} )</td>
</tr>
</tbody>
</table>

### Lower bound

<table>
<thead>
<tr>
<th>Known Generalization bound</th>
<th>( \propto \frac{1}{n} \sigma^2 s^* \tilde{\kappa} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Our Generalization bound</td>
<td>( \propto \frac{1}{n} \sigma^2 s^* \tilde{\kappa} )</td>
</tr>
</tbody>
</table>

- **Support Expansion**
  - Known \( \propto s^* \tilde{\kappa}^2 \)
  - Our's \( \propto s^* \tilde{\kappa} \log \tilde{\kappa} \)

Unconditional lower bounds for OMP.
Support recovery guarantees and its lower bounds.
Known results and our contribution

<table>
<thead>
<tr>
<th>Upper bound</th>
<th>Lower bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \frac{1}{n} \sigma^2 s^* \tilde{\kappa}^2 )</td>
<td>( \frac{1}{n} \sigma^2 s^* \tilde{\kappa} )</td>
</tr>
</tbody>
</table>

Known Generalization bound \( \propto \)

Our Generalization bound \( \propto \)

- Unconditional lower bounds for OMP.
- Support recovery guarantees and its lower bounds.

Support Expansion

Knowing \( \propto \)

Our’s \( \propto \)

\( s^* \tilde{\kappa}^2 \)

\( s^* \tilde{\kappa} \log \tilde{\kappa} \)
## Known results and our contribution

<table>
<thead>
<tr>
<th></th>
<th>Upper bound</th>
<th>Lower bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Known Generalization bound</td>
<td>$\propto \frac{1}{n} \sigma^2 s^* \tilde{\kappa}^2$</td>
<td>$\propto \frac{1}{n} \sigma^2 s^* \tilde{\kappa}$</td>
</tr>
<tr>
<td>Our Generalization bound</td>
<td>$\propto \frac{1}{n} \sigma^2 s^* \tilde{\kappa} \log \tilde{\kappa}$</td>
<td>$\propto \frac{1}{n} \sigma^2 s^* \tilde{\kappa}$</td>
</tr>
</tbody>
</table>

- Unconditional lower bounds for OMP.
- Support recovery guarantees and its lower bounds.

### Support Expansion

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Known</td>
<td>$s^* \tilde{\kappa}^2$</td>
</tr>
<tr>
<td>Our’s</td>
<td>$s^* \tilde{\kappa} \log \tilde{\kappa}$</td>
</tr>
</tbody>
</table>
A key idea

\[ f(x) = \|Ax - y\|_2^2 \]

- If any support is unrecovered, then there is a large additive decrease.
- \( f(x) \geq 0 \implies \) support recovery will happen soon.
- Recovery with small support \( \implies \) small generalization error.
A key idea

\[ f(x) = \|Ax - y\|^2 \]

- If *any* support is unrecovered, then there is a *large additive decrease*.
- \( f(x) \geq 0 \implies \) support recovery will happen soon.
- Recovery with small support \( \implies \) small generalization error.
A key idea

\[ f(x) = \|Ax - y\|_2^2 \]

- If *any* support is unrecovered, then there is a *large additive decrease*.
- \( f(x) \geq 0 \implies \) support recovery will happen soon.
- Recovery with small support \( \implies \) small generalization error.
A key idea

\[ f(x) = \| Ax - y \|_2^2 \]

- If *any* support is unrecovered, then there is a *large additive decrease*.
- \( f(x) \geq 0 \implies \) support recovery will happen soon.
- Recovery with small support \( \implies \) small generalization error.
Thank You!