Average-Case Analysis of OMP

Joel A. Tropp

jtropp@umich.edu

Department of Mathematics
The University of Michigan

With contributions from
Anna C. Gilbert and Martin Strauss

Identification of Sparse Signals

Suppose that we measure a signal

$$\boldsymbol{s} = \boldsymbol{\Phi} \boldsymbol{c}_{\mathrm{opt}} + \boldsymbol{\nu}$$

- where $\boldsymbol{\Phi}$ is a known matrix
- $\boldsymbol{c}_{\mathrm{opt}}$ is an unknown sparse coefficient vector
- $\boldsymbol{\nu}$ is an unknown noise vector

Goal:

Find a sparse coefficient vector $\widehat{\boldsymbol{c}}$ that approximates $\boldsymbol{c}_{\mathrm{opt}}$. In particular, $\widehat{\boldsymbol{c}}$ should correctly identify the support of $\boldsymbol{c}_{\mathrm{opt}}$.

Model I: Random Noise

Suppose that we measure a signal

$$\boldsymbol{s} = \boldsymbol{\Phi} \boldsymbol{c}_{\text{opt}} + \boldsymbol{\nu}$$

where the noise vector $\boldsymbol{\nu}$ is random

- Goal: Recover a sparse superposition contaminated with additive noise
- Intuition: It is easy to reject the noise, which is unlikely to look like any column of the matrix
- Related work: [Candès et al., Donoho, Elad, Fletcher et al., . . .]

Model II: Random Matrix

Suppose that we measure a signal

$$\boldsymbol{s} = \boldsymbol{\Phi} \boldsymbol{c}_{\text{opt}} + \boldsymbol{\nu}$$

where the matrix $\boldsymbol{\Phi}$ is random

- Goal: Recover a sparse signal from imperfect random measurements
- Intuition: The information content of a sparse signal is preserved under random projections
- Related work: [Alon et al., Candès et al., Donoho, Gilbert et al., Johnson–Lindenstrauss, Nowak, . . .]

Model III: Random Coefficients

Suppose that we measure a signal

$$\boldsymbol{s} = \boldsymbol{\Phi} \boldsymbol{c}_{\text{opt}} + \boldsymbol{\nu}$$

where the coefficient vector $\boldsymbol{c}_{\text{opt}}$ is sparse and random

- Goal: Recover a random sparse superposition contaminated with noise
- Intuition: It is unlikely that a random superposition looks like another column of the matrix
- Related work: [Candès et al., Donoho, Elad–Zibulevsky, . . .]

Orthogonal Matching Pursuit (OMP)

Input: A matrix $\mathbf{\Phi}$, an input signal $\boldsymbol{s}$, a stopping criterion

Initialize the residual $\boldsymbol{r}_0 = \boldsymbol{s}$ and the counter $t = 0$

Until the stopping criterion holds do increment t and

A. Find the column index ω_t that solves

$$\omega_t \;=\; \arg\max_j \; |\langle \boldsymbol{r}_{t-1},\; \boldsymbol{\varphi}_j \rangle|$$

B. Calculate the next residual

$$\boldsymbol{r}_t = \boldsymbol{v} - \boldsymbol{P}_t\, \boldsymbol{s}$$

where $\boldsymbol{P}_t$ is the orthogonal projector onto $\operatorname{span}\{\boldsymbol{\varphi}_{\omega_1}, \dots, \boldsymbol{\varphi}_{\omega_t}\}$

Output: An sparse estimate $\widehat{\boldsymbol{c}}$ with nonzero entries in components $\omega_1, \dots, \omega_t$. These entries appear in the expansion

$$\boldsymbol{P}_t\, \boldsymbol{s} \;=\; \sum_{k=1}^{t} \widehat{c}_{\omega_k}\, \boldsymbol{\varphi}_{\omega_k}$$

Worst-Case Performance of OMP

Assumptions:

- We measure $\boldsymbol{s} = \boldsymbol{\Phi} \boldsymbol{c}_{\text{opt}} + \boldsymbol{\nu}$
- The matrix $\boldsymbol{\Phi}$ has coherence μ
- The sparsity m of $\boldsymbol{c}_{\text{opt}}$ satisfies $\mu\, m \leq \frac{1}{3}$
- $c_{\min}$ is the smallest nonzero component of $\boldsymbol{c}_{\text{opt}}$
- The norm of the noise $\|\boldsymbol{\nu}\|_2 \leq 0.25\, c_{\min}$
- The OMP halting criterion is $\|\boldsymbol{\Phi}^* \boldsymbol{r}_t\|_\infty \leq 0.5\, c_{\min}$

Theorem 1. [TGS] *The support of $\widehat{c}$ equals the support of $\boldsymbol{c}_{\text{opt}}$.*

OMP with Random Noise

Assumptions:

- We measure $\boldsymbol{s} = \boldsymbol{\Phi} \boldsymbol{c}_{\text{opt}} + \boldsymbol{\nu}$
- The matrix $\boldsymbol{\Phi}$ has coherence μ and N columns
- The sparsity m of $\boldsymbol{c}_{\text{opt}}$ satisfies $\mu\, m \le \frac{1}{3}$
- $c_{\min}$ is the smallest nonzero component of $\boldsymbol{c}_{\text{opt}}$
- The noise $\boldsymbol{\nu}$ is white Gaussian with variance σ^2
- The OMP halting criterion is $\|\boldsymbol{\Phi}^* \boldsymbol{r}_t\|_\infty \le 0.472\, c_{\min}$

Theorem 2. [JAT] *The support of $\widehat{\boldsymbol{c}}$ equals the support of $\boldsymbol{c}_{\text{opt}}$ with probability at least $(1-\delta)$ provided that*

$$\sigma < \frac{0.167}{\ln^{1/2}(N/\delta)}\, c_{\min}.$$

Worst-Case versus Average-Case Noise

- We measure $\boldsymbol{s} = \boldsymbol{\Phi}\boldsymbol{c}_{\text{opt}} + \boldsymbol{\nu}$
- The matrix $\boldsymbol{\Phi}$ is $2^{10} \times 2^{12}$ with coherence $\mu = 2^{-5}$
- The sparsity $m = 10$ and $c_{\min} = 1$

- For the worst-case noise, success requires an SNR around 22.04 dB
- For Gaussian noise, 99% success probability when SNR is 6.57 dB

OMP with Random Matrix

Assumptions:

- The matrix $\mathbf{\Phi}$ is $d \times N$ with iid $\textsc{normal}(0, 1)$ entries
- The vector $\boldsymbol{c}_{\text{opt}}$ has sparsity m
- We measure $\boldsymbol{s} = \mathbf{\Phi}\boldsymbol{c}_{\text{opt}}$

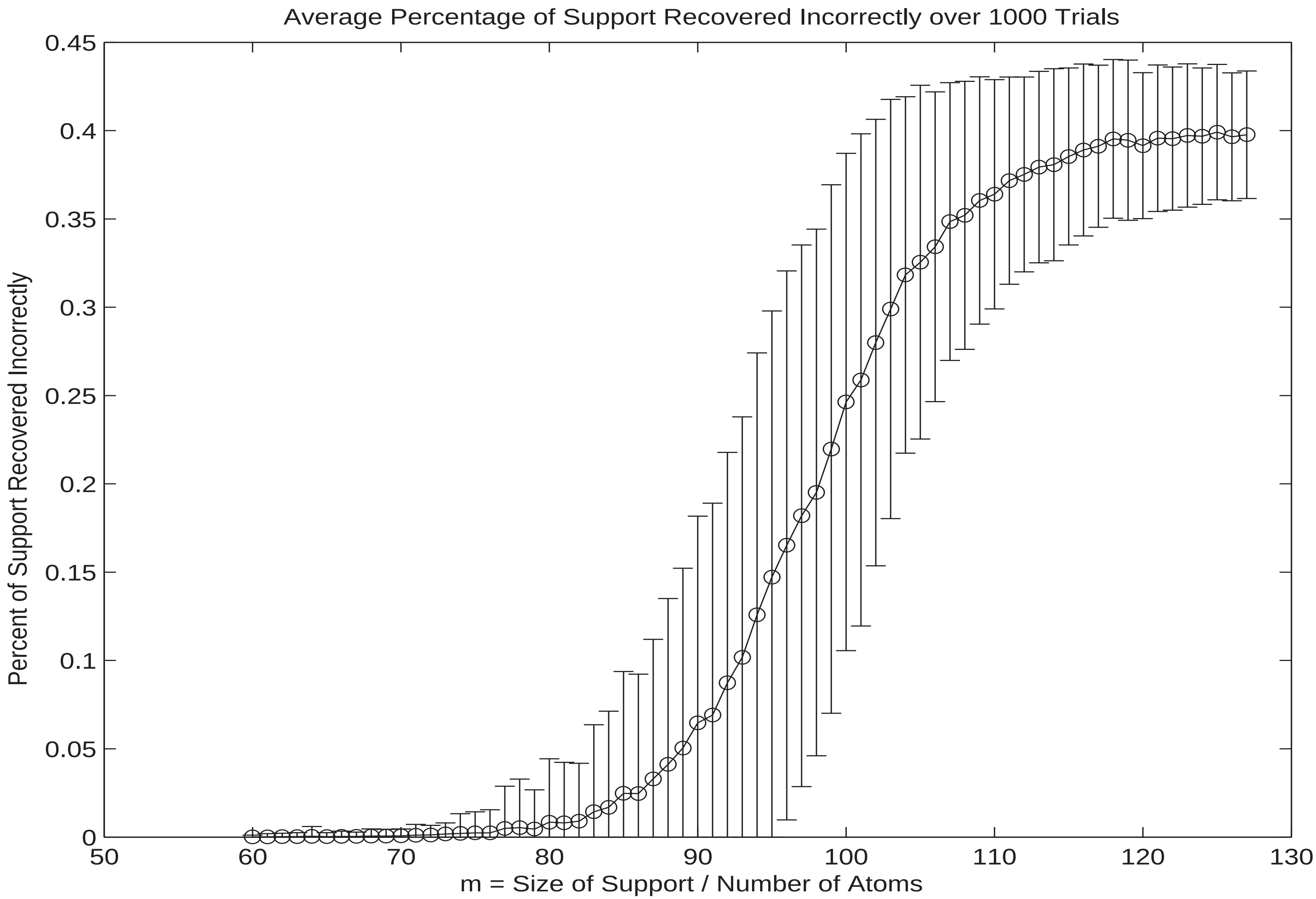
Theorem 3. [JAT–ACG] *Fix a positive number p. The probability that $\widehat{\boldsymbol{c}} = \boldsymbol{c}_{\text{opt}}$ exceeds $(1 - 2N^{-p})$ provided that*

$$m \leq \frac{d}{8\,(p+1)\ln N}.$$

To ensure recovery for general $\mathbf{\Phi}$, $m = O(\sqrt{d})$

OMP with Random Coefficients

- Let $\mathbf{\Phi} = \begin{bmatrix} \mathbf{I} & \mathscr{F} \end{bmatrix}^{128 \times 256}$
- The vector $\boldsymbol{c}_{\text{opt}}$ is m-sparse
- The location of the nonzero entries of $\boldsymbol{c}_{\text{opt}}$ are random
- The nonzero entries of $\boldsymbol{c}_{\text{opt}}$ are iid $\textsc{normal}(0, 1)$
- We measure $\boldsymbol{s} = \mathbf{\Phi}\boldsymbol{c}_{\text{opt}}$
- Theory requires $m \leq 6$ to ensure exact recovery
- But...


Average Percentage of Support Recovered Incorrectly over 1000 Trials
Percent of Support Recovered Incorrectly
m = Size of Support / Number of Atoms
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
50
60
70
80
90
100
110
120
130

Related Papers and Contact Information

- "Signal recovery from partial information via Orthogonal Matching Pursuit," submitted April 2005
- "Algorithms for simultaneous sparse approximation. Parts I and II," accepted to *EURASIP J. Signal Processing*, April 2005
- "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Info. Theory*, October 2004
- "Just Relax: Convex programming methods for identifying sparse signals," submitted February 2004

All papers available from `http://www.umich.edu/~jtropp`

E-mail: `jtropp@umich.edu`