Approximate Support Recovery of Atomic Line Spectral Estimation

Qiuwei Li & **Gongguo Tang**

EECS, Colorado School of Mines

SIgnals and NEtworks

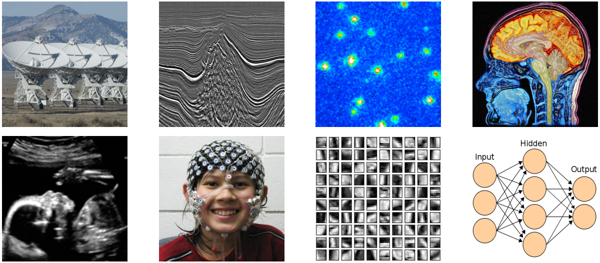
Motivations and Basic Ideas

Signal Model: Superposition of parameterized atoms/building-blocks

$$\mathbf{x} = \sum_{k=1}^{r} c_k \mathbf{a}(\boldsymbol{\theta}_k)$$

Different atoms correspond to different applications

- Radar/Seismology/Microscope/MRI
- Ultrasound imaging/Array proc./Dictionary learning/Neural networks



Line Spectral Estimation

- $\mathbf{a}(f) = [e^{j2\pi(-n)f}, e^{j2\pi(-n+1)f}, \ldots, e^{j2\pi(n-1)f}, e^{j2\pi nf}]^T, f \in [0, 1)$
  $2n+1$ equspaced samples of a normalized band limited complex sinusoid
- Sparse recovery problem with a continuous DFT dictionary $\mathcal{A} = \{\mathbf{a}(f)\}$
- **Infer** the frequencies of a mixture of r complex sinusoids in white noise

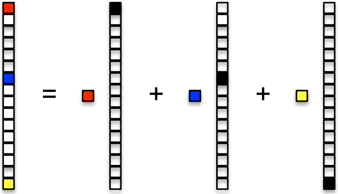
$$\mathbf{y} = \mathbf{x}^{\star} + \mathbf{w} = \sum_{k=1}^{r} c_k^{\star} \mathbf{a}(f_k^{\star}) + \mathbf{w}$$

Parameter estimation/Support recovery is very important here!!!

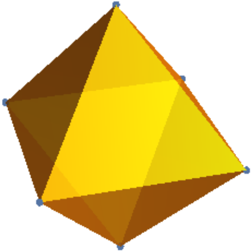
e.x., in radar, parameters correspond to the **position** *and* **velocity** *information of the targets!*

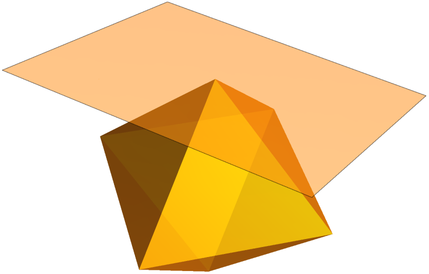
The Atomic Norm Minimization Algorithm I

In **compressed sensing,**
minimize $\|\mathbf{x}\|_1$ subject to $\mathbf{y} = A\mathbf{x}$
often produces a sparse solution.



A hyperplane will most likely touch the ℓ_1 norm ball at spiky points.





The Atomic Norm Minimization Algorithm II

✓ **Atomic norm:** $\|\mathbf{x}\|_{\mathcal{A}} = \inf\{\sum_k |c_k| : \mathbf{x} = \sum_k c_k \mathbf{a}(f_k)\}$ which has an equivalent SDP characterization.

1. Solve **ALASSO** for unique primal solution $\hat{\mathbf{x}}$:

$$\left(\hat{\mathbf{x}} = \sum_{k=1}^{\hat{r}} \hat{c}_k \mathbf{a}(\hat{f}_k)\right) = \operatorname{argmin} \frac{1}{2}\|\mathbf{x} - \mathbf{y}\|_W^2 + \lambda\|\mathbf{x}\|_{\mathcal{A}} \qquad \text{(ALASSO)}$$

   whose dual optimal solution is

$$\hat{\mathbf{q}} = \frac{W(\mathbf{y} - \hat{\mathbf{x}})}{\lambda}$$

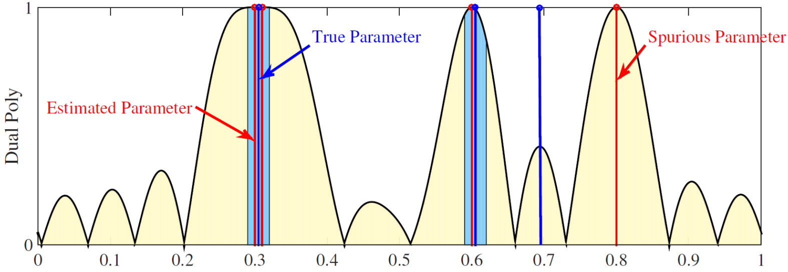
2. Extract the frequencies $\{\hat{f}_k\}$ by solving dual polynomial equation $|\hat{Q}(f)| := |\langle \hat{\mathbf{q}}, \mathbf{a}(f)\rangle| = 1$ as long as we can show the **BIP**

$$|\hat{Q}(f)| < 1, \; f \notin \{\hat{f}_k\} \qquad \text{(Boundedness)}$$

$$\hat{Q}(\hat{f}_k) = \operatorname{sign}(\hat{c}_k), k \in [\hat{r}] \qquad \text{(Interpolation)}$$

The Atomic Norm Minimization Algorithm III

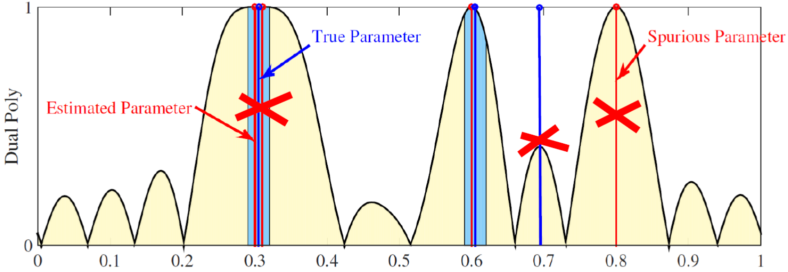
1. How well can **ALASSO** localize the frequencies?
2. Does **ALASSO** recovers exactly r frequencies?



V. Duval, G. Peyré. "Exact support recovery for sparse spikes deconvolution."
C. Fernandez-Granda. "Support detection in super-resolution."
G. Tang, B. Bhaskar, B. Recht. "Near minimax line spectral estimation."

The Atomic Norm Minimization Algorithm IV

1. Error bound matches **CRB**
2. Recover **exactly** r frequencies



Q. Li, G. Tang. "Approximate support recovery of atomic line spectral estimation: A tale of resolution and precision."

Main Contributions I

- Noise is Gaussian with variance σ^2
- Noise level is measured by $\gamma_0 := \sigma\sqrt{\frac{\log n}{n}}$

Theorem (Li & Tang, 2016)

Suppose

- *SNR as measured by* $|c_{\min}|/\gamma_0$ *is large.*
- *The dynamic range of the coefficients is small.*
- *Regularization parameter* λ *is large compared to* γ_0*.*
- *The frequencies are well-separated.*

Then w.h.p. we can extract **exactly** r *parameters from* $\hat{\mathbf{x}}$ *or* $\hat{\mathbf{q}}$*, which satisfy*

$$\max |c_k^\star||\hat{f}_k - f_k^\star| = O(\gamma_0/n) = O(\frac{\sqrt{\log n}}{n^{3/2}}\sigma)$$

$$\max |\hat{c}_k - c_k^\star| = O(\lambda) = O(\sqrt{\frac{\log n}{n}}\sigma)$$

Q. Li, G. Tang. "Approximate support recovery of atomic line spectral estimation: A tale of resolution and precision."

Main Contributions II

Comparison with CRB, MUSIC, and MLE.

- Only asymptotic bounds available when the snapshots number $T \to \infty$.
- **Our algorithm:** work for single snapshot, i.e., $T = 1$.

- CRB: $O(\frac{\sigma^2}{T|c|^2 n^3})$
- Atomic: $O(\frac{\sigma^2 \log n}{|c|^2 n^3})$
- MLE: $O(\frac{\sigma^2}{T|c|^2 n^3} + \frac{\sigma^4}{T|c|^4 n^4})$
- MUSIC: $O(\frac{\sigma^2}{T|c|^2 n^3} + \frac{\sigma^4}{T|c|^4 n^4})$

P. Stoica, A. Nehorai. "MUSIC, maximum likelihood, and Cramer-Rao bound."

Proof by Primal-Dual Witness Construction

- The unique primal optimal solution is $\hat{\mathbf{x}} = \sum_{k=1}^{\hat{r}} \hat{c}_k \mathbf{a}(\hat{f}_k)$.
- The unique dual optimal solution is $\hat{\mathbf{q}} = W(\mathbf{y} - \hat{\mathbf{x}})/\lambda$ satisfying **BIP**.
- They **certify the optimality of each other.**

1. Fix $\hat{r} = r$ and construct the primal candidate solution by solving

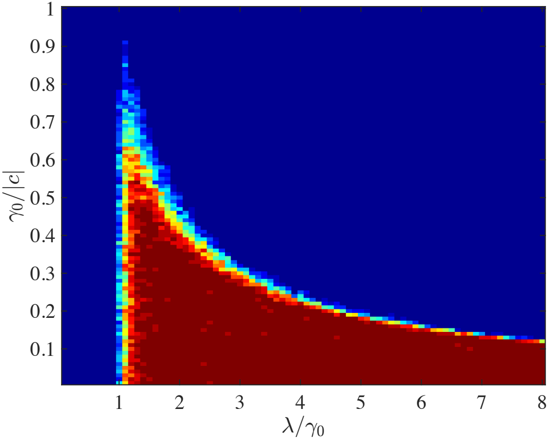
$$\{\hat{f}_k\}, \{\hat{c}_k\} = \operatorname*{argmin}_{\{f_k\},\{c_k\}} \frac{1}{2}\|\mathbf{y} - \sum_{k=1}^{r} c_k \mathbf{a}(f_k)\|_W^2 + \lambda \sum_{k=1}^{r} |c_k|. \quad \text{(NLASSO)}$$

2. Run gradient descent initialized by true parameters $\{f_k^\star\}, \{c_k^\star\}$.
3. Construct the primal candidate $\hat{\mathbf{x}} = \sum_{k=1}^{\hat{r}} \hat{c}_k \mathbf{a}(\hat{f}_k)$ by $\{\hat{f}_k\}, \{\hat{c}_k\}$
4. Show $W(\mathbf{y} - \hat{\mathbf{x}})/\lambda$ satisfies **BIP**.

M. Wainwright. "Sharp thresholds for high-dimensional and noisy sparsity recovery using constrained quadratic programming (Lasso)"

Numerical Experiment

1. SNR as measured by $|c_{\min}|/\gamma_0$ is large.
2. The dynamic range of the coefficients is small.
3. Regularization parameter λ is large compared to γ_0.
4. The frequencies are well-separated.



Setup:

- $n = 130$
- $|c_k^\star| = 1$
- Separation $\geq 2.5/n$

Success means:

- $\max_k |c_k^\star||\hat{f}_k - f_k^\star| \leq \frac{\gamma_0}{2n}$
- $\max_k |\hat{c}_k - c_k^\star| \leq 2\lambda$

Acknowledgements

- This work is supported by National Science Foundation under grants CCF-1464205.

- **arXiv version**: https://arxiv.org/pdf/1612.01459.pdf