Super-resolution and sensor calibration in imaging

Wenjing Liao
School of Mathematics, Georgia Institute of Technology

ICERM
September 25, 2017

My collaborators

Albert Fannjiang
UC Davis

Weilin Li
UMD

Yonina C. Eldar
Tel-Aviv University, Israel

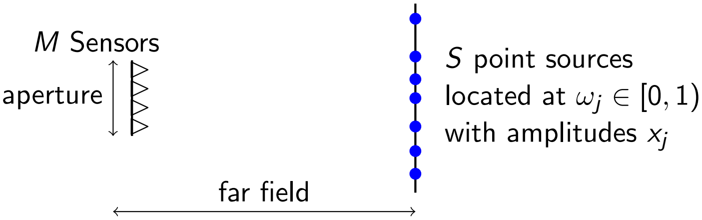
Sui Tang
JHU

Outline

1. Super-resolution
    - Resolution in imaging
    - Super-resolution limit and min-max error
    - Super-resolution algorithms

2. Sensor calibration
    - Problem formulation
    - Uniqueness
    - An optimization approach
    - Numerical simulations

Source localization with sensor array



Point sources: $x(t) = \sum_{j=1}^{S} x_j \delta(t - \omega_j), \quad \omega_j \in [0, 1)$

Measurement at the mth sensor, $m = 0, \ldots, M-1$:

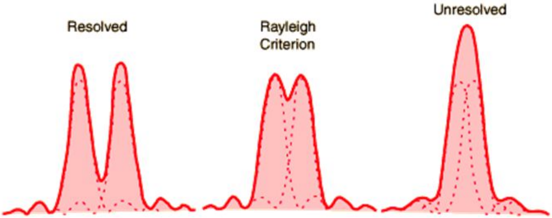
$$y_m = \sum_{j=1}^{S} x_j e^{-2\pi i m \omega_j} + e_m$$

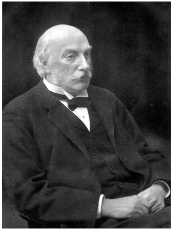
Measurements: $\{y_m : m = 0, \ldots, M-1\}$

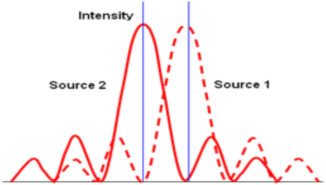
To recover: source locations $\{\omega_j\}_{j=1}^{S}$ and source amplitudes $\{x_j\}_{j=1}^{S}$.

Rayleigh criterion

$$\hat{x}(\omega) = \sum_{m=0}^{M-1} y_m \frac{e^{2\pi i m \omega}}{M}$$

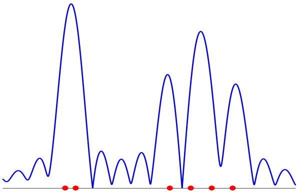




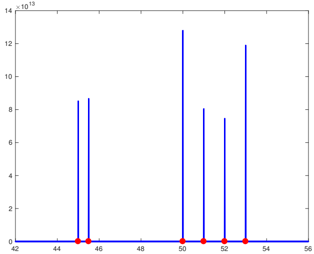


Rayleigh length $= 1/M$

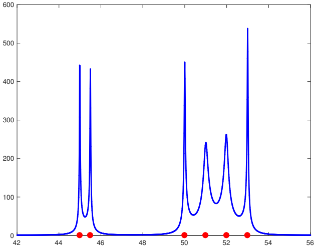
Inverse Fourier transform and the MUSIC algorithm



Multiple Signal Classification (MUSIC): [Schmidt 1981]



noise-free



noisy

Interesting questions

1. What is the super-resolution limit of the "best" algorithm?

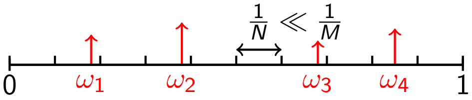
2. What is the super-resolution limit of a specific algorithm?

    - MUSIC [Schmidt 1981]
    - ESPRIT [Roy and Kailath 1989]
    - the matrix pencil method [Hua and Sarkar 1990]

Existing works

1. Super-resolution limit with continuous measurements
   - Donoho 1992, Demanet and Nguyen 2015
2. Performance guarantees for well separated point sources
   - Total variation minimization [Candès and Fernandez-Granda 2013,2014, Tang, Bhaskar, Shah and Recht 2013, Duval and Peyré 2015, Li 2017]
   - Greedy algorithms [Duarte and Baraniuk 2013, Fannjiang and L. 2012]
   - MUSIC [L. and Fannjiang 2016]
   - The matrix pencil method [Moitra 2015]
3. Performance guarantees for super-resolution
   - Total variation min for *positive* sources [Morgenshtern and Candès 2016] or sources with certain sign pattern [Benedetto and Li 2016]
   - Lasso for *positive* sources [Denoyelle, Duval and Peyré 2016]

Discretization on a fine grid



- Point sources: $\mu = \sum_{n=0}^{N-1} x_n \delta_{n/N}$ with $x \in \mathbb{C}_S^N$
- Measurement vector

$$y = \Phi x + e$$

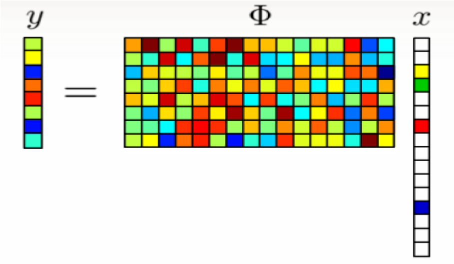
where $\Phi \in \mathbb{C}^{M \times N}$ is the first M rows of the $N \times N$ DFT matrix:

$$\Phi_{m,n} = e^{-2\pi i mn/N}$$

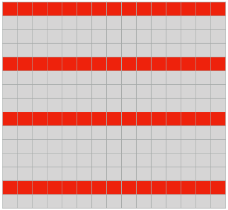
and $\|e\|_2 \leq \delta$.

Super-resolution factor (SRF) $:= \frac{N}{M}$

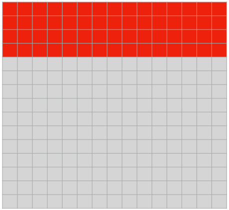
Connection to compressive sensing



Sensing matrices contain certain rows of the DFT matrix.



(a) compressive sensing



(b) super-resolution

Min-max error

Definition (S-min-max error)

Fix positive integers M, N, S such that $S \leq M \leq N$ and let $\delta > 0$. The S-min-max error is

$$\mathcal{E}(M,N,S,\delta) = \inf_{\substack{\tilde{x}=\tilde{x}(y,M,N,S,\delta)\in\mathbb{C}^N \\ y=\Phi x+e}} \sup_{x\in\mathbb{C}^N_S} \sup_{e\in\mathbb{C}^M:\|e\|_2\leq\delta} \|\tilde{x}-x\|_2.$$

Sharp bound on the min-max error

Theorem (Li and L. 2017)

There exist constants $A(S), B(S), C(S)$ such that:

1. *Lower bound. If $M \geq 2S$ and $N \geq C(2S)M^{3/2}$, then*

$$\mathcal{E}(M,N,S,\delta) \geq \frac{\delta}{2B(2S)\sqrt{M}} \mathrm{SRF}^{2S-1}.$$

2. *Upper bound. If $M \geq 4S(2S+1)$ and $N \geq M^2/(2S^2)$, then*

$$\mathcal{E}(M,N,S,\delta) \leq \frac{2\delta}{A(2S)\sqrt{M}} \mathrm{SRF}^{2S-1}.$$

The best algorithm in the upper bound:

$$\min \|z\|_0 \quad \text{subject to} \|\Phi z - y\|_2 \leq \delta$$

W. Li and W. Liao, "Stable super-resolution limit and smallest singular value of restricted Fourier matrices," preprint, arXiv:1709.03146.

Multiple Signal Classification (MUSIC)

- **Pioneering work:** Prony 1795
- **MUSIC in signal processing:** Schmidt 1981
- **MUSIC in imaging:** Devaney 2000, Devaney, Marengo and Gruber 2005, Cheney 2001, Kirsch 2002
- **Related:** the linear sampling method [Cakoni, Colton and Monk 2011], factorization method [Kirsch and Grinsberg 2008]

MUSIC

Assumption: S is known.

$$y_m = \sum_{j=1}^{S} x_j e^{-2\pi i m \omega_j}, \; m = 0, \ldots, M-1.$$

$$H = \mathrm{Hankel}(y) = \begin{bmatrix} y_0 & y_1 & \cdots & y_{M-L} \\ y_1 & y_2 & \cdots & y_{M-L+1} \\ \vdots & \vdots & \vdots & \vdots \\ y_{L-1} & y_L & \cdots & y_{M-1} \end{bmatrix} = \underbrace{\Phi^L}_{L\times S} \underbrace{X}_{S\times S} \underbrace{(\Phi^{M-L+1})^T}_{S\times(M-L+1)}$$

where

$$X = \mathrm{diag}(x_1, \ldots, x_S)$$

$$\phi^L(\omega) = \begin{bmatrix} 1 & e^{-2\pi i \omega} & \ldots & e^{-2\pi i (L-1)\omega} \end{bmatrix}^T \in \mathbb{C}^L$$

$$\Phi^L = [\phi^L(\omega_1) \; \ldots \; \phi^L(\omega_S)] \in \mathbb{C}^{L\times S}.$$

MUSIC with noiseless measurements

$$H = \Phi^L X (\Phi^{M-L+1})^T$$

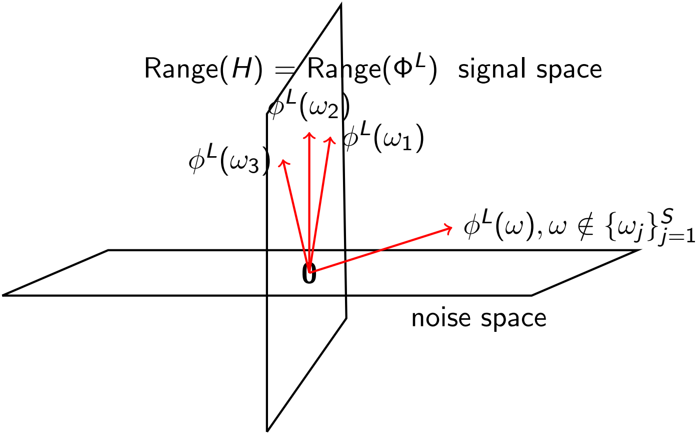
Suppose $\{\omega_j\}_{j=1}^S$ are distinct.

1. If $L \geq S$, $\mathrm{rank}(\Phi^L) = S$.
2. If $M - L + 1 \geq S$, $\mathrm{Range}(H) = \mathrm{Range}(\Phi^L)$.
3. If $L \geq S + 1$, $\mathrm{rank}([\Phi^L \;\; \phi^L(\omega)]) = S + 1$ if and only if $\omega \notin \{\omega_j\}_{j=1}^S$.

Theorem

If $L \geq S + 1$ and $M - L + 1 \geq S$, $\omega \in \{\omega_j\}_{j=1}^S$ iff $\phi^L(\omega) \in \mathrm{Range}(H)$.

Exact recovery with $M \geq 2S$ regardless of the support .

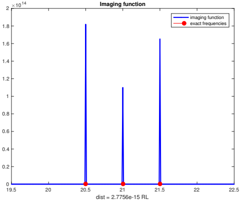


- Noise-space correlation function: $\mathcal{N}(\omega) = \frac{\|\mathcal{P}_{\text{noise}}\phi^L(\omega)\|_2}{\|\phi^L(\omega)\|_2}$
- Imaging function: $\mathcal{J}(\omega) = \frac{1}{\mathcal{N}(\omega)}$

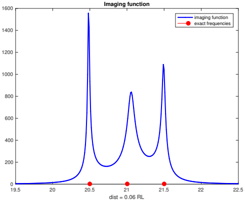
$$\mathcal{N}(\omega_j) = 0 \text{ and } \mathcal{J}(\omega_j) = \infty,\ j = 1, \ldots, S.$$

MUSIC with noisy measurements

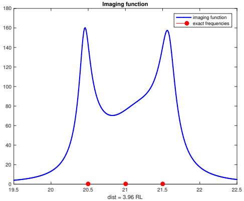
Three sources separated by 0.5 RL, $e \sim N(0, \sigma^2 I_M)$



(c) $\sigma = 0$



(d) $\sigma = 0.001$



(e) $\sigma = 0.01$

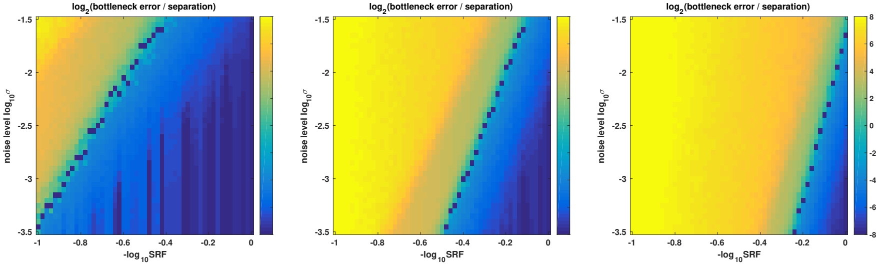
Recall upper bound of the min-max error

$$\mathcal{E}(M, N, S, \delta) \lesssim \frac{\delta}{\sqrt{M}} \mathrm{SRF}^{2S-1}$$

The noise that the “best” algorithm can handle is $\delta \sim \left(\frac{1}{\mathrm{SRF}}\right)^{2S-1}$.

Phase transition

- S consecutive point sources on the grid with spacing $1/N$
- Support error: $d(\{\omega_j\}_{j=1}^S, \{\hat{\omega}_j\}_{j=1}^S)$
- Noise $e \sim N(0, \sigma^2 I_M) + i \cdot N(0, \sigma^2 I_M)$, so $\mathbb{E}\|e\|_2 = \sqrt{2M}\sigma$.



$S = 2$ $\qquad$ $S = 3$ $\qquad$ $S = 4$

Figure: The average $\log_2[\frac{\text{Support error}}{1/N}]$ over 100 trials with respect to $\log_{10} \frac{1}{\text{SRF}}$ (x-axis) and $\log_{10} \sigma$ (y-axis).

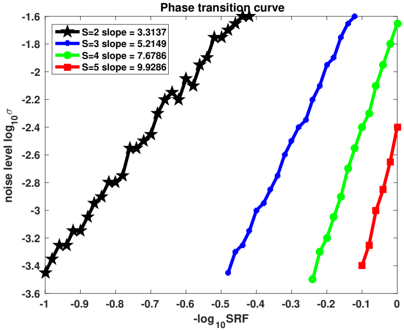
Super-resolution limit of MUSIC

The phase transition curve is

$$\sigma \sim \left(\frac{1}{\mathrm{SRF}}\right)^{p(S)}$$

where

$$2S - 1 \leq p(S) \leq 2S.$$



Future work:

$$\text{Support error by MUSC} \lesssim \mathrm{SRF}^{p(S)} \cdot \sigma.$$

Outline

1. Super-resolution
   - Resolution in imaging
   - Super-resolution limit and min-max error
   - Super-resolution algorithms
2. Sensor calibration
   - Problem formulation
   - Uniqueness
   - An optimization approach
   - Numerical simulations

Sensor calibration

Measurement at the m-th sensor, $m = 0, \ldots, M-1$:

$$y_m(t) = g_m \sum_{j=1}^{S} x_j(t) e^{-2\pi i m \omega_j} + e_m(t)$$

Multiple snapshots of measurements:

$$\{y_m(t), m = 0, \ldots, M-1, \ t \in \Gamma\}$$

To recover:

- Calibration parameters $g = \{g_m\}_{m=0}^{M-1} \in \mathbb{C}^M$
- Source locations $\{\omega_j\}_{j=1}^{S}$ and source amplitudes $x_j(t)$

Assumptions

Matrix form:

$$\underbrace{y(t)}_{\mathbb{C}^M} = \underbrace{\mathrm{diag}(g)}_{\mathbb{C}^{M\times M}} \underbrace{A}_{\mathbb{C}^{M\times S}} \underbrace{x(t)}_{\mathbb{C}^S} + \underbrace{e(t)}_{\mathbb{C}^M}$$

$$A_{n,j} = e^{-2\pi i m\omega_j}$$

$x(t) = [x_1(t) \ \ldots x_S(t)]^T, y(t) = [y_0(t) \ldots y_{M-1}(t)]^T, e(t) = [e_0(t) \ldots e_{M-1}(t)]^T$

Assumptions:

1. $|g_m| \neq 0, \ m = 0, \ldots, M-1;$
2. $\mathbb{E}x(t) = 0$ and $\mathbb{E}e(t) = 0;$
3. $R^x := \mathbb{E}x(t)x^*(t) = \mathrm{diag}(\{\gamma_j^2\}_{j=1}^S);$
4. $\mathbb{E}x(t)e^*(t) = 0;$
5. $\mathbb{E}e(t)e^*(t) = \sigma^2 I_M$ where σ represents noise level.

Uniqueness up to a trivial ambiguity

Trivial ambiguity: $\{\tilde{g}, \{\tilde{\omega}_j\}_{j=1}^S, \tilde{x}(t)\}$ is called equivalent to $\{g, \{\omega_j\}_{j=1}^S, x(t)\}$ up to a trivial ambiguity if there exist $c_0 > 0, c_1, c_2 \in \mathbb{R}$:

$$\begin{aligned} \tilde{g} &= \{\tilde{g}_m = c_0 e^{i(c_1 + mc_2)} g_m\}_{m=0}^{M-1} \\ \tilde{\mathcal{S}} &= \{\tilde{\omega}_j : \tilde{\omega}_j = \omega_j - c_2/(2\pi)\}_{j=1}^S \\ \tilde{x}(t) &= x(t) c_0^{-1} e^{-ic_1}. \end{aligned}$$

Uniqueness with infinite snapshots of noiseless measurements:

Let $f_m = \sum_{j=1}^s \gamma_j^2 e^{2\pi i m \omega_j}, \; m = 0, \ldots, M-1.$

Theorem

Suppose $|f_1| > 0$ and $M \geq S+1$. Let $\{g, \{\omega_j\}_{j=1}^S, x(t)\}$ be a solution to the calibration problem. If there is another solution $\{\tilde{g}, \{\tilde{\omega}_j\}_{j=1}^S, \tilde{x}(t)\}$, then $\{\tilde{g}, \{\tilde{\omega}_j\}_{j=1}^S, \tilde{x}(t)\}$ is equivalent to $\{g, \{\omega_j\}_{j=1}^S, x(t)\}$.

Covariance matrix

Pioneering work: Full algebraic method [Paulraj and Kailath, 1985], Partial algebraic method [Wylie, Roy and Schmitt, 1993]

$$R^y := \mathbb{E}y(t)y^*(t) = \mathrm{diag}(g)AR^xA^*\mathrm{diag}(\bar{g})$$

$$\mathcal{H} : \mathbb{C}^M \to \mathbb{C}^{M\times M} : \ \mathcal{H}(f) := \begin{bmatrix} f_0 & \bar{f}_1 & \ddots & \bar{f}_{N-1} \\ f_1 & f_0 & \ddots & \bar{f}_{N-2} \\ \ddots & \ddots & \ddots & \ddots \\ f_{N-1} & f_{N-2} & \ddots & f_0 \end{bmatrix} = AR^xA^*. \text{ Then}$$

$$R^y = \mathrm{diag}(g)\mathcal{H}(f)\mathrm{diag}(\bar{g})$$
$$R^y_{m,n} = g_m\bar{g}_n f_{m-n}$$

When $f_1 \neq 0$, the diagonal and subdiagonal entries in R^y determine the solution up to a trivial ambiguity.

Algebraic methods

Sensitivity of the partial algebraic method:

- $N \geq s+1$, $|f_1| > 0$ and sources are separated by $1/M$.
- Empirical covariance matrix is computed with L snapshots of measurements.

We proved that,

$$\mathbb{E} \min_{c_0>0, c_1, c_2 \in \mathbb{R}} \max_m |c_0 \widehat{g}_m - e^{i(c_1+mc_2)} g_m| \leq O\left(\frac{\max(\sigma, \sigma^2)}{\sqrt{L}}\right),$$

Partial algebraic method: only diagonal and subdiagonal entries in the covariance matrix are used

Full algebraic method: problem of phase wrapping

An optimization approach

$$R^y = GAR^x A^* G^* = \mathrm{diag}(g)\mathcal{H}(f)\mathrm{diag}(\bar{g})$$

Optimization problem:

$$\min_{\mathbf{g},\mathbf{f}\in\mathbb{C}^M} \mathcal{L}(\mathbf{g},\mathbf{f}) := \left\| \mathrm{diag}(\mathbf{g})\mathcal{H}(\mathbf{f})\mathrm{diag}(\bar{\mathbf{g}}) - \widehat{R}^y \right\|_F^2 .$$

- If $\widehat{R}^y = R^y$, the global minimizer of $\mathcal{L}$ is equivalent to the ground truth (g, f).

Regularized optimization

Goal: prevent $\|\mathbf{g}\| \to \infty$ and $\|\mathbf{f}\| \to 0$ (or vice versa)
$\widehat{n}_0$ is an estimator of $n_0 := \|g\|^2 \|f\|$ from the partial algebraic method.

Regularized optimization:

$$\min_{\mathbf{g},\mathbf{f} \in \mathbb{C}^N} \tilde{\mathcal{L}}(\mathbf{g},\mathbf{f}) := \mathcal{L}(\mathbf{g},\mathbf{f}) + \mathcal{G}(\mathbf{g},\mathbf{f})$$

$$\mathcal{G}(\mathbf{g},\mathbf{f}) = \rho \left[\mathcal{G}_0\left(\frac{\|\mathbf{f}\|^2}{2\widehat{n}_0} \right) + \mathcal{G}_0\left(\frac{\|\mathbf{g}\|^2}{\sqrt{2\widehat{n}_0}} \right) \right]$$

where $\mathcal{G}_0(z) = (\max(z-1,0))^2$ and $\rho \geq \frac{3\widehat{n}_0 + \|R^y - \widehat{R}^y\|_F}{(\sqrt{2}-1)^2}$

Initialization: $(\mathbf{g}^0, \mathbf{f}^0) : \|\mathbf{g}^0\|^2 \leq \sqrt{2\widehat{n}_0}, \|\mathbf{f}^0\| \leq \sqrt{2\widehat{n}_0}$

Feasible set: $\mathcal{N}_{\widehat{n}_0} = \{(\mathbf{g},\mathbf{f}) : \|\mathbf{g}\|^2 \leq 2\sqrt{\widehat{n}_0}, \|\mathbf{f}\| \leq 2\sqrt{\widehat{n}_0}\}$

Wirtinger gradient descent

for $k = 1, 2, \dots,$

- $\mathbf{g}^k = \mathbf{g}^{k-1} - \eta^k \nabla_{\mathbf{g}} \tilde{\mathcal{L}}(\mathbf{g}^{k-1}, \mathbf{f}^{k-1})$
- $\mathbf{f}^k = \mathbf{f}^{k-1} - \eta^k \nabla_{\mathbf{f}} \tilde{\mathcal{L}}(\mathbf{g}^{k-1}, \mathbf{f}^{k-1})$

end

Theorem (Eldar, L. and Tang)

If the step length is chosen such that

$$\eta^k \le \frac{2}{146\widehat{n}_0 \max(\sqrt{\widehat{n}_0}, \sqrt[4]{\widehat{n}_0}) + 8\widehat{n}_0 + 16 \max(\sqrt{\widehat{n}_0}, \sqrt[4]{\widehat{n}_0}) \|R^y - \widehat{R}^y\|_F + \frac{8\rho}{\min(\widehat{n}_0, \sqrt{\widehat{n}_0})}},$$

then Wirtinger gradient descent gives rise to $(\mathbf{g}^k, \mathbf{f}^k) \in \mathcal{N}_{\widehat{n}_0}$, *and*

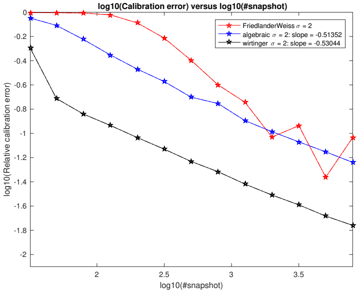
$$\|\nabla \tilde{\mathcal{L}}(\mathbf{g}^k, \mathbf{f}^k)\| \to 0, \text{ as } k \to \infty.$$

Y. C. Eldar, W. Liao and S. Tang, "Sensor calibration for off-the-grid spectral estimation," preprint, arXiv: 1707.03378.

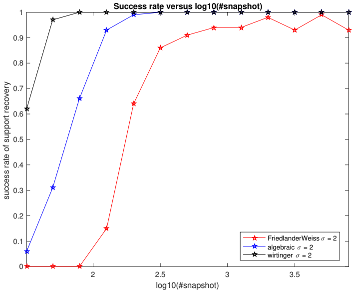
Sensitivity to the number of snapshots

1. the partial algebraic method
2. our optimization approach
3. an alternating minimization: [Friedlander and Weiss 1990]

- 20 sources separated by $2/M$ and noise level $\sigma = 2$



Relative calibration error versus L

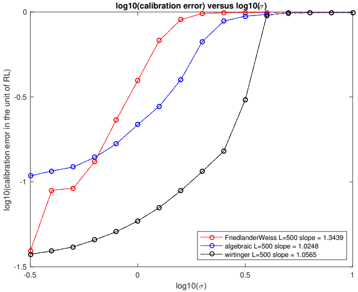


Support success rate versus L

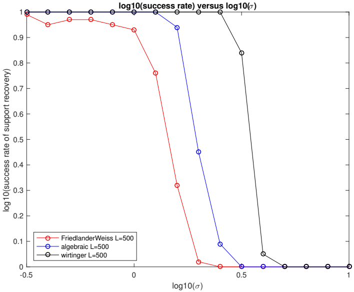
Observation: Calibration error $= O(L^{-\frac{1}{2}})$

Sensitivity to noise level σ

- 20 sources separated by $2/M$ and $L = 500$



Relative calibration error versus σ



Support success rate versus σ

Observation: Calibration error $= O(\sigma)$

Conclusion

1. Super-resolution
   - Resolution limit and a sharp bound on the min-max error
   - Resolution limit of the MUSIC algorithm
2. Sensor calibration
   - Uniqueness with infinite snapshots of noiseless data
   - The partial algebraic method and a stability analysis
   - An optimization approach and convergence to a stationary point

Thank you for your attention!

Wenjing Liao

Georgia Institute of Technology

wliao60@gatech.edu

http://people.math.gatech.edu/~wliao60