UC San Diego
JACOBS SCHOOL OF ENGINEERING
Electrical and Computer Engineering

Understanding Trade-offs in Super-resolution Imaging with Spatiotemporal Measurements

Piya Pal

Department of Electrical and Computer Engineering
University of California, San Diego

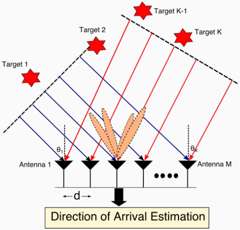
One World Signal Processing Seminar,
November 11, 2020

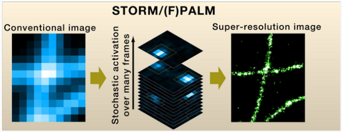
Outline

1. Super-Resolution: Motivation and Background
2. Super-Resolution, Sparsity and Correlation Priors
3. Super-Resolution via Parameter Estimation: Going Off the Grid
4. Conclusion

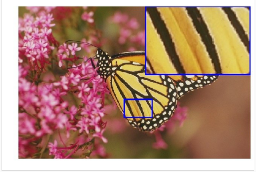
Super-Resolution: Motivation and Background

Super-Resolution: Motivation and Applications





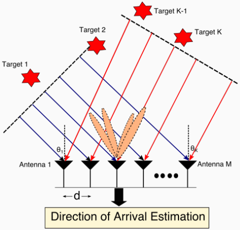
B. Huang, CELL , 2010

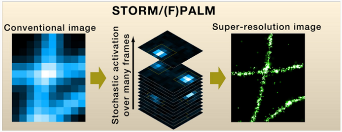


W. Shi, CVPR , 2016

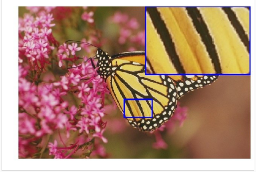
- The goal of Super-resolution is to recover "lost" details (typically high frequency components) from noisy, low-resolution (typically low-frequency) measurements acquired by a physical system.

Super-Resolution: Motivation and Applications





B. Huang, CELL, 2010



W. Shi, CVPR, 2016

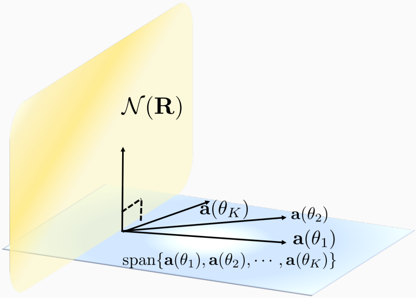
- The goal of Super-resolution is to recover “lost” details (typically high frequency components) from noisy, low-resolution (typically low-frequency) measurements acquired by a physical system.
- The problem has origins in optics. Features widely across many applications, including radar, microscopy, medical imaging, radio astronomy, image processing/computer vision...

Super-Resolution: Classical Methods and Recent Advances

- Harmonic Retrieval Problem:

$$y_m = \sum_{k=1}^{K} e^{jm\omega_k} c_k + n_m, 1 \leqslant m \leqslant M$$

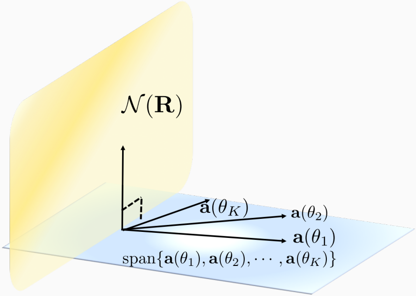
- Classical Methods are algebraic, and they utilize the structure of subspace spanned by Vandermonde vectors.



- Harmonic Retrieval Problem:

$$y_m = \sum_{k=1}^{K} e^{jm\omega_k} c_k + n_m, 1 \leqslant m \leqslant M$$

- Classical Methods are algebraic, and they utilize the structure of subspace spanned by Vandermonde vectors.



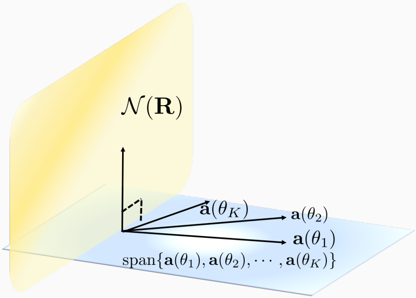
- Many impactful results spanning decades: [Schmidt'86],[Kailath'89],[Hua et.al'90], [Kaveh'86], [Rao,Hari'89], [Stoica,Nehorai'89], [Vaccaro'93],[Krim'96]...
- Guarantees are mostly asymptotic/perturbation-based.

Super-Resolution: Classical Methods and Recent Advances

- Harmonic Retrieval Problem:

$$y_m = \sum_{k=1}^{K} e^{jm\omega_k} c_k + n_m, 1 \leqslant m \leqslant M$$

- Classical Methods are algebraic, and they utilize the structure of subspace spanned by Vandermonde vectors.



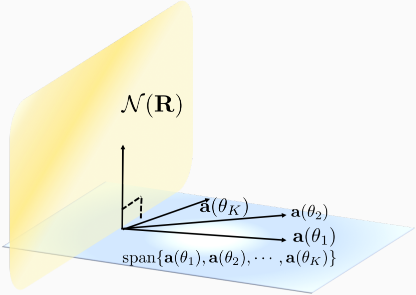
- Many impactful results spanning decades: [Schmidt'86],[Kailath'89],[Hua et.al'90], [Kaveh'86], [Rao,Hari'89], [Stoica,Nehorai'89], [Vaccaro'93],[Krim'96]...
- Guarantees are mostly asymptotic/perturbation-based.
- Recent advances in non-asymptotic guarantees of classical methods: [Liao'14], [Moitra'15,'20], [Li'19],[Qiao,Pal'19],[Hucumenoglu,Pal'20]..

Super-Resolution: Classical Methods and Recent Advances

- Harmonic Retrieval Problem:

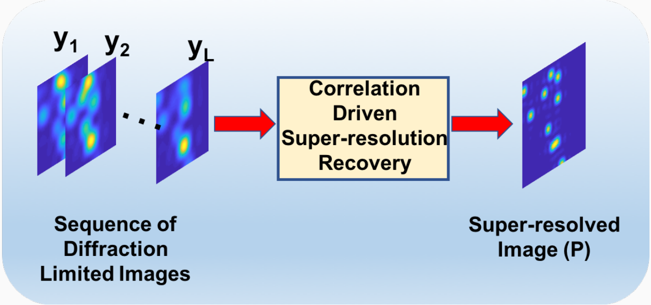
$$y_m = \sum_{k=1}^{K} e^{jm\omega_k} c_k + n_m, 1 \leqslant m \leqslant M$$

- Classical Methods are algebraic, and they utilize the structure of subspace spanned by Vandermonde vectors.



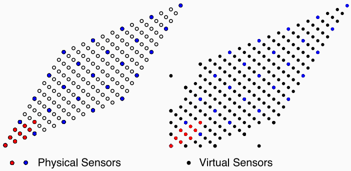
- Many impactful results spanning decades: [Schmidt'86],[Kailath'89],[Hua et.al'90], [Kaveh'86], [Rao,Hari'89], [Stoica,Nehorai'89], [Vaccaro'93],[Krim'96]...
- Guarantees are mostly asymptotic/perturbation-based.
- Recent advances in non-asymptotic guarantees of classical methods: [Liao'14], [Moitra'15,'20], [Li'19],[Qiao,Pal'19],[Hucumenoglu,Pal'20]..
- Modern Convex algorithms for Super-resolution: Atomic norm/TV norm minimization. Robusntess guarantees, minimax optimality: [Candes,Fernandez-Granda'12-'20],[Tang et al.'12-20]...

This Talk: Multiple (Temporal) Measurements and Correlation Priors



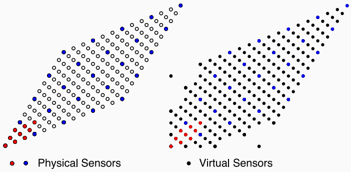
- In many applications (such as microscopy, radar target localization, interferometry), we acquire several low-resolution measurements of a scene of interest over time.
- Incorporation of temporal measurements and correlation priors can significantly enhance super-resolution capabilities.

Sparse Arrays and Aperture Synthesis



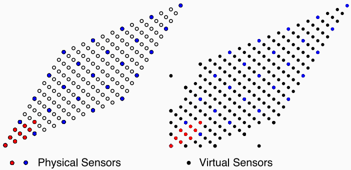
- In many applications sources are assumed to be spatially incoherent (or statistically uncorrelated).
- By utilizing a sparse sensing geometry and computing spatial correlation between sensor pairs, it is possible to generate the effect of a virtual difference co-array. [Moffet'68],[Pillai'85],[Kassam'90],[Abramovich],[Pal,Vaidyanathan'10],[Amin'15],[Wang,Nehorai'17],[Koochakzadeh,Pal'16],[Qiao,Pal'20]...

Sparse Arrays and Aperture Synthesis



- In many applications sources are assumed to be spatially incoherent (or statistically uncorrelated).
- By utilizing a sparse sensing geometry and computing spatial correlation between sensor pairs, it is possible to generate the effect of a virtual difference co-array. [Moffet'68],[Pillai'85],[Kassam'90],[Abramovich],[Pal,Vaidyanathan'10],[Amin'15],[Wang,Nehorai'17], [Koochakzadeh,Pal'16],[Qiao,Pal'20]...
- Asymptotic guarantees for resolving more sources than sensors, significantly smaller Cramér-Rao Bounds.

Sparse Arrays and Aperture Synthesis



- In many applications sources are assumed to be spatially incoherent (or statistically uncorrelated).
- By utilizing a sparse sensing geometry and computing spatial correlation between sensor pairs, it is possible to generate the effect of a virtual difference co-array. [Moffet'68],[Pillai'85],[Kassam'90],[Abramovich],[Pal,Vaidyanathan'10],[Amin'15],[Wang,Nehorai'17], [Koochakzadeh,Pal'16],[Qiao,Pal'20]...
- Asymptotic guarantees for resolving more sources than sensors, significantly smaller Cramér-Rao Bounds.
- **Non-asymptotic Guarantees: Largely open.**

Open Questions of Interest

- Classical Subspace based algorithms do not explicitly need separation condition, but their guarantees are mostly asymptotic in the number of snapshots.
- Modern TV-norm and atomic norm based algorithms offer non-asymptotic robustness guarantees, but require a minimum separation condition, even in absence of noise (reminiscent of Rayleigh resolution limit).

Open Questions of Interest

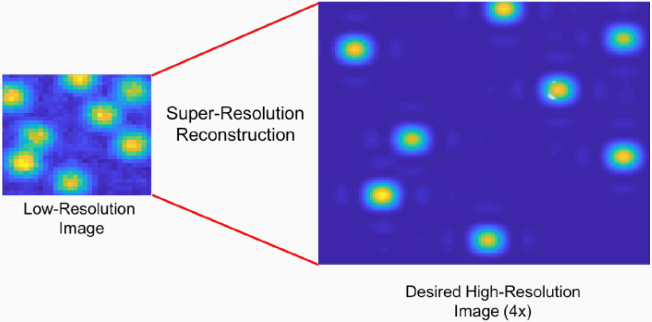
- Classical Subspace based algorithms do not explicitly need separation condition, but their guarantees are mostly asymptotic in the number of snapshots.
- Modern TV-norm and atomic norm based algorithms offer non-asymptotic robustness guarantees, but require a minimum separation condition, even in absence of noise (reminiscent of Rayleigh resolution limit).

Can correlation priors and aperture synthesis provably lead to improvements in resolution? Can a *strict separation condition* be relaxed and *noise amplification be tamed* by exploiting

- Sensing geometry?
- Temporal snapshots?
- Inherent conic constraints?

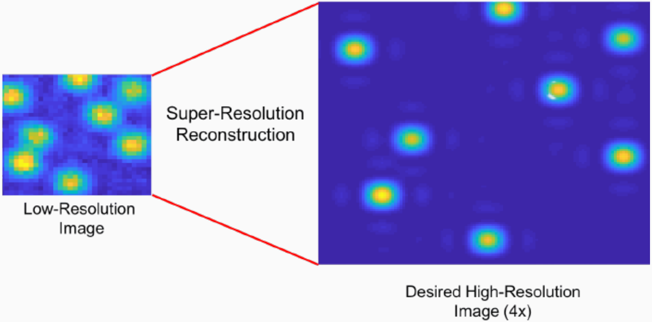
Super-Resolution, Sparsity and Correlation Priors

Discrete Setup of Super-Resolution Image Reconstruction



- Discrete Super-resolution: The goal is to reconstruct a desired image on a high-resolution grid, given low-resolution measurements collected by a sensor array.

Discrete Setup of Super-Resolution Image Reconstruction



- Discrete Super-resolution: The goal is to reconstruct a desired image on a high-resolution grid, given low-resolution measurements collected by a sensor array.
- Widely used in optical super-resolution imaging [Solomon,Eldar,Segev'18,Goodman et. al'17]

Noise Amplification in Super-resolution: Discrete Setup

- The discrete version of the super-resolution problem has been studied extensively, following pioneering works by [Donoho'90]
- Discrete version appears frequency in applications where the goal is to display a super-resolved image on a desired high resolution grid [Solomon,Eldar et. al'18].

Noise Amplification in Super-resolution: Discrete Setup

- The discrete version of the super-resolution problem has been studied extensively, following pioneering works by [Donoho'90]
- Discrete version appears frequency in applications where the goal is to display a super-resolved image on a desired high resolution grid [Solomon,Eldar et. al'18].

Measurement Model [Morgenshtern,Candes'16]

$$\boldsymbol{y} = \boldsymbol{Q}\boldsymbol{x} + \boldsymbol{n}$$

- $\boldsymbol{y} \in \mathbb{C}^N$: Low-resolution measurements, contaminated with noise $\boldsymbol{n}$.

Noise Amplification in Super-resolution: Discrete Setup

- The discrete version of the super-resolution problem has been studied extensively, following pioneering works by [Donoho'90]
- Discrete version appears frequency in applications where the goal is to display a super-resolved image on a desired high resolution grid [Solomon,Eldar et. al'18].

Measurement Model [Morgenshtern,Candes'16]

$$\boldsymbol{y} = \boldsymbol{Q}\boldsymbol{x} + \boldsymbol{n}$$

- $\boldsymbol{y} \in \mathbb{C}^N$: Low-resolution measurements, contaminated with noise $\boldsymbol{n}$.
- $\boldsymbol{Q} \in \mathbb{C}^{N \times N}$: Discrete Convolution operator, representing a low-pass filter with cut-off $f_c < N$:

$$\boldsymbol{Q} = \boldsymbol{W}^H \boldsymbol{\Lambda} \boldsymbol{W}, \quad [\boldsymbol{W}]_{m,n} = \frac{1}{\sqrt{N}} e^{-j2\pi mn/N}, \ {}_{-N/2+1 \leqslant m \leqslant N/2, \ 0 \leqslant n \leqslant N-1}$$

where $\Lambda = \text{diag}(p_{-N/2+1}, p_2, \cdots, p_{N/2})$ with $p_n = 0, |n| > f_c$.

Noise Amplification in Super-resolution: Discrete Setup

- The discrete version of the super-resolution problem has been studied extensively, following pioneering works by [Donoho'90]
- Discrete version appears frequency in applications where the goal is to display a super-resolved image on a desired high resolution grid [Solomon,Eldar et. al'18].

Measurement Model [Morgenshtern,Candes'16]

$$\boldsymbol{y} = \boldsymbol{Q}\boldsymbol{x} + \boldsymbol{n}$$

- $\boldsymbol{y} \in \mathbb{C}^N$: Low-resolution measurements, contaminated with noise $\boldsymbol{n}$.
- $\boldsymbol{Q} \in \mathbb{C}^{N \times N}$: Discrete Convolution operator, representing a low-pass filter with cut-off $f_c < N$:

$$\boldsymbol{Q} = \boldsymbol{W}^H \boldsymbol{\Lambda} \boldsymbol{W}, \quad [\boldsymbol{W}]_{m,n} = \frac{1}{\sqrt{N}} e^{-j2\pi mn/N}, \; {\scriptstyle -N/2+1 \leqslant m \leqslant N/2, \; 0 \leqslant n \leqslant N-1}$$

  where $\Lambda = \text{diag}(p_{-N/2+1}, p_2, \cdots, p_{N/2})$ with $p_n = 0, |n| > f_c$.
- $\boldsymbol{x} \in \mathbb{C}^N$: Desired high-resolution signal.

Stable Super-resolution

Representation in Frequency Domain:

$$\boldsymbol{y} = \boldsymbol{Q}\boldsymbol{x} + \boldsymbol{n} \Rightarrow \boldsymbol{W}\boldsymbol{y} = \boldsymbol{\Lambda}\boldsymbol{W}\boldsymbol{x} + \boldsymbol{W}\boldsymbol{n}$$

Stable Super-resolution

Representation in Frequency Domain:

$$\boldsymbol{y} = \boldsymbol{Q}\boldsymbol{x} + \boldsymbol{n} \Rightarrow \boldsymbol{W}\boldsymbol{y} = \boldsymbol{\Lambda}\boldsymbol{W}\boldsymbol{x} + \boldsymbol{W}\boldsymbol{n}$$

- If $\boldsymbol{\Lambda}$ has a flat spectrum, then we need to recover $\boldsymbol{x} \in \boldsymbol{C}^N$ from $M = 2f_c + 1 \ll N$ low frequency components (DFT coefficients), corrupted with noise.

Stable Super-resolution

Representation in Frequency Domain:

$$\boldsymbol{y} = \boldsymbol{Q}\boldsymbol{x} + \boldsymbol{n} \Rightarrow \boldsymbol{W}\boldsymbol{y} = \boldsymbol{\Lambda}\boldsymbol{W}\boldsymbol{x} + \boldsymbol{W}\boldsymbol{n}$$

- If $\boldsymbol{\Lambda}$ has a flat spectrum, then we need to recover $\boldsymbol{x} \in \boldsymbol{C}^N$ from $M = 2f_c + 1 \ll N$ low frequency components (DFT coefficients), corrupted with noise.
- Unambiguous and stable recovery of $\boldsymbol{x}$ from $\boldsymbol{y}$ is not possible unless we exploit **priors** on $\boldsymbol{x}$.

Stable Super-resolution

Representation in Frequency Domain:

$$\boldsymbol{y} = \boldsymbol{Q}\boldsymbol{x} + \boldsymbol{n} \Rightarrow \boldsymbol{W}\boldsymbol{y} = \boldsymbol{\Lambda}\boldsymbol{W}\boldsymbol{x} + \boldsymbol{W}\boldsymbol{n}$$

- If $\boldsymbol{\Lambda}$ has a flat spectrum, then we need to recover $\boldsymbol{x} \in \boldsymbol{C}^N$ from $M = 2f_c + 1 \ll N$ low frequency components (DFT coefficients), corrupted with noise.
- Unambiguous and stable recovery of $\boldsymbol{x}$ from $\boldsymbol{y}$ is not possible unless we exploit **priors** on $\boldsymbol{x}$.

Let $\mathcal{C}$ denote a class of signals that the desired $\boldsymbol{x}$ belongs to (captures apriori information).

Stable Super-resolution

Representation in Frequency Domain:

$$\boldsymbol{y} = \boldsymbol{Q}\boldsymbol{x} + \boldsymbol{n} \Rightarrow \boldsymbol{W}\boldsymbol{y} = \boldsymbol{\Lambda}\boldsymbol{W}\boldsymbol{x} + \boldsymbol{W}\boldsymbol{n}$$

- If $\boldsymbol{\Lambda}$ has a flat spectrum, then we need to recover $\boldsymbol{x} \in \boldsymbol{C}^N$ from $M = 2f_c + 1 \ll N$ low frequency components (DFT coefficients), corrupted with noise.
- Unambiguous and stable recovery of x from $\boldsymbol{y}$ is not possible unless we exploit **priors** on x.

Let $\mathcal{C}$ denote a class of signals that the desired $\boldsymbol{x}$ belongs to (captures apriori information).

Stable Recovery

We say that an estimate $\hat{\boldsymbol{x}}$ leads to stable recovery of $\boldsymbol{x}$ (using the apriori information in $\mathcal{C}$), if

$$\|\boldsymbol{x} - \hat{\boldsymbol{x}}\| \leqslant NA(\mathcal{C}, n, N)\ .\ \|\boldsymbol{n}\|$$

$$NA(\mathcal{C}, n, N): \text{ Noise Amplification Factor}$$

Non-negative Super-Resolution and Noise Amplification

- Suppose we have the apriori information that $\boldsymbol{x} \geqslant \mathbf{0}$, i.e. $\boldsymbol{x}$ is non-negative.
- Can this prior information enable stable recovery?

Non-negative Super-Resolution and Noise Amplification

- Suppose we have the apriori information that $\boldsymbol{x} \geqslant \mathbf{0}$, i.e. $\boldsymbol{x}$ is non-negative.
- Can this prior information enable stable recovery?

Non negative Super-resolution [Morgenshtern,Candes 2016]

$$\min_{\boldsymbol{z}} \|\boldsymbol{y} - \boldsymbol{Q}\boldsymbol{z}\|_1 \quad \text{subjet to } \boldsymbol{z} \geqslant \mathbf{0}$$

No explicit regularizer (such as sparsity enforcing l_1 norm, or TV norm) utilized, other than non-negative constraint on $\boldsymbol{x}$.

Stable recovery is still possible if the ground truth $\boldsymbol{x}$ is non-negative and satisfies Rayleigh-Regularity.

Stability of Non-negative Super-resolution

Rayleigh Regularity [Morgenshtern,Candes'16]: Informally, a signal obeys Rayleigh regularity with parameters (d, r) if it contains no more than r spikes in any d consecutive intervals, each of length $\frac{1}{f_c}$.

Stability of Non-negative Super-resolution

Rayleigh Regularity [Morgenshtern,Candes'16]: Informally, a signal obeys Rayleigh regularity with parameters (d, r) if it contains no more than r spikes in any d consecutive intervals, each of length $\frac{1}{f_c}$.

Theorem (Stable Non-negative Super-resolution [Morgenshtern,Candes'16])

Suppose $\boldsymbol{x}$ satisfies Rayleigh regularity condition with parameters $(3.724r, r)$, and the filter $\boldsymbol{Q}$ has a flat or triangular spectrum. Then the solution $\hat{\boldsymbol{x}}$ to (CVX) obeys

$$\|\hat{\boldsymbol{x}} - \boldsymbol{x}\|_1 \leqslant C \left(\frac{N}{M-1} \right)^{2r} \|\boldsymbol{n}\|_1$$

Stability of Non-negative Super-resolution

Rayleigh Regularity [Morgenshtern,Candes'16]: Informally, a signal obeys Rayleigh regularity with parameters (d, r) if it contains no more than r spikes in any d consecutive intervals, each of length $\frac{1}{f_c}$.

Theorem (Stable Non-negative Super-resolution [Morgenshtern,Candes'16])

Suppose $\boldsymbol{x}$ satisfies Rayleigh regularity condition with parameters $(3.724r, r)$, and the filter $\boldsymbol{Q}$ has a flat or triangular spectrum. Then the solution $\hat{\boldsymbol{x}}$ to (CVX) obeys

$$\|\hat{\boldsymbol{x}} - \boldsymbol{x}\|_1 \leqslant C \left(\frac{N}{M-1}\right)^{2r} \|\boldsymbol{n}\|_1$$

When the sparsity pattern of $\boldsymbol{x}$ obeys the conventional separation condition $\Delta > \frac{c}{M-1}$ (with $r = 1$), noise amplifies by a factor of $\left(\frac{N}{M-1}\right)^2$=SRF^2.

Our Goal: Super-Resolution with Spatiotemporal Measurements

Suppose we collect a set of L temporal measurement vectors $\boldsymbol{y}_l \in \mathbb{C}^M$

$$\boldsymbol{y}_l = \boldsymbol{A}\boldsymbol{x}_l + \boldsymbol{n}_l, \quad 1 \leqslant l \leqslant L$$

- $\boldsymbol{A} \in \mathbb{C}^{M \times N}$ $(M < N)$ is an undersampled (fat) DFT matrix:

$$\boldsymbol{A}_{m,n} = e^{j2\pi d_m n/N}, \ {\scriptstyle 1 \leqslant m \leqslant M, \ 0 \leqslant n \leqslant N-1}$$

  where d_m denotes the (normalized) location of the mth sensing element.
- Common Support: Supp$(\boldsymbol{x}_l) = \mathcal{S}, \quad l = 1, 2, \cdots, L$
- Special Case: When $\{d_m\}_{m=1}^M$ is a set of consecutive integers, each measurement vector follows the same model as [Morghenstern,Candes16].
- Appears widely in Mulitple Measurement Vector (MMV) models.

Super-resolution Correlation-Imaging

- In many problems, the sources are assumed to be spatially incoherent

$$E\left(x_i x_j^*\right) = p_i \delta[i-j], \quad 1 \leqslant i, j \leqslant K$$

- Such assumptions are heavily exploited in correlation microscopy (e.g. SOFI, SPARCOM) to exploit the independent statistical fluctuation of fluoresecent emitters to aid super-resolution in the discrete setting.

Super-resolution Correlation-Imaging

- In many problems, the sources are assumed to be spatially incoherent

$$E\left(x_i x_j^*\right) = p_i \delta[i-j], \quad 1 \leqslant i, j \leqslant K$$

- Such assumptions are heavily exploited in correlation microscopy (e.g. SOFI, SPARCOM) to exploit the independent statistical fluctuation of fluoresecent emitters to aid super-resolution in the discrete setting.

Goal of Correlation-Driven Super-resolution

Obtain a super-resolved image $\boldsymbol{p} \in \mathbb{R}^N$, where each pixel represents the source power, i.e. $p_i = E\left(|x_i|^2\right)$

Super-resolution Correlation-Imaging

- In many problems, the sources are assumed to be spatially incoherent

$$E\left(x_i x_j^*\right) = p_i \delta[i-j], \quad 1 \leqslant i, j \leqslant K$$

- Such assumptions are heavily exploited in correlation microscopy (e.g. SOFI, SPARCOM) to exploit the independent statistical fluctuation of fluoresecent emitters to aid super-resolution in the discrete setting.

Goal of Correlation-Driven Super-resolution

Obtain a super-resolved image $\boldsymbol{p} \in \mathbb{R}^N$, where each pixel represents the source power, i.e. $p_i = E\left(|x_i|^2\right)$

- Utilization of correlation priors can lead to significant improvement in super-resolution performance [Solomon,Eldar,Mutzafi,Segev'18].

Mathematical Theory of Correlation-Driven Super-resolution

Key Questions of Interest

- Can the separation condition be relaxed in correlation-driven Super-resolution?
- Can we tame the noise amplification (typically SRF^2) using correlation Priors?
- What roles will the geometry of spatial sampling (choice of $d_1, d_2, \cdots d_M$) and positivity play?
- What is the underlying trade-off between Spatial and Temporal Measurements?

Key Ingredient I: Khatri-Rao Product and Difference Set

$$\boldsymbol{R_{yy}} = \boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^H + \sigma^2\boldsymbol{I} \Longleftrightarrow \mathsf{vec}\left(\boldsymbol{R_{yy}}\right) = \left(\boldsymbol{A}^* \odot \boldsymbol{A}\right)\boldsymbol{p} + \sigma^2\mathsf{vec}(\boldsymbol{I})$$

Key Ingredient I: Khatri-Rao Product and Difference Set

$$\boldsymbol{R_{yy}} = \boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^H + \sigma^2\boldsymbol{I} \Longleftrightarrow \mathsf{vec}\left(\boldsymbol{R_{yy}}\right) = \left(\boldsymbol{A}^* \odot \boldsymbol{A}\right)\boldsymbol{p} + \sigma^2\mathsf{vec}(\boldsymbol{I})$$

Fact: Desired correlation image $\boldsymbol{p}$ is mapped to the data covariance $\boldsymbol{R_{yy}}$ via the Khatri-Rao product of $\boldsymbol{A}$:

$$\boldsymbol{A}^* \odot \boldsymbol{A} = [\boldsymbol{a}_1^* \otimes \boldsymbol{a}_1, \boldsymbol{a}_2^* \otimes \boldsymbol{a}_2, \cdots, \boldsymbol{a}_N^* \otimes \boldsymbol{a}_N]$$

Key Ingredient I: Khatri-Rao Product and Difference Set

$$\boldsymbol{R_{yy}} = \boldsymbol{APA}^H + \sigma^2 \boldsymbol{I} \Longleftrightarrow \text{vec}(\boldsymbol{R_{yy}}) = (\boldsymbol{A}^* \odot \boldsymbol{A})\,\boldsymbol{p} + \sigma^2 \text{vec}(\boldsymbol{I})$$

Fact: Desired correlation image $\boldsymbol{p}$ is mapped to the data covariance $\boldsymbol{R_{yy}}$ via the Khatri-Rao product of $\boldsymbol{A}$:

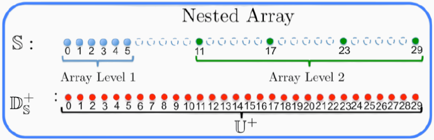
$$\boldsymbol{A}^* \odot \boldsymbol{A} = [\boldsymbol{a}_1^* \otimes \boldsymbol{a}_1, \boldsymbol{a}_2^* \otimes \boldsymbol{a}_2, \cdots, \boldsymbol{a}_N^* \otimes \boldsymbol{a}_N]$$

Difference Set

$\mathbb{S} = \{d_1, d_2, \cdots, d_M\}$

$\mathbb{D}_{\mathbb{S}} = \{d_m - d_n, \quad d_m.d_n \in \mathbb{S}\}$

$2M_{\text{diff}} + 1$ = cardinality of largest subset of consecutive integers in $\mathbb{D}_{\mathbb{S}}$



Key Ingredient I: Khatri-Rao Product and Difference Set

$$\boldsymbol{R_{yy}} = \boldsymbol{APA}^H + \sigma^2 \boldsymbol{I} \Longleftrightarrow \text{vec}\,(\boldsymbol{R_{yy}}) = (\boldsymbol{A}^* \odot \boldsymbol{A})\,\boldsymbol{p} + \sigma^2 \text{vec}(\boldsymbol{I})$$

Fact: Desired correlation image $\boldsymbol{p}$ is mapped to the data covariance $\boldsymbol{R_{yy}}$ via the Khatri-Rao product of $\boldsymbol{A}$:

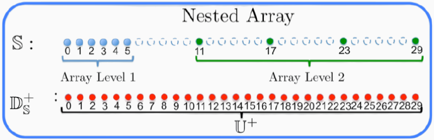
$$\boldsymbol{A}^* \odot \boldsymbol{A} = [\boldsymbol{a}_1^* \otimes \boldsymbol{a}_1, \boldsymbol{a}_2^* \otimes \boldsymbol{a}_2, \cdots, \boldsymbol{a}_N^* \otimes \boldsymbol{a}_N]$$

Difference Set

$\mathbb{S} = \{d_1, d_2, \cdots, d_M\}$

$\mathbb{D}_\mathbb{S} = \{d_m - d_n, \quad d_m.d_n \in \mathbb{S}\}$

$2M_{\text{diff}} + 1 =$ cardinality of largest subset of consecutive integers in $\mathbb{D}_\mathbb{S}$



- The quantity M_{diff} will be used to relax the separation condition, and reduce noise amplification in correlation-driven super-resolution.

Key Ingredient II: Role of Positive Constraints—Warm-Up

Solving an ill-posed system of equations (in $\boldsymbol{p}, \sigma$) :

$$\boldsymbol{R_{yy}} = \boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^H + \sigma^2 \boldsymbol{I} \tag{1}$$

[1] H. Qiao and P. Pal, "Guaranteed Localization of More Sources Than Sensors With Finite Snapshots in Multiple Measurement Vector Models Using Difference Co-Arrays," in IEEE Transactions on Signal Processing, vol. 67, no. 22, pp. 5715-5729, 15 Nov.15, 2019.

Key Ingredient II: Role of Positive Constraints—Warm-Up

Solving an ill-posed system of equations (in $\boldsymbol{p}, \sigma$) :

$$\boldsymbol{R_{yy}} = \boldsymbol{APA}^H + \sigma^2 \boldsymbol{I} \tag{1}$$

Asymptotic Unique Recovery without sparsity Constraints [Qiao,Pal2019]

As long as $\|\mathcal{S}\|_0 \leqslant M_{\text{diff}}$, there is a unique non negative pair $(\boldsymbol{p}, \sigma)$ that satisfies (1)

[1] H. Qiao and P. Pal, "Guaranteed Localization of More Sources Than Sensors With Finite Snapshots in Multiple Measurement Vector Models Using Difference Co-Arrays," in IEEE Transactions on Signal Processing, vol. 67, no. 22, pp. 5715-5729, 15 Nov.15, 2019.

Key Ingredient II: Role of Positive Constraints—Warm-Up

Solving an ill-posed system of equations (in $\boldsymbol{p}, \sigma$) :

$$\boldsymbol{R}_{\boldsymbol{yy}} = \boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^H + \sigma^2 \boldsymbol{I} \tag{1}$$

Asymptotic Unique Recovery without sparsity Constraints [Qiao,Pal2019]

As long as $\|\mathcal{S}\|_0 \leqslant M_{\text{diff}}$, there is a unique non negative pair $(\boldsymbol{p}, \sigma)$ that satisfies (1)

- No need for separation (asymptotically in number of snapshots L).
- Explicit Sparsity constraint not necessary.

[1] H. Qiao and P. Pal, "Guaranteed Localization of More Sources Than Sensors With Finite Snapshots in Multiple Measurement Vector Models Using Difference Co-Arrays," in IEEE Transactions on Signal Processing, vol. 67, no. 22, pp. 5715-5729, 15 Nov.15, 2019.

Key Ingredient II: Role of Positive Constraints—Warm-Up

Solving an ill-posed system of equations (in $\boldsymbol{p}, \sigma$) :

$$\boldsymbol{R_{yy}} = \boldsymbol{APA}^H + \sigma^2 \boldsymbol{I} \tag{1}$$

Asymptotic Unique Recovery without sparsity Constraints [Qiao,Pal2019]

As long as $\|\mathcal{S}\|_0 \leqslant M_{\text{diff}}$, there is a unique non negative pair $(\boldsymbol{p}, \sigma)$ that satisfies (1)

- No need for separation (asymptotically in number of snapshots L).
- Explicit Sparsity constraint not necessary.
- Proof Technique:
  - Lift to higher dimension: $\boldsymbol{R_{yy}} \to \boldsymbol{T} \in \mathbb{C}^{M_{\text{diff}} \times M_{\text{diff}}}, \boldsymbol{T} \succeq \boldsymbol{0}$, $\boldsymbol{T}$ is Toeplitz.
  - Invoke Caratheodory:

$$\sigma^2 = \sigma_{\min(\boldsymbol{T})} \tag{2}$$

$$(\boldsymbol{a}_i^* \otimes \boldsymbol{a}_i)_{\mathbb{U}} \perp \mathcal{N}(\boldsymbol{T} - \sigma^2 \boldsymbol{I}), \quad \forall i \in \mathcal{S} \tag{3}$$

[1] H. Qiao and P. Pal, "Guaranteed Localization of More Sources Than Sensors With Finite Snapshots in Multiple Measurement Vector Models Using Difference Co-Arrays," in IEEE Transactions on Signal Processing, vol. 67, no. 22, pp. 5715-5729, 15 Nov.15, 2019.

Non Asymptotic Guarantees

- In practice we have access to an estimate $\hat{\boldsymbol{R}}_L$ of the covariance matrix $\boldsymbol{R}_{\boldsymbol{yy}}$ computed using finite snapshots L.

$$\hat{\boldsymbol{R}}_L = \boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^H + \underbrace{\sigma^2\boldsymbol{I} + \boldsymbol{E}_L}_{\boldsymbol{\Delta}_L}$$

- **Key Questions of Interest:**
  - Noise + Finite snapshot error both can potentially degrade the ability to super-resolve.
  - Can (i) positivity of the desired correlation-image and (ii) geometry of sensing still lead to stable super-resolution with relaxed separation, and reduction in noise amplification?

Geometry of Conic Constraints

Feasible Set

$$\mathcal{F}_{\boldsymbol{\Delta}_L} = \left\{ \boldsymbol{z} \geqslant \mathbf{0}, \quad \left\| \mathsf{vec}(\hat{\boldsymbol{R}}_L) - (\boldsymbol{A}^* \odot \boldsymbol{A})\, \boldsymbol{z} \right\|_2 \leqslant \|\boldsymbol{\Delta}_L\|_F \right\}$$

- Feasible set $\mathcal{F}_{\boldsymbol{\Delta}_L}$ characterized by snapshots, and contains the true source power $\boldsymbol{p}$.

Geometry of Conic Constraints

Feasible Set

$$\mathcal{F}_{\boldsymbol{\Delta}_L} = \left\{ \boldsymbol{z} \geqslant \boldsymbol{0}, \quad \left\| \text{vec}(\hat{\boldsymbol{R}}_L) - (\boldsymbol{A}^* \odot \boldsymbol{A})\, \boldsymbol{z} \right\|_2 \leqslant \|\boldsymbol{\Delta}_L\|_F \right\}$$

- Feasible set $\mathcal{F}_{\boldsymbol{\Delta}_L}$ characterized by snapshots, and contains the true source power $\boldsymbol{p}$.
- Is it possible to bound the distance between any two points $\boldsymbol{z}_1, \boldsymbol{z}_2 \in \mathcal{F}_{\boldsymbol{\Delta}_L}$ in terms of $\|\boldsymbol{\Delta}_L\|_F$, despite $\boldsymbol{A}^* \odot \boldsymbol{A}$ being a fat matrix?
- Such a bound can lead us to universal stability guarantees for correlation-driven super-resolution.

Geometry of Conic Constraints

Feasible Set

$$\mathcal{F}_{\boldsymbol{\Delta}_L} = \left\{ \boldsymbol{z} \geqslant \boldsymbol{0}, \quad \left\| \text{vec}(\hat{\boldsymbol{R}}_L) - (\boldsymbol{A}^* \odot \boldsymbol{A})\, \boldsymbol{z} \right\|_2 \leqslant \|\boldsymbol{\Delta}_L\|_F \right\}$$

- Feasible set $\mathcal{F}_{\boldsymbol{\Delta}_L}$ characterized by snapshots, and contains the true source power $\boldsymbol{p}$.
- Is it possible to bound the distance between any two points $\boldsymbol{z}_1, \boldsymbol{z}_2 \in \mathcal{F}_{\boldsymbol{\Delta}_L}$ in terms of $\|\boldsymbol{\Delta}_L\|_F$, despite $\boldsymbol{A}^* \odot \boldsymbol{A}$ being a fat matrix?
- Such a bound can lead us to universal stability guarantees for correlation-driven super-resolution.
- Main challenge: $\boldsymbol{A}^* \odot \boldsymbol{A}$ has a non-trivial null-space.

Positivity to the Rescue

$$\mathcal{F}_{\boldsymbol{\Delta}_L} = \left\{ \boldsymbol{z} \geqslant \boldsymbol{0}, \quad \left\| \mathsf{vec}(\hat{\boldsymbol{R}}_L) - (\boldsymbol{A}^* \odot \boldsymbol{A})\, \boldsymbol{z} \right\|_2 \leqslant \|\boldsymbol{\Delta}_L\|_F \right\}$$

How does the conic constraint help?

Positivity to the Rescue

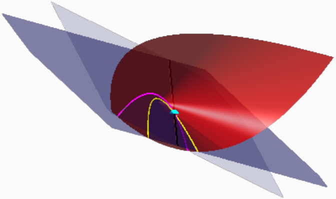
$$\mathcal{F}_{\mathbf{\Delta}_L} = \left\{ \boldsymbol{z} \geqslant \mathbf{0}, \quad \left\| \text{vec}(\hat{\boldsymbol{R}}_L) - (\boldsymbol{A}^* \odot \boldsymbol{A})\, \boldsymbol{z} \right\|_2 \leqslant \|\mathbf{\Delta}_L\|_F \right\}$$

How does the conic constraint help?

- Without non-negative constraint

$$\mathcal{B}_\epsilon = \left\{ \boldsymbol{z} \in \mathbb{R}^N, \left\| \text{vec}(\boldsymbol{R}_{\boldsymbol{yy}}) - \left(\boldsymbol{A}^* \odot \boldsymbol{A} \right) \boldsymbol{z} \right\|_2 \leqslant \|\mathbf{\Delta}_L\|_F \right\}$$

- Let $\boldsymbol{p} \in \mathcal{B}$ and let $\boldsymbol{z}_1 = \boldsymbol{p} + \alpha \boldsymbol{v}$, where $\boldsymbol{v} \in \mathcal{N}(\boldsymbol{A}^* \odot \boldsymbol{A})$. Then $\boldsymbol{z}_1 \in \mathcal{B}$ but $\|\boldsymbol{p} - \boldsymbol{z}_1\|$ diverges with α.
- Geometry of conic constraint crucial to make $\mathcal{F}_{\mathbf{\Delta}_L}$ bounded.



Stability of Convex Feasibility Test

Definition

Define the set of sparse signals obeying relaxed Difference-Set Separation (DS-SEP) condition as

$$\mathcal{P}_{\text{DS-SEP}} \triangleq \{\mathbf{p} \in \mathbb{C}^N \mid \phi(\frac{k}{N}, \frac{l}{N}) \geqslant \frac{2}{M_{\text{diff}}}, \forall k \neq l \in \text{Supp}(\mathbf{p})\}$$

Stability of Convex Feasibility Test

Definition

Define the set of sparse signals obeying relaxed Difference-Set Separation (DS-SEP) condition as

$$\mathcal{P}_{\text{DS-SEP}} \triangleq \{\mathbf{p} \in \mathbb{C}^N \mid \phi(\frac{k}{N}, \frac{l}{N}) \geqslant \frac{2}{M_{\text{diff}}}, \forall k \neq l \in \text{Supp}(\mathbf{p})\}$$

Theorem (Qiao,Pal.19)

Suppose the ground truth $\mathbf{p}$ *satisfies the relaxed difference-set separation condition, i.e.* $\mathbf{p} \in \mathcal{P}_{\text{DS-SEP}}$. *Further suppose* $M_{\text{diff}} \geqslant 128$ *and* $N \geqslant 3.03(2M_{\text{diff}} + 1)$. *Then, for any* $\mathbf{p}^{\#} \in \mathcal{F}_{\mathbf{\Delta}_L}$, *we have*

$$\|\mathbf{p}^{\#} - \mathbf{p}\|_1 = O\left(\frac{1-\rho}{\rho}\|\mathbf{\Delta}_L\|_F\right) \quad (4)$$

where $\rho = c_1\left(\frac{M_{\text{diff}}}{N}\right)^2$, c_1 *being a universal constant.*

H. Qiao and P. Pal, "Guaranteed Localization of More Sources Than Sensors With Finite Snapshots in Multiple Measurement Vector Models Using Difference Co-Arrays," in IEEE Transactions on Signal Processing, vol. 67, no. 22, pp. 5715-5729, 15 Nov.15, 2019.

Significance Of the Bound: Universal Stability in Correlation-driven super-resolution

Consider the Feasibility Problem

$$\begin{aligned} \text{find} \quad & \boldsymbol{z} \\ \text{subject to} \quad & \|\text{vec}(\boldsymbol{R_{yy}}) - (\boldsymbol{A}^* \odot \boldsymbol{A})\,\boldsymbol{z}\|_2 \leqslant \|\boldsymbol{\Delta}_L\|_F, \\ & \boldsymbol{z} \succeq 0. \end{aligned} \tag{FEAS}$$

Significance Of the Bound: Universal Stability in Correlation-driven super-resolution

Consider the Feasibility Problem

$$\begin{aligned} \text{find} \quad & \boldsymbol{z} \\ \text{subject to} \quad & \|\text{vec}(\boldsymbol{R_{yy}}) - (\boldsymbol{A}^* \odot \boldsymbol{A})\, \boldsymbol{z}\|_2 \leqslant \|\boldsymbol{\Delta}_L\|_F, \\ & \boldsymbol{z} \geq 0. \end{aligned} \tag{FEAS}$$

- Any solution $\boldsymbol{z}^*$ to (FEAS) will satisfy $\|\mathbf{z}^* - \mathbf{p}\|_1 = O\left(\frac{1-\rho}{\rho}\|\boldsymbol{\Delta}_L\|_F\right)$
- Captures how correlation estimation error $\|\boldsymbol{\Delta}_L\|_F$ controls the (worst-case) reconstruction error.

Significance Of the Bound: Universal Stability in Correlation-driven super-resolution

Consider the Feasibility Problem

$$\begin{aligned} &\text{find} && \boldsymbol{z} \\ &\text{subject to} && \|\text{vec}(\boldsymbol{R_{yy}}) - (\boldsymbol{A}^* \odot \boldsymbol{A})\,\boldsymbol{z}\|_2 \leqslant \|\boldsymbol{\Delta}_L\|_F, \\ &&& \boldsymbol{z} \geq 0. \end{aligned} \tag{FEAS}$$

- Any solution $\boldsymbol{z}^*$ to (FEAS) will satisfy $\|\mathbf{z}^* - \mathbf{p}\|_1 = O\left(\frac{1-\rho}{\rho}\|\boldsymbol{\Delta}_L\|_F\right)$
- Captures how correlation estimation error $\|\boldsymbol{\Delta}_L\|_F$ controls the (worst-case) reconstruction error.
- Algorithm-independent upper bound on the reconstruction error, depending only on the geometry of the Feasible set $\mathcal{F}_{\boldsymbol{\Delta}_L}$. Universal benchmakr to determine objective functions can do better than picking arbitrary point from Feasible set.

Error Amplification Can be Reduced

$$\begin{aligned} \|\mathbf{p}^{\#} - \mathbf{p}\|_1 &= O\left(\frac{1-\rho}{\rho}\|\mathbf{\Delta}_L\|_F\right) \\ \rho &= c_1\left(\frac{M_{\mathsf{diff}}}{N}\right)^2 \end{aligned} \tag{5}$$

- Covariance estimation error gets scaled by a factor of

$$\frac{1}{\rho} = c_1'\left(\frac{N}{M_{\mathsf{diff}}}\right)^2$$

Error Amplification Can be Reduced

$$\|\mathbf{p}^{\#} - \mathbf{p}\|_1 = O\left(\frac{1-\rho}{\rho}\|\boldsymbol{\Delta}_L\|_F\right)$$
$$\rho = c_1 \left(\frac{M_{\text{diff}}}{N}\right)^2 \tag{5}$$

- Covariance estimation error gets scaled by a factor of

$$\frac{1}{\rho} = c_1' \left(\frac{N}{M_{\text{diff}}}\right)^2$$

- **Role of Sensing Geometry:**
  - $M_{\text{diff}} = \Theta(M)$, corresponds to ULA: $\|\boldsymbol{\Delta}_L\|_F$ amplifies by $\text{SRF}^2 = \left(\frac{N}{M}\right)^2$. Similar to existing analysis.

Error Amplification Can be Reduced

$$
\begin{aligned}
\|\mathbf{p}^{\#} - \mathbf{p}\|_1 &= O\left(\frac{1-\rho}{\rho}\|\mathbf{\Delta}_L\|_F\right) \\
\rho &= c_1 \left(\frac{M_{\mathsf{diff}}}{N}\right)^2
\end{aligned}
\tag{5}
$$

- Covariance estimation error gets scaled by a factor of

$$
\frac{1}{\rho} = c_1' \left(\frac{N}{M_{\mathsf{diff}}}\right)^2
$$

- **Role of Sensing Geometry:**
  - $M_{\mathsf{diff}} = \Theta(M)$, corresponds to ULA: $\|\mathbf{\Delta}_L\|_F$ amplifies by $\mathsf{SRF}^2 = \left(\frac{N}{M}\right)^2$. Similar to existing analysis.
  - $M_{\mathsf{diff}} = \Theta(M^2)$, corresponds to sparse arrays: Covariance error scales by $\frac{N^2}{M^4}$.

Error Amplification Can be Reduced

$$\|\mathbf{p}^{\#} - \mathbf{p}\|_1 = O\left(\frac{1-\rho}{\rho}\|\mathbf{\Delta}_L\|_F\right)$$
$$\rho = c_1 \left(\frac{M_{\text{diff}}}{N}\right)^2 \tag{5}$$

- Covariance estimation error gets scaled by a factor of

$$\frac{1}{\rho} = c_1' \left(\frac{N}{M_{\text{diff}}}\right)^2$$

- **Role of Sensing Geometry:**
  - $M_{\text{diff}} = \Theta(M)$, corresponds to ULA: $\|\mathbf{\Delta}_L\|_F$ amplifies by $\text{SRF}^2 = \left(\frac{N}{M}\right)^2$. Similar to existing analysis.
  - $M_{\text{diff}} = \Theta(M^2)$, corresponds to sparse arrays: Covariance error scales by $\frac{N^2}{M^4}$.
  - Covariance error can be potentially compensated in the final correlation image, thanks to the large difference set of sparse arrays, as long as $N = o(M^2)$.

Tightness of the Amplification Factor

Amplification is quadratic in N: $\frac{1}{\rho} \sim \frac{N^2}{M_{\text{diff}}^2}$.

Is the quadratic scaling tight?

[1] H. Qiao and P. Pal, "Guaranteed Localization of More Sources Than Sensors With Finite Snapshots in Multiple Measurement Vector Models Using Difference Co-Arrays," in IEEE Transactions on Signal Processing, vol. 67, no. 22, pp. 5715-5729, 15 Nov.15, 2019.

Tightness of the Amplification Factor

Amplification is quadratic in N: $\frac{1}{\rho} \sim \frac{N^2}{M_{\text{diff}}^2}$.

Is the quadratic scaling tight?

Theorem (Qiao,Pal19)

There exist $\boldsymbol{p}_1, \boldsymbol{p}_2 \in \mathcal{F}_{\boldsymbol{\Delta}_L}$ (with $\boldsymbol{p}_1$ also obeying separation condition) such that whenever $M_{diff} \geqslant 128$ and $N \geqslant 3.03(2M_{diff} + 1)$, we have

$$\|\boldsymbol{p}_1 - \boldsymbol{p}_2\|_1 \leqslant C_1(M) N^2 \|\boldsymbol{\Delta}_L\|_F$$

and

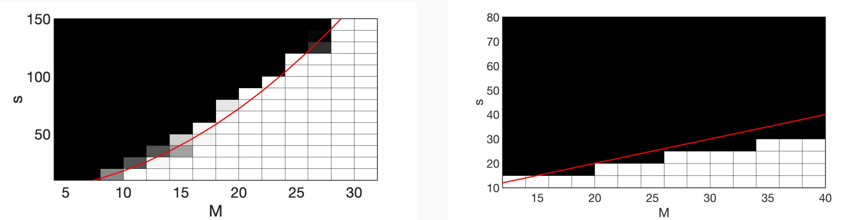
$$\|\boldsymbol{p}_1 - \boldsymbol{p}_2\|_1 \geqslant C_2(M) N^2 \|\boldsymbol{\Delta}_L\|_F$$

where $C_1(M)$ and $C_2(M)$ are only functions of M.

[1] H. Qiao and P. Pal, "Guaranteed Localization of More Sources Than Sensors With Finite Snapshots in Multiple Measurement Vector Models Using Difference Co-Arrays," in IEEE Transactions on Signal Processing, vol. 67, no. 22, pp. 5715-5729, 15 Nov.15, 2019.

Numerical Results 1

Phase Transition and Sample Complexity

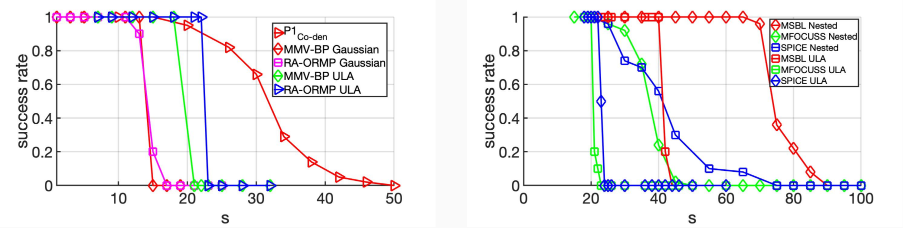


(a) (b)

Figure 1: Phase transition of success rate as function of sparsity s and number of measurements M: (a) $(P1_{\text{Co-den}})$, (b) MMV-BP. White pixels indicate perfect recovery and black pixels denote total failure. Here $L = 2000, N = 600$ and the results are averaged over 50 runs. The overlaid red curve represents $s = 0.18M^2$ in (a) and $s = M$ in (b) and (c).

Numerical Results 2

Empirical Support Recovery versus Sparsity



(a) (b)

Figure 2: (a) Probability of successful support recovery as a function of sparsity s. (b) Success rate of M-SBL, M-FOCUSS and SPICE as a function of sparsity s. For both cases, $M = 24, N = 300, L = 100$.

Super-Resolution via Parameter Estimation: Going Off the Grid

Super-Resolution and Line Spectrum Estimation

Measurement Model:

$$\boldsymbol{y}_l = \sum_{k=1}^{K} \boldsymbol{a}(\omega_k) c_{k,l} + \boldsymbol{n}_l, \quad l = 1, 2, \cdots L$$

- $\boldsymbol{y}_l \in \mathbb{C}^M$ — lth temporal snapshot of measurements collected by an array of M sensors.
- $\boldsymbol{a}(\omega) \in \mathbb{C}^M$ — steering vector of the array corresponding to spatial frequency ω.
- $c_{k,l}$ — (Time varying) amplitude of the kth source
- $\boldsymbol{n}_l$— Additive noise at the sensor array.
- Model is widely adopted for the problem of point source localization.

Goal: Recover $\{\omega_k\}_{k=1}^{K}$ from measurements $\boldsymbol{y}_l$

Atomic Norm Minimization: Basics

- Point source model: $x(t) = \sum_{k=1}^{K} c_k \delta(t - t_k), \tau_k \in [0, 1)$

[2] For an arbitrary point-spread function $g(t)$ bandlimited to $|f| \leqslant B/2$, the Fourier-domain measurement model has been typically modified as [Chi '16,'20]

$$y_m = \int e^{j2\pi\omega_m t}(g * x)(t)dt + n_m = \sum_{k=1}^{K} c_k e^{j2\pi\omega_m \tau_k} \hat{g}_\omega + n_m,\ \omega_m \in [-B/2, \cdots, B/2]$$

However, as argued in [Batenkov,Bhandari,Blu'19],[Chen,Moitra'20], bandwidth selection is an issue, and the frequency domain model may not be fully representative of the actual physical measurements.

Atomic Norm Minimization: Basics

- Point source model: $x(t) = \sum_{k=1}^{K} c_k \delta(t - t_k), \tau_k \in [0, 1)$
- Bandlimited Measurement Model: [2]

$$y_m = \int_0^1 e^{j2\pi mt} x(t) + n_m = \sum_{k=1}^{K} c_k e^{j2\pi m t_k} + n_m,\ |m| \leqslant M/2$$

[2] For an arbitrary point-spread function $g(t)$ bandlimited to $|f| \leqslant B/2$, the Fourier-domain measurement model has been typically modified as [Chi '16,'20]

$$y_m = \int e^{j2\pi\omega_m t} (g * x)(t) dt + n_m = \sum_{k=1}^{K} c_k e^{j2\pi\omega_m \tau_k} \hat{g}_\omega + n_m,\ \omega_m \in [-B/2, \cdots, B/2]$$

However, as argued in [Batenkov,Bhandari,Blu'19],[Chen,Moitra'20], bandwidth selection is an issue, and the frequency domain model may not be fully representative of the actual physical measurements.

Atomic Norm Minimization: Basics

- Point source model: $x(t) = \sum_{k=1}^{K} c_k \delta(t - t_k), \tau_k \in [0, 1)$
- Bandlimited Measurement Model: [2]

$$y_m = \int_0^1 e^{j2\pi mt} x(t) + n_m = \sum_{k=1}^{K} c_k e^{j2\pi m t_k} + n_m,\ |m| \leqslant M/2$$

- Atomic Set: $\mathcal{A} = \{e^{j2\pi\phi}[1, e^{j2\pi\tau}, e^{j2\pi 2\tau}, \cdots, e^{j2\pi(M-1)\tau}], \phi, \tau \in [0, 1)\}$

[2] For an arbitrary point-spread function $g(t)$ bandlimited to $|f| \leqslant B/2$, the Fourier-domain measurement model has been typically modified as [Chi '16,'20]

$$y_m = \int e^{j2\pi\omega_m t} (g * x)(t) dt + n_m = \sum_{k=1}^{K} c_k e^{j2\pi\omega_m \tau_k} \hat{g}_\omega + n_m,\ \omega_m \in [-B/2, \cdots, B/2]$$

However, as argued in [Batenkov,Bhandari,Blu'19],[Chen,Moitra'20], bandwidth selection is an issue, and the frequency domain model may not be fully representative of the actual physical measurements.

Atomic Norm Minimization: Basics

- Point source model: $x(t) = \sum_{k=1}^{K} c_k \delta(t - t_k), \tau_k \in [0, 1)$
- Bandlimited Measurement Model: [2]

$$y_m = \int_0^1 e^{j2\pi m t} x(t) + n_m = \sum_{k=1}^{K} c_k e^{j2\pi m t_k} + n_m,\ |m| \leqslant M/2$$

- Atomic Set: $\mathcal{A} = \{e^{j2\pi\phi}[1, e^{j2\pi\tau}, e^{j2\pi 2\tau}, \cdots, e^{j2\pi(M-1)\tau}], \phi, \tau \in [0, 1)\}$
- Atomic Norm: $\|\boldsymbol{x}\|_{\mathcal{A}} := \inf\{t \geqslant 0, \boldsymbol{x} \in t\,.\,\mathrm{conv}(\mathcal{A})\}$

TV or atomic norm minimization rely on a "separation condition" between spikes/sources for developing theoretical guarantees.

[2] For an arbitrary point-spread function $g(t)$ bandlimited to $|f| \leqslant B/2$, the Fourier-domain measurement model has been typically modified as [Chi '16,'20]

$$y_m = \int e^{j2\pi\omega_m t} (g * x)(t) dt + n_m = \sum_{k=1}^{K} c_k e^{j2\pi\omega_m \tau_k} \hat{g}_\omega + n_m,\ \omega_m \in [-B/2, \cdots, B/2]$$

However, as argued in [Batenkov,Bhandari,Blu'19],[Chen,Moitra'20], bandwidth selection is an issue, and the frequency domain model may not be fully representative of the actual physical measurements.

Atomic Norm Denoising and Separation Condition

$$\min_{\boldsymbol{x}} \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{x}\|_{\boldsymbol{Z}}^2 + \lambda\|\boldsymbol{x}\|_{\mathcal{A}}$$

- **Separation Condition:** $\Delta := \min_{i \neq j} \phi(\tau_i, \tau_j) > \frac{c}{M}$ (wrap-around distance)

Atomic Norm Denoising and Separation Condition

$$\min_{\boldsymbol{x}} \frac{1}{2}\|\boldsymbol{y}-\boldsymbol{x}\|_{\boldsymbol{Z}}^2 + \lambda\|\boldsymbol{x}\|_{\mathcal{A}}$$

- **Separation Condition:** $\Delta := \min_{i \neq j} \phi(\tau_i, \tau_j) > \frac{c}{M}$ (wrap-around distance)

Recovery Guarantee [Li,Tang 2020]

Assume that the noise $\boldsymbol{n}$ is zero mean Gaussian with independent entries and variance σ. If Separation condition holds, the complex amplitudes c_k have approximately the same magnitude, and $\boldsymbol{Z}, \lambda$ are suitably chosen, then

$$|c_k||\tau_k - \hat{\tau}_K| = O\left(\sigma \frac{\sqrt{\log M}}{M^{3/2}}\right), |c_k - \hat{c}_k| = O\left(\sigma\sqrt{\frac{\log M}{M}}\right)$$

Atomic Norm Denoising and Separation Condition

$$\min_{\boldsymbol{x}} \frac{1}{2}\|\boldsymbol{y}-\boldsymbol{x}\|_{\boldsymbol{Z}}^2 + \lambda\|\boldsymbol{x}\|_{\mathcal{A}}$$

- **Separation Condition:** $\Delta := \min_{i \neq j} \phi(\tau_i, \tau_j) > \frac{c}{M}$ (wrap-around distance)

Recovery Guarantee [Li,Tang 2020]

Assume that the noise $\boldsymbol{n}$ is zero mean Gaussian with independent entries and variance σ. If Separation condition holds, the complex amplitudes c_k have approximately the same magnitude, and $\boldsymbol{Z}, \lambda$ are suitably chosen, then

$$|c_k||\tau_k - \hat{\tau}_K| = O\left(\sigma \frac{\sqrt{\log M}}{M^{3/2}}\right), |c_k - \hat{c}_k| = O\left(\sigma \sqrt{\frac{\log M}{M}}\right)$$

- Separation condition is needed even in *noiseless setting* and is shown to be necessary for success of atomic and TV norm minimization [Da Costa,Dai'18],[Fernandez-Granda'18,'20].

Revisiting Separation Condition

- Role of additional Measurements available due to temporal dimension.
- Role of correlation priors (or sources being statistically independent)?

Revisiting Separation Condition

- Role of additional Measurements available due to temporal dimension.
- Role of correlation priors (or sources being statistically independent)?
- Can correlation priors lead us to fundamentally relax the separation condition, and re-parameterize it by bringing out the integrated effect of number of temporal measurements, noise power in addition to spatial measurements?

Correlation Priors and Sparse Arrays

Sources are statistically uncorrelated: $E(c_j c_k^*) = p_k \delta[j-k]$

Physical Array

- Measurement Covariance Matrix: $\boldsymbol{R_{yy}} = \boldsymbol{S} \boldsymbol{T}_{\text{diff}} \boldsymbol{S}^T$
- $\boldsymbol{R_{yy}} \in \mathbb{C}^{M \times M}$ is Toeplitz for a ULA, not Toeplitz for sparse arrays.

Difference Co-Array

- Difference-set Covariance Matrix $\boldsymbol{T}_{\text{diff}} \in \mathbb{C}^{M_{\text{diff}} \times M_{\text{diff}}}$ is Toeplitz, and $\boldsymbol{T}_{\text{diff}} \succeq \boldsymbol{0}$.
- $\boldsymbol{T}_{\text{diff}} = \mathcal{A}_{\text{Toeplitz}}(\boldsymbol{R_{yy}})$

Correlation Priors and Sparse Arrays

Sources are statistically uncorrelated: $E(c_j c_k^*) = p_k \delta[j-k]$

Physical Array

- Measurement Covariance Matrix: $\boldsymbol{R_{yy}} = \boldsymbol{S}\boldsymbol{T}_{\text{diff}}\boldsymbol{S}^T$
- $\boldsymbol{R_{yy}} \in \mathbb{C}^{M\times M}$ is Toeplitz for a ULA, not Toeplitz for sparse arrays.

Difference Co-Array

- Difference-set Covariance Matrix $\boldsymbol{T}_{\text{diff}} \in \mathbb{C}^{M_{\text{diff}}\times M_{\text{diff}}}$ is Toeplitz, and $\boldsymbol{T}_{\text{diff}} \succeq \mathbf{0}$.
- $\boldsymbol{T}_{\text{diff}} = \mathcal{A}_{\text{Toeplitz}}(\boldsymbol{R_{yy}})$

- Difference-set based super-resolution methods utilize the subspace-structure of $\boldsymbol{T}_{\text{diff}}$ (and the large difference set of sparse arrays) to recover $\{\omega_i\}_{i=1}^K$
- Can correlation priors and **temporal** measurements help overcome the need for a strict separation condition ($\Delta > \frac{c}{M}$) which is dictated only by the number M of **spatial** measurements ?

Analyzing Co-array Super-resolution with Spatiotemporal Measurements

Theorem [Hucumenoglu,P.20]

Suppose $\sigma_K^2(\boldsymbol{A}^* \odot \boldsymbol{A}) > \frac{\sigma^2}{p_{\min}}$. Given any $\epsilon > 0$, and $0 < \delta < 1$, the matching distance error in frequency estimation by co-array ESPRIT satisfies $\mathrm{md}(\boldsymbol{\omega}, \hat{\boldsymbol{\omega}}) \leqslant \epsilon$ with probability at least $1 - \delta$ if

$$T \geqslant \max\left(T_0, \frac{\phi(M_{\mathsf{diff}}, \sigma, K)}{\epsilon^2 \log \delta}\right)$$

Analyzing Co-array Super-resolution with Spatiotemporal Measurements

Theorem [Hucumenoglu,P.20]

Suppose $\sigma_K^2(\boldsymbol{A}^* \odot \boldsymbol{A}) > \frac{\sigma^2}{p_{\min}}$. Given any $\epsilon > 0$, and $0 < \delta < 1$, the matching distance error in frequency estimation by co-array ESPRIT satisfies $\mathrm{md}(\boldsymbol{\omega}, \hat{\boldsymbol{\omega}}) \leqslant \epsilon$ with probability at least $1 - \delta$ if

$$T \geqslant \max\left(T_0, \frac{\phi(M_{\mathsf{diff}}, \sigma, K)}{\epsilon^2 \log \delta}\right)$$

- The condition $\sigma_K^2(\boldsymbol{A}^* \odot \boldsymbol{A}) > \frac{\sigma^2}{p_{\min}}$ can be simplified to produce a minimum separation condition that depends on both M_{diff} and SNR.

Analyzing Co-array Super-resolution with Spatiotemporal Measurements

Theorem [Hucumenoglu,P.20]

Suppose $\sigma_K^2(\boldsymbol{A}^* \odot \boldsymbol{A}) > \frac{\sigma^2}{p_{\min}}$. Given any $\epsilon > 0$, and $0 < \delta < 1$, the matching distance error in frequency estimation by co-array ESPRIT satisfies $\mathrm{md}(\boldsymbol{\omega}, \hat{\boldsymbol{\omega}}) \leqslant \epsilon$ with probability at least $1 - \delta$ if

$$T \geqslant \max\left(T_0, \frac{\phi(M_{\mathsf{diff}}, \sigma, K)}{\epsilon^2 \log \delta}\right)$$

- The condition $\sigma_K^2(\boldsymbol{A}^* \odot \boldsymbol{A}) > \frac{\sigma^2}{p_{\min}}$ can be simplified to produce a minimum separation condition that depends on both M_{diff} and SNR.
- The number of snapshots needs to be larger than a threshold T_0 that depends on the minimum separation Δ, number of sources K, M_{diff} and SNR.

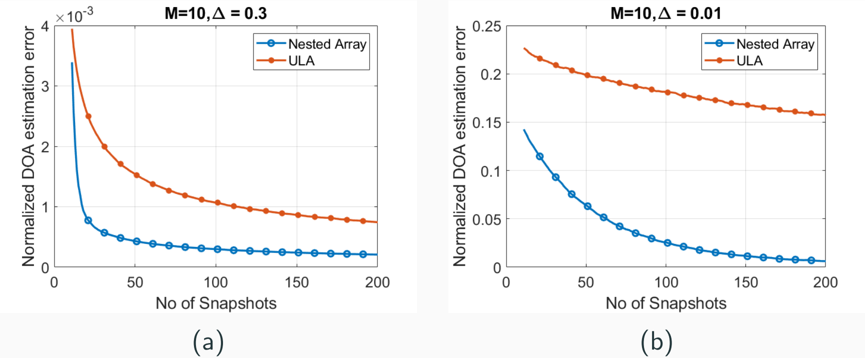
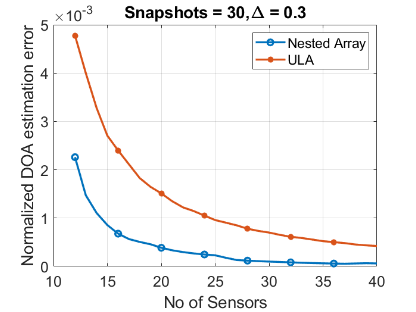
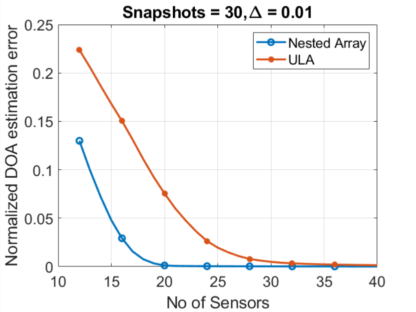


Figure 3: Comparison of DoA Estimation error of Nested Array and ULA as a function of L for (a) $\Delta = 0.3$ and (b) $\Delta = 0.01$

Numerical Results: Frequency Error and Separation



(a)



(b)

Figure 4: Comparison of DoA Estimation error of Nested Array and ULA as a function of M for (a) $\Delta = 0.3$ and (b) $\Delta = 0.01$

Numerical Results: MUSIC Spectrum as a function of Separation

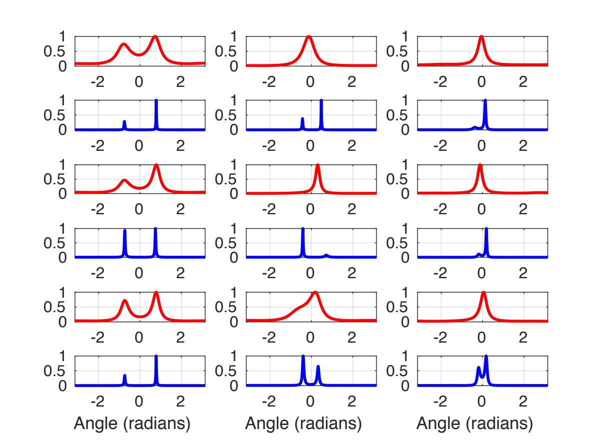


Figure 5: MUSIC Spectrum of ULA (red) and a Nested array (blue). The SNR varies row-wise with values $\{-1, -0.5, 0\}$ dB. Source separation varies column-wise with values $\{\frac{\pi}{4}, \frac{\pi}{8}, \frac{\pi}{16}\}$

A note on Covariance Estimation and Frequency Estimation Error with Sparse Arrays

Let $\mathbb{S}$ denote the set of sensor locations. Let $\hat{\boldsymbol{T}}_{\text{diff},\mathbb{S}}$ be an estimate of the co-array covariance matrix, obtained by spatially averaging entries of $\hat{\boldsymbol{R}}_L$.

- **Nested Geometry with M Sensors:** $\|\hat{\boldsymbol{T}}_{\text{diff,nest}} - \boldsymbol{T}_{\text{diff,nest}}\|_2 \leqslant \epsilon \|\boldsymbol{T}_{\text{diff,nest}}\|_2$ with probability at least $1 - \delta$ if $L \geqslant c_1 \frac{M_{\text{diff,nest}} \log(M_{\text{diff,nest}}/\epsilon\delta)}{\epsilon^2}$

A note on Covariance Estimation and Frequency Estimation Error with Sparse Arrays

Let $\mathbb{S}$ denote the set of sensor locations. Let $\hat{\boldsymbol{T}}_{\text{diff},\mathbb{S}}$ be an estimate of the co-array covariance matrix, obtained by spatially averaging entries of $\hat{\boldsymbol{R}}_L$.

- **Nested Geometry with M Sensors:** $\|\hat{\boldsymbol{T}}_{\text{diff,nest}} - \boldsymbol{T}_{\text{diff,nest}}\|_2 \leqslant \epsilon \|\boldsymbol{T}_{\text{diff,nest}}\|_2$ with probability at least $1-\delta$ if $L \geqslant c_1 \frac{M_{\text{diff,nest}} \log(M_{\text{diff,nest}}/\epsilon\delta)}{\epsilon^2}$
- **Uniform Geometry with M Sensors:** $\|\hat{\boldsymbol{T}}_{\text{diff,ULA}} - \boldsymbol{T}_{\text{diff,ULA}}\|_2 \leqslant \epsilon \|\boldsymbol{T}_{\text{diff,ULA}}\|_2$ with probability at least $1-\delta$ if $L \geqslant c_1 \frac{\log M_{\text{diff,ULA}} \log(M_{\text{diff,ULA}}/\epsilon\delta)}{\epsilon^2}$

A note on Covariance Estimation and Frequency Estimation Error with Sparse Arrays

Let $\mathbb{S}$ denote the set of sensor locations. Let $\hat{\boldsymbol{T}}_{\text{diff},\mathbb{S}}$ be an estimate of the co-array covariance matrix, obtained by spatially averaging entries of $\hat{\boldsymbol{R}}_L$.

- **Nested Geometry with M Sensors:** $\|\hat{\boldsymbol{T}}_{\text{diff,nest}} - \boldsymbol{T}_{\text{diff,nest}}\|_2 \leqslant \epsilon \|\boldsymbol{T}_{\text{diff,nest}}\|_2$ with probability at least $1-\delta$ if $L \geqslant c_1 \frac{M_{\text{diff,nest}} \log(M_{\text{diff,nest}}/\epsilon\delta)}{\epsilon^2}$
- **Uniform Geometry with M Sensors:** $\|\hat{\boldsymbol{T}}_{\text{diff,ULA}} - \boldsymbol{T}_{\text{diff,ULA}}\|_2 \leqslant \epsilon \|\boldsymbol{T}_{\text{diff,ULA}}\|_2$ with probability at least $1-\delta$ if $L \geqslant c_1 \frac{\log M_{\text{diff,ULA}} \log(M_{\text{diff,ULA}}/\epsilon\delta)}{\epsilon^2}$

Estimating the co-array covariance matrix (by simple sample averaging) entails *higher error for sparse arrays* for a given budget of spatial (M) and temporal (L) measurements.

Is this true for frequency estimation error as well?

Covariance versus Frequency Estimation: A reversal of Trend

- Study the Cramér-Rao Bound for covariance versus frequency estimation from measurements

$$\boxed{\boldsymbol{y}_l = \sum_{k=1}^{K} \boldsymbol{a}(\omega_k) c_{k,l} + \boldsymbol{n}_l}$$

Covariance versus Frequency Estimation: A reversal of Trend

- Study the Cramér-Rao Bound for covariance versus frequency estimation from measurements

$$\boldsymbol{y}_l = \sum_{k=1}^{K} \boldsymbol{a}(\omega_k) c_{k,l} + \boldsymbol{n}_l$$

Covariance Estimation

$$\boldsymbol{y}_l \sim \mathcal{CN}\left(\boldsymbol{0}, \boldsymbol{S}\boldsymbol{T}_{\text{diff}}\boldsymbol{S}^T\right)$$

Parameter: $\boldsymbol{\theta} = [\boldsymbol{T}_{\text{diff}}]$

Frequency Estimation

$$\boldsymbol{y}_l \sim \mathcal{CN}\left(\boldsymbol{0}, \boldsymbol{A}(\boldsymbol{\omega})\boldsymbol{P}\boldsymbol{A}^H(\boldsymbol{\omega}) + \sigma^2 \boldsymbol{I}\right)$$

Parameter: $\boldsymbol{\theta} = [\{\omega_k, p_k\}_{k=1}^K, \sigma]$

Covariance versus Frequency Estimation: A reversal of Trend

- Study the Cramér-Rao Bound for covariance versus frequency estimation from measurements

$$\boldsymbol{y}_l = \sum_{k=1}^{K} \boldsymbol{a}(\omega_k) c_{k,l} + \boldsymbol{n}_l$$

Covariance Estimation

$$\boldsymbol{y}_l \sim \mathcal{CN}\left(\boldsymbol{0}, \boldsymbol{S}\boldsymbol{T}_{\text{diff}}\boldsymbol{S}^T\right)$$

Parameter: $\boldsymbol{\theta} = [\boldsymbol{T}_{\text{diff}}]$

Frequency Estimation

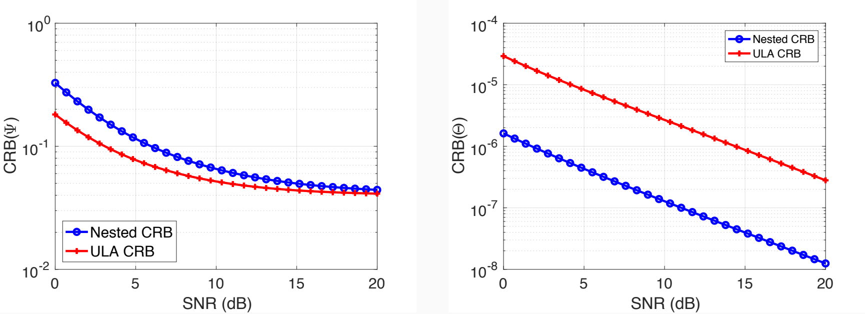
$$\boldsymbol{y}_l \sim \mathcal{CN}\left(\boldsymbol{0}, \boldsymbol{A}(\boldsymbol{\omega})\boldsymbol{P}\boldsymbol{A}^H(\boldsymbol{\omega}) + \sigma^2\boldsymbol{I}\right)$$

Parameter: $\boldsymbol{\theta} = [\{\omega_k, p_k\}_{k=1}^{K}, \sigma]$

$$[\boldsymbol{J}_{\boldsymbol{\theta}}]_{m,n} = \text{vec}^H\left(\frac{\partial \boldsymbol{R}_{\boldsymbol{yy}}(\boldsymbol{\theta})}{\partial \theta_m}\right)\boldsymbol{F}(\boldsymbol{\theta})\text{vec}\left(\frac{\partial \boldsymbol{R}_{\boldsymbol{yy}}(\boldsymbol{\theta})}{\partial \Psi_n}\right), \boldsymbol{F}(\boldsymbol{\theta}) = \boldsymbol{R}(\boldsymbol{\theta})^{-T} \otimes \boldsymbol{R}(\boldsymbol{\theta})^{-1}$$

Cramér-Rao Bound of Covariance versus Frequency Estimation

- Number of antennas $M = 10$
- Number of sources $K = 4$
- Number of snapshots $L = 1000$



(a) (b)

Figure 6: CRB for (a) Estimating $\boldsymbol{T}_{\text{diff}}$ (b) AOA Estimation

Conclusion

Conclusions

- Noisy super-resolution is a challenging task. Utilization of appropriate priors can make a significant difference.

Conclusions

- Noisy super-resolution is a challenging task. Utilization of appropriate priors can make a significant difference.
- Sensing geometry and algorithms can work hand-in-hand to overcome limitations of existing techniques.

Conclusions

- Noisy super-resolution is a challenging task. Utilization of appropriate priors can make a significant difference.
- Sensing geometry and algorithms can work hand-in-hand to overcome limitations of existing techniques.
- "Resolvability" of point sources depends on both spatial and temporal measurements. Judicious use of temporal measurements can significantly improve achievable resolution.

Conclusions

- Noisy super-resolution is a challenging task. Utilization of appropriate priors can make a significant difference.
- Sensing geometry and algorithms can work hand-in-hand to overcome limitations of existing techniques.
- “Resolvability” of point sources depends on both spatial and temporal measurements. Judicious use of temporal measurements can significantly improve achievable resolution.
- These results can be generalized to incorporate different types of PSFs.