Sparse Signal Reconstruction in Bilinear Inverse Problems

Yoram Bresler

Coordinated Science Laboratory and the Department of ECE
University of Illinois at Urbana-Champaign

Work with
• Yanjun Li
• Kiryung Lee
• Marius Junge

Acknowledgements: NSF CCF 10-18789, DMS 12-01886, and IIS 14-47879

SPARS 2017
From Sensing to Data
Outline

• Bilinear Inverse Problems
• Key Questions
• Identifiability in Bilinear Inverse Problems
• Two prototypical problems
  ➢ Blind Deconvolution with sparsity models
    □ Identifiability
    □ Data-efficient algorithm
  ➢ Blind Gain and Phase Calibration (BGPC) with sparsity models
    □ Identifiability
    □ Data-efficient algorithm
• Conclusions
1 Bilinear Inverse Problems

2 Identifiability in Bilinear Inverse Problems

3 Blind Deconvolution
   • Identifiability
   • Data-Efficient Algorithm

4 Blind Gain and Phase Calibration (BGPC)
   • Identifiability in BGPC
   • Blind Gain and Phase Calibration via Power Iteration

5 Discussion
Inverse Problems

- Image deblurring
- Tomography (CT)
- MRI
- Sensor Arrays
- ...

Solving equation $Ax = b$, subject to $x \in \Omega$

$\Omega$: Smoothness, Subspace, Sparsity, ...
Inverse Problems

- Image deblurring
- Tomography (CT)
- MRI
- Sensor Arrays
- ...

Solving equation $Ax = b$, subject to $x \in \Omega$

$\Omega$: Smoothness, Subspace, Sparsity, ...
Inverse Problems

- Image deblurring
- Tomography (CT)
- MRI
- Sensor Arrays

\[
Ax = b, \quad \text{subject to } x \in \Omega
\]

\(\Omega\): Smoothness, Subspace, Sparsity, ...

Solving equation \(Ax = b\), subject to \(x \in \Omega\)
Inverse Problems

- Image deblurring
- Tomography (CT)
- MRI
- Sensor Arrays
- ...

Solving equation $Ax = b$, subject to $x \in \Omega$

$\Omega$: Smoothness, Subspace, Sparsity, ...
Bilinear Inverse Problems

What if both the input and the system are unknown?

- Blind Signal Processing $\implies$ Bilinear Inverse Problems
- Blind image deblurring
- Calibrationless pMRI
- Blind gain and phase calibration in antenna arrays
- Dictionary learning
- ...
Bilinear Inverse Problems

What if both the input and the system are unknown?

Blind Signal Processing $\implies$ Bilinear Inverse Problems

- Blind image deblurring
  - Calibrationless pMRI
  - Blind gain and phase calibration in antenna arrays
  - Dictionary learning
  - ...
Bilinear Inverse Problems

What if both the input and the system are unknown?

- Blind Signal Processing $\implies$ Bilinear Inverse Problems

- Blind image deblurring
- Calibrationless pMRI
- Blind gain and phase calibration in antenna arrays
- Dictionary learning
- ...
Bilinear Inverse Problems

What if both the input and the system are unknown?

Blind Signal Processing $\implies$ Bilinear Inverse Problems

- Blind image deblurring
- Calibrationless pMRI
- Blind gain and phase calibration in antenna arrays
  - Dictionary learning
  - ...
Bilinear Inverse Problems

What if both the input and the system are unknown?

- Blind Signal Processing $\implies$ Bilinear Inverse Problems
- Blind image deblurring
- Calibrationless pMRI
- Blind gain and phase calibration in antenna arrays
- Dictionary learning
- ...
Bilinear Inverse Problem (BLIP)

Definition (Bilinear Mapping)

\[ \mathcal{F} : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z} \] such that:

\[ \mathcal{F}(a_1 x_1 + a_2 x_2, y) = a_1 \mathcal{F}(x_1, y) + a_2 \mathcal{F}(x_2, y), \]
\[ \mathcal{F}(x, b_1 y_1 + b_2 y_2) = b_1 \mathcal{F}(x, y_1) + b_2 \mathcal{F}(x, y_2). \]

Definition (Bilinear Inverse Problem)

Given bilinear measurement \( z = \mathcal{F}(x_0, y_0) \), and constraint sets \( \Omega_x, \Omega_y \):

\[ \text{(BLIP)} \quad \text{find} \quad (x, y), \]
\[ \text{s.t.} \quad \mathcal{F}(x, y) = z, \]
\[ x \in \Omega_x, \quad y \in \Omega_y. \]
Example: Blind Deconvolution

Given the measurement \( z = x_0 \otimes y_0 \):

- find \((x, y)\),
- s.t. \( x \otimes y = z \),
- \( x \in \mathbb{C}^n, y \in \mathbb{C}^n \).

- Scaling ambiguity and shift ambiguity

\[
\begin{align*}
2S_1(x_0) & \otimes \frac{1}{2}S_{-1}(y_0) \\
\frac{1}{2}S_2(x_0) & \otimes 2S_{-2}(y_0)
\end{align*}
\]
**Bilinear Inverse Problems (BLIPs)**

**Definition (Bilinear Inverse Problem)**

Given bilinear measurement $z = \mathcal{F}(x_0, y_0)$, and constraint sets $\Omega_X, \Omega_Y$:

\[(\text{BLIP}) \quad \text{find} \quad (x, y), \quad \text{s.t.} \quad \mathcal{F}(x, y) = z, \quad x \in \Omega_X, \ y \in \Omega_Y.\]

**Questions:**

- **Identifiability**
  - In what sense can the solution be unique?
  - Can $x$ be determined uniquely without determining $y$ uniquely?
  - Uniqueness conditions?
  - Stability?

- **algorithms**
  - Data efficiency (sample complexity optimal)?
  - Computational efficiency?
1 Bilinear Inverse Problems

2 Identifiability in Bilinear Inverse Problems

3 Blind Deconvolution
   • Identifiability
   • Data-Efficient Algorithm

4 Blind Gain and Phase Calibration (BGPC)
   • Identifiability in BGPC
   • Blind Gain and Phase Calibration via Power Iteration

5 Discussion
Identifiability up to a Transformation Group

Given the bilinear measurement $z = \mathcal{F}(x_0, y_0)$:

\[(\text{BLIP}) \quad \text{find } (x, y), \quad \text{s.t. } \mathcal{F}(x, y) = z, \quad x \in \Omega_x, \ y \in \Omega_y.\]

Transformation group $\mathcal{T}$ on $\Omega_x \times \Omega_y$ associated with $\mathcal{F}$ satisfies:

- Closed under composition.
- Contains identity transformation $1$.
- Contains inverse transformations: $\mathcal{W}$ and $\mathcal{W}^{-1}$.
- $\mathcal{F}$ is preserved: $\mathcal{F} = \mathcal{F} \circ \mathcal{W}$, $\forall \mathcal{W} \in \mathcal{T}$.

Definition (Identifiability up to Transformation Group $\mathcal{T}$)

Every solution satisfies $(x, y) = \mathcal{W}(x_0, y_0)$, for some $\mathcal{W} \in \mathcal{T}$.

Example: Every solution is a scaled and/or circularly shifted version of $(x_0, y_0)$. 

Example: Blind Deconvolution Problem

Given the measurement $z = x_0 \ast y_0$:

\[
\text{find } (x, y),
\]

\[
\text{s.t. } x \ast y = z,
\]

\[
x \in \mathbb{C}^n, \ y \in \mathbb{C}^n.
\]

- **Transformation group** on $\mathbb{C}^n \times \mathbb{C}^n$ associated with circular convolution:

\[
\mathcal{T} = \left\{ \mathcal{W} : \mathcal{W}(x, y) = \left( \sigma S_\ell(x), \frac{1}{\sigma} S_{-\ell}(y) \right), \text{ for } \sigma \neq 0, \ \ell \in \mathbb{Z} \right\}.
\]

- **Equivalence class**:

\[
[(x_0, y_0)]_{\mathcal{T}} = \left\{ (x, y) \in \mathbb{C}^n \times \mathbb{C}^n : (x, y) = \left( \sigma S_\ell(x_0), \frac{1}{\sigma} S_{-\ell}(y_0) \right), \text{ for } \sigma \neq 0, \ \ell \in \mathbb{Z} \right\}.
\]
Main Result: identifiability up to a transformation group

Given the bilinear measurement $z = \mathcal{F}(x_0, y_0)$:

\[(\text{BLIP}) \quad \text{find } (x, y), \quad \text{s.t. } \mathcal{F}(x, y) = z, \quad x \in \Omega_X, \; y \in \Omega_Y.\]

Theorem (Y. Li, K. Lee, and Bresler, 2015)

A sufficient condition for identifiability is:

1. Vector $x_0$ can be identified up to the transformation group.
2. Once $x_0$ is identified, the recovery of $y_0$ is unique.

Under mild conditions on $\mathcal{F}$ and $T$, the above conditions are also necessary.

Example - Blind Deconvolution:

1. Every solution $x$ is a scaled and/or circularly shifted version of $x_0$.
2. Given $x_0$, the solution $y_0$ to non-blind deconvolution is unique.
The rest of this talk ...

Two problems:
- Blind deconvolution (BD)
- Blind gain and phase calibration (BGPC)

Main assumptions:
- Subspace or sparsity structures
- Generic or random bases/frames (benchmark setup for theoretical analysis)

Two questions:
- Optimal identifiability conditions
- Data-Efficient solutions
1 Bilinear Inverse Problems

2 Identifiability in Bilinear Inverse Problems

3 Blind Deconvolution
   - Identifiability
   - Data-Efficient Algorithm

4 Blind Gain and Phase Calibration (BGPC)
   - Identifiability in BGPC
   - Blind Gain and Phase Calibration via Power Iteration

5 Discussion
Blind Deconvolution

Both \( u \) and \( v \) are unknown \( \rightarrow \) \textit{ill-posed} bilinear inverse problem

- Solved with “good” priors (e.g., subspace, sparsity)

✓ Empirical success in various applications (e.g., blind image deblurring, speech dereverberation, seismic data analysis, etc.)
  - Theoretical results are limited. \( \rightarrow \) The focus of our work
Problem Statement

- **Signal**: $u_0 \in \mathbb{C}^n$
- **Filter**: $v_0 \in \mathbb{C}^n$
- **Measurement**: $z = u_0 \odot v_0 \in \mathbb{C}^n$

Find $(u, v)$

s.t. $u \odot v = z$,

$u \in \Omega_u$, $v \in \Omega_v$.

Three scenarios:

1. **Subspace constraints**
2. **Sparsity constraints**
3. **Mixed constraints**
Problem Statement

- Signal: \( u_0 \in \mathbb{C}^n \)
- Filter: \( v_0 \in \mathbb{C}^n \)
- Measurement: \( z = u_0 \ast v_0 \in \mathbb{C}^n \)

Find \((u, v)\)

\[
\text{s.t.} \quad u \ast v = z, \\
\quad u \in \Omega_U, \ v \in \Omega_V.
\]

Three scenarios:
1. Subspace constraints
2. Sparsity constraints
3. Mixed constraints
Problem Statement

- **Signal:** \( u_0 \in \mathbb{C}^n \)
- **Filter:** \( v_0 \in \mathbb{C}^n \)
- **Measurement:** \( z = u_0 \ast v_0 \in \mathbb{C}^n \)

Find \((u, v)\)

s.t. \( u \ast v = z, \)

\( u \in \Omega_U, \ v \in \Omega_V. \)

Three scenarios:

1. Subspace constraints
2. Sparsity constraints
3. Mixed constraints
Problem Statement

- **Signal:** $u_0 = Dx_0$, the columns of $D \in \mathbb{C}^{n \times m_1}$ form a basis or a frame
- **Filter:** $v_0 = Ey_0$, the columns of $E \in \mathbb{C}^{n \times m_2}$ form a basis or a frame
- **Measurement:** $z = u_0 \circledast v_0 = (Dx_0) \circledast (Ey_0) \in \mathbb{C}^n$

(BD) Find $(x, y)$

s.t. $(Dx) \circledast (Ey) = z$

$x \in \Omega_X, y \in \Omega_Y$.

Three scenarios:

1. **Subspace constraints:**
   
   $\Omega_X = \mathbb{C}^{m_1}$
   
   and $\Omega_Y = \mathbb{C}^{m_2}$

2. **Sparsity constraints:**

   $\Omega_X = \{x \in \mathbb{C}^{m_1} : \|x\|_0 \leq s_1\}$
   
   and $\Omega_Y = \{y \in \mathbb{C}^{m_2} : \|y\|_0 \leq s_2\}$

3. **Mixed constraints:**

   $\Omega_X = \{x \in \mathbb{C}^{m_1} : \|x\|_0 \leq s_1\}$
   
   and $\Omega_Y = \mathbb{C}^{m_2}$
Weak and Strong Identifiability

**Definition (Identifiability up to scaling)**

- **Weak Identifiability** ($(x_0, y_0)$ is identifiable): every solution $(x, y)$ satisfies $x = \sigma x_0$ and $y = \frac{1}{\sigma} y_0$ for some nonzero scalar $\sigma$.
- **Strong Identifiability** ($\Omega_X \times \Omega_Y$ is identifiable): every $(x_0, y_0) \in \Omega_X \times \Omega_Y$ is identifiable up to scaling.

**Lifting**

Define $G_{DE} : \mathbb{C}^{m_1 \times m_2} \rightarrow \mathbb{C}^n$ such that $G_{DE}(xy^\top) = (Dx) \otimes (Ey)$, and $M_0 = x_0y_0^\top \in \Omega_M = \{xy^\top : x \in \Omega_X, y \in \Omega_Y\}$.

(BD) Find $(x, y)$, s.t. $(Dx) \otimes (Ey) = z$, $x \in \Omega_X$, $y \in \Omega_Y$.

$\implies$ (Lifted BD) Find $M$, s.t. $G_{DE}(M) = z$, $M \in \Omega_M$.

- Weak Identifiability $\iff$ Unique recovery of a **single** arbitrary matrix $M_0 \in \Omega_M$
- Strong Identifiability $\iff$ Uniform unique recovery of all matrices in $\Omega_M$
Single-point and Uniform Stability

\begin{align*}
\text{(Noisy BD)} & \quad \min_M \|G_{DE}(M) - z\|_2, \\
\text{s.t.} & \quad M \in \Omega_B := \Omega_M \cap B_{C^{m_1 \times m_2}}.
\end{align*}

Definition (Stability)

Single-point stability at $M_0$: $\|G_{DE}(M) - G_{DE}(M_0)\|_2 \leq \delta$ for $M \in \Omega_B \implies \|M - M_0\|_2 \leq \varepsilon.$

Uniform Stability on $\Omega_B$: $\|G_{DE}(M_1) - G_{DE}(M_2)\|_2 \leq \delta$ for $M_1, M_2 \in \Omega_B \implies \|M_1 - M_2\|_2 \leq \varepsilon.$

- Strong identifiability + single-point stability at $M_0$ \implies $G_{DE}^{-1}$ is continuous at $G_{DE}(M_0)$

- Uniform stability on $\Omega_B$ \implies $G_{DE}^{-1}$ is uniformly continuous on $\Omega_B$

- Stability \implies solution to (Noisy BD) is accurate
Prior Work

- **Identifiability analysis**
  - [Choudhary and Mitra, 2014]: canonical sparsity constraints
    - Lacks sample-complexity type interpretation
  - [Y. Li, K. Lee, and Bresler, 2016b]: generic subspace or sparsity constraints
    - Suboptimal sample complexities

- **Guaranteed recovery algorithms**
  - [Ahmed et al., 2014]: nuclear norm minimization
  - [Ling and Strohmer, 2015]: $\ell_1$ norm minimization
  - [Chi, 2016]: atomic norm minimization
  - [Li et al., 2016a]: gradient descent
  - [K. Lee, Y. Li, Junge, and Bresler, 2017]: alternating minimization

✓ Constructive proof of uniqueness and stability
  - Requires probabilistic assumptions and interpretations
  - Sample complexities contain large constants and log factors

**Goal**

- Identifiability in BD with generic bases or frames
- Optimal sample complexities
Main Results: Identifiability in Blind Deconvolution

\[ (BD) \text{ Find } (x, y) \text{ s.t. } (Dx) \otimes (Ey) = z, \]
\[ x \in \Omega_X, \ y \in \Omega_Y. \]

Sample complexity constant

\[
d = \begin{cases} 
m_1 + m_2 & \text{subspace constraints} \\
 s_1 + m_2 & \text{mixed constraints} \\
 s_1 + s_2 & \text{sparsity constraints} 
\end{cases}
\]

**Theorem (Identifiability, Y. Li, K. Lee, and Bresler, 2015, 2017a)**

- \( n > d \implies \text{Single-Point Identifiability for almost all } D \in \mathbb{C}^{n \times m_1} \text{ and } E \in \mathbb{C}^{n \times m_2}. \)
- \( n > 2d \implies \text{Uniform (strong) Identifiability for almost all } D \in \mathbb{C}^{n \times m_1} \text{ and } E \in \mathbb{C}^{n \times m_2}. \)

The same sample complexities hold for the case where \( x, y, D, E \) are real.
Main Results: Stability in Blind Deconvolution

(BD) Find \((x, y)\) s.t. \((Dx) \otimes (Ey) = z,\)
\[ x \in \Omega_x, \ y \in \Omega_y. \]

Stable recovery under the same sample complexity:

**Theorem (Stability, Y. Li, K. Lee, and Bresler, 2015, 2017a)**

\[ D \in \mathbb{C}^{n \times m_1} \text{ and } E \in \mathbb{C}^{n \times m_2} \ \text{independent random, s.t.:} \]
\[ \{(FD)(j,:)\}^n_{j=1} \ \text{i.i.d. uniform distribution on a ball in } \mathbb{C}^{m_1} \]
\[ \{(FE)(j,:)\}^n_{j=1} \ \text{i.i.d. uniform distribution on a ball in } \mathbb{C}^{m_2} \]

- If \(n > d\), single-point stability w.h.p.
- If \(n > 2d\), uniform stability w.h.p.

Similar stability results hold for the case where \(x, y, D, E\) are real
Summary: Identifiability and Stability in Blind Deconvolution

- The sample complexities are optimal (up to a small additive number):
  - Sufficient condition: $n > s_1 + s_2$
  - Necessary condition: $n \geq s_1 + s_2 - 1$
    (information-theoretic lower bound [Lee et al., 2013])

- Identifiability for almost all $D, E$
  $\implies$ Unique recovery for random $D, E$ w.p. 1

- Stability for random $D, E$ w.h.p.

- More recently slightly improved identifiability and stability results for blind deconvolution using techniques from algebraic geometry [Kech and Krahmer, 2016, 2017].
  - Necessary & sufficient conditions differ from [Y. Li, K. Lee, Bresler, 2015, 2017] by an additive term of at most 5 samples.
Outline of Theoretical Analysis of Identifiability in BD
Lifting: BD as a Matrix Recovery Problem

Let $M_0 = x_0 y_0^\top$, $a_j = (FD)^{(j,:)}^\ast$, $b_j = (FE)^{(j,:)}^\ast$, and $\tilde{e} = \frac{1}{\sqrt{n}} Fe$. Then the frequency domain measurements are

$$
\frac{1}{\sqrt{n}} F G_{DE}(M) = [a_1^* M b_1, a_2^* M b_2, \ldots, a_n^* M b_n]^\top
$$

$\implies$ study unique / stable recovery of structured matrices!

- For almost all $a_j \in \mathbb{C}^{m_1}$ and $b_j \in \mathbb{C}^{m_2}$,
  - $n > d \implies$ Single-Point (Weak) Identifiability
  - $n > 2d \implies$ Uniform (Strong) Identifiability

- For $\{a_j\}_{j=1}^n \overset{i.i.d.}{\sim} \text{Uniform}(RB_{\mathbb{C}^{m_1}})$, and $\{b_j\}_{j=1}^n \overset{i.i.d.}{\sim} \text{Uniform}(RB_{\mathbb{C}^{m_2}})$
  - $n > d \implies$ Single-Point Stability w.h.p.
  - $n > 2d \implies$ Uniform Stability w.h.p.
Two main ingredients of analysis:

- The constraint set $\Omega_B$ is “small”:
  It has a small covering number

- The measurement vectors $\{a_j, b_j\}_{j=1}^n$ are “generic”:
  Their probability distribution satisfies certain concentration of measure bounds

Refinement and extension of similar results on unique recovery of real-valued matrices [Riegler et al., 2015]
1 Bilinear Inverse Problems

2 Identifiability in Bilinear Inverse Problems

3 Blind Deconvolution
   - Identifiability
   - Data-Efficient Algorithm

4 Blind Gain and Phase Calibration (BGPC)
   - Identifiability in BGPC
   - Blind Gain and Phase Calibration via Power Iteration

5 Discussion
Subsampled Blind Deconvolution

Subsampled measurements – $m$ samples at $\Omega$:

$$b = \sqrt{\frac{n}{m}} S_{\Omega}(h \ast x) + z$$

- $x \in \mathbb{C}^n$ is $s_1$-sparse over a dictionary $\Phi$: $x = \Phi u$, $\|u\|_0 \leq s_1$
- $h \in \mathbb{C}^n$ is $s_2$-sparse over a dictionary $\Psi$: $h = \Psi v$, $\|v\|_0 \leq s_2$
- Relevant applications: super-resolution, blind calibration in pMRI
- More challenging than fully sampled blind deconvolution!
Alternating Minimization

- Natural heuristic for a bilinear inverse problem
- Finding a good initialization is critical.
- Solution to subproblems with performance guarantees

**Why alternating minimization for blind deconvolution?**
An alternating minimization algorithm, sparse power factorization (SPF) [K. Lee, Y. Wu, and Bresler, 2013] empirically outperformed convex methods (mixed norm, nuclear norm, and their best combination).

SPF also provides a near optimal performance guarantee while convex methods were proven suboptimal [Oymak et al., 2015].
Main Results - Sparse Blind Deconvolution [Lee, Y. Li, Junge, and Bresler, 2015, 2017]

- **Sparsity models** (union of subspaces)
  - $x \in \mathbb{C}^n$ is $s_1$-sparse over a dictionary $\Phi$
  - $h \in \mathbb{C}^n$ is $s_2$-sparse over a dictionary $\Psi$

- $h$ and $x$ are spectrally $\mu$-flat:

\[
\frac{\|Fh\|_\infty}{\|Fh\|_2 / \sqrt{n}} \leq \sqrt{\mu}, \quad F: \text{unitary DFT matrix}
\]

Suppose $\Phi, \Psi$ are i.i.d. Gaussian matrices and $\mu = O(\log n)$. Then stable blind deconvolution is achieved by alternating minimization at near optimal sample complexity of $m = O((s_1 + s_2) \log^6 n)$.*

* With additional technical conditions.
Blind Deconvolution - Comparison with Previous Results

- Number of degrees of freedom in the problem is proportional to $s$
- The dimension of the ambient space is $N$
- We compare
  - ARR [Ahmed et al., 2014]
  - LS [Ling and Strohmer, 2015]
  - C [Chi, 2016]
  - LLSW [Li et al., 2016a]
  - LLJB [Lee, Y. Li, Junge, and Bresler, 2017]

<table>
<thead>
<tr>
<th></th>
<th>Signal Model</th>
<th>#Measurements</th>
<th>Subsampling</th>
<th>Additional Assumptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARR</td>
<td>subspace</td>
<td>$O(s \log N)$</td>
<td>No</td>
<td>spectral flatness</td>
</tr>
<tr>
<td>LS</td>
<td>union of subspaces</td>
<td>$O(s^2 \log N)$</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>mixed (off-grid)</td>
<td>$O(s^3 \log N)$</td>
<td>No</td>
<td>separation of frequencies</td>
</tr>
<tr>
<td>LLSW</td>
<td>subspace</td>
<td>$O(s \log^2 N)$</td>
<td>No</td>
<td>spectral flatness</td>
</tr>
<tr>
<td>LLJB</td>
<td>union of subspaces</td>
<td>$O(s \log^6 N)$</td>
<td>✓</td>
<td>spectral flatness</td>
</tr>
</tbody>
</table>
Application: Astronomical Image Deblurring

Original Image

Blurry Image

Initial Estimates

Final Reconstruction
Summary: Data-Efficient Algorithm for Blind Deconvolution

- Blind deconvolution with a sparsity model
  
  \[\text{subspace} \rightarrow \text{union of subspaces}\]
  
  \[\text{sparsity with known support} \rightarrow \text{sparsity with unknown support}\]

- Theory of blind deconvolution extended to subsampled blind deconvolution.

- Blind deconvolution of sparse signals by an alternating minimization with exact projection is guaranteed, under the spectral flatness condition and some additional technical conditions, for near optimal sparsity levels.
1 Bilinear Inverse Problems

2 Identifiability in Bilinear Inverse Problems

3 Blind Deconvolution
  - Identifiability
  - Data-Efficient Algorithm

4 Blind Gain and Phase Calibration (BGPC)
  - Identifiability in BGPC
  - Blind Gain and Phase Calibration via Power Iteration

5 Discussion
Blind Gain and Phase Calibration (BGPC)

Find \((\lambda, \Phi)\),
s.t. \(\text{diag}(\lambda)\Phi = Y\),
\[\lambda \in \mathbb{C}^n, \quad \Phi \in \mathbb{C}^{n \times N}.\]

- \(\Phi_0\): unknown signals; \(\lambda_0\): unknown gain and phase, 
  \[\lambda_0^{(j)} = g^{(j)} e^{i p^{(j)}}\]

- Ill-posed! \(\implies\) Need extra constraints for unique solution

\[\Omega_{\Phi} = \{\Phi = AX : X \in \Omega_X\}\]:
- Subspace: \(A\) is a known basis for a lower-dimensional subspace.
- Joint sparsity: \(A\) is a known basis or frame, columns of \(X\) are jointly \(s\)-sparse.

\[\text{(BGPC)} \quad \text{find} \quad (\lambda, X), \quad \text{s.t.} \quad \text{diag}(\lambda)AX = Y, \quad \lambda \in \mathbb{C}^n, \quad X \in \Omega_X.\]
Applications in Which BGPC Arises

- Blind gain and phase calibration in sensor array processing [Paulraj and Kailath, 1985]
- Multichannel blind deconvolution [Abed-Meraim et al., 1997]
- Synthetic aperture radar (SAR) autofocus [Morrison et al., 2009]
- Inverse rendering in computational relighting [Nguyen et al., 2013]

\[
\text{(BGPC)} \quad \text{find } (\lambda, X), \\
\text{s.t.} \quad \text{diag}(\lambda)AX = Y, \\
\lambda \in \mathbb{C}^n, \ X \in \Omega_x.
\]
Applications in Which BGPC Arises

- Blind gain and phase calibration in sensor array processing [Paulraj and Kailath, 1985]
- Multichannel blind deconvolution [Abed-Meraim et al., 1997]
- Synthetic aperture radar (SAR) autofocus [Morrison et al., 2009]
- Inverse rendering in computational relighting [Nguyen et al., 2013]

\[
\text{(BGPC) } \text{find } (\lambda, X), \quad \text{s.t. } \text{diag}(\lambda)AX = Y, \\
\lambda \in \mathbb{C}^n, \quad X \in \Omega_X.
\]

\[
\begin{align*}
  s[n] & \rightarrow h_1[n] \rightarrow y_1[n] = s[n] * h_1[n] \\
  & \quad \rightarrow h_2[n] \rightarrow y_2[n] = s[n] * h_2[n] \\
  & \quad \quad \vdots \quad \vdots \\
  & \quad \quad \rightarrow h_N[n] \rightarrow y_N[n] = s[n] * h_N[n]
\end{align*}
\]
Applications in Which BGPC Arises

- Blind gain and phase calibration in sensor array processing [Paulraj and Kailath, 1985]
- Multichannel blind deconvolution [Abed-Meraim et al., 1997]
- Synthetic aperture radar (SAR) autofocus [Morrison et al., 2009]
- Inverse rendering in computational relighting [Nguyen et al., 2013]

\[
\text{(BGPC)} \quad \text{find} \quad (\lambda, X), \\
\text{s.t.} \quad \text{diag}(\lambda)AX = Y, \\
\lambda \in \mathbb{C}^n, \quad X \in \Omega_X.
\]
Applications in Which BGPC Arises

- Blind gain and phase calibration in sensor array processing [Paulraj and Kailath, 1985]
- Multichannel blind deconvolution [Abed-Meraim et al., 1997]
- Synthetic aperture radar (SAR) autofocus [Morrison et al., 2009]
- Inverse rendering in computational relighting [Nguyen et al., 2013]
- ...

(BGPC) find \((\lambda, X)\),

s.t. \(
\text{diag}(\lambda)AX = Y, \\
\lambda \in \mathbb{C}^n, \ X \in \Omega_X. 
\)
How much data? – Sample Complexity

Find \((\lambda, X)\),

\[
\text{s.t. } \text{diag}(\lambda)AX = Y, \\
\lambda \in \mathbb{C}^n, \ X \in \Omega_X.
\]

Sample Complexity: condition on \(N\) (number of channels, snapshots, lighting conditions, etc.) in terms of \(n\) (length of the signals), \(m\) (dimension of subspace), or \(s\) (sparsity level)

<table>
<thead>
<tr>
<th>Subspace Constraint</th>
<th>Joint Sparsity Constraint</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sufficient Condition</strong></td>
<td>(N \geq m^2)</td>
</tr>
<tr>
<td><strong>Necessary Condition</strong></td>
<td>(N \geq \frac{n-1}{n-m})</td>
</tr>
<tr>
<td><strong>Gap</strong></td>
<td>?</td>
</tr>
</tbody>
</table>

\(^2\)Balzano and Nowak, 2007\] sensor arrays, \([\text{Nguyen et al., 2013}]\) Inverse rendering

\(^3\)Balzano and Nowak, 2007, [Morrison et al., 2009] SAR autofocus
How much data? – Sample Complexity

Find \((\lambda, X)\),
\[
\text{s.t. } \text{diag}(\lambda)AX = Y,
\]
\(\lambda \in \mathbb{C}^n, \ X \in \Omega_X\).

Sample Complexity: condition on \(N\) (number of channels, snapshots, lighting conditions, etc.) in terms of \(n\) (length of the signals), \(m\) (dimension of subspace), or \(s\) (sparsity level)

<table>
<thead>
<tr>
<th>Subspace Constraint</th>
<th>Joint Sparsity Constraint</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sufficient Condition</strong></td>
<td>(N \geq m^2)</td>
</tr>
<tr>
<td><strong>Necessary Condition</strong></td>
<td>(N \geq \frac{n-1}{n-m})</td>
</tr>
<tr>
<td><strong>Gap</strong></td>
<td>?</td>
</tr>
</tbody>
</table>

\(^2[\text{Balzano and Nowak, 2007}]\) sensor arrays, \(^3[\text{Nguyen et al., 2013}]\) Inverse rendering

\(^2[\text{Balzano and Nowak, 2007}], \ [\text{Morrison et al., 2009}]\) SAR autofocus
Sample Complexity (with explicit model conditions)

Sample complexity: condition on $N$ in terms of $n, m, or s$

\[ Y = \text{diag}(\lambda) \]

Subspace Constraint

\[ \text{Sufficient Condition} \]
\[ N \geq m \]

\[ \text{Necessary Condition} \]
\[ N \geq \frac{n-1}{n-m} \]

Joint Sparsity Constraint (DFT)

\[ \text{Sufficient Condition} \]
\[ N \geq s \]

\[ \text{Necessary Condition} \]
\[ N \geq \frac{n-1}{n-s} \]

Gap?

[Y. Li, K. Lee, and Bresler, 2015, 2017]
Sample Complexity (with explicit model conditions)

Sample complexity: condition on $N$ in terms of $n$, $m$, or $s$

\[ Y = \text{diag}(\lambda) A X \]

Subspace Constraint

**Sufficient Condition**

\[ N \geq m \]

**Necessary Condition**

\[ N \geq \frac{n-1}{n-m} \]

Joint Sparsity Constraint (DFT)

\[ N \geq s \]

\[ N \geq \frac{n-1}{n-s} \]

Numerical Experiment

[Y. Li, K. Lee, and Bresler, 2015, 2017]
Sample Complexity (with explicit model conditions)

Sample complexity: condition on $N$ in terms of $n$, $m$, or $s$

$Y = \text{diag}(\lambda) A X$

Subspace Constraint

Sufficient Condition

$N \geq m$

Necessary Condition

$N \geq \frac{n-1}{n-m}$

Joint Sparsity Constraint (DFT)

$N \geq s$

$N \geq \frac{n-1}{n-s}$

Numerical Experiment

Necessary conditions almost sufficient!
BGPC with a Subspace Constraint - Tight Conditions

Find \((\lambda, X)\),

\[
\text{s.t. } \text{diag}(\lambda)AX = Y,
\]

\[
\lambda \in \mathbb{C}^n, \ X \in \mathbb{C}^{m \times N}.
\]

\[
Y = \text{diag}(\lambda)A_{n \times m}X_{m \times N}.
\]

Theorem (Y. Li, K. Lee, and Bresler, 2016c)

If \(n > m\) and \(N \geq \frac{n - 1}{n - m}\), then for almost all \(\lambda_0 \in \mathbb{C}^n\), almost all \(X_0 \in \mathbb{C}^{m \times N}\), and almost all \(A \in \mathbb{C}^{n \times m}\), the pair \((\lambda_0, X_0)\) is identifiable up to an unknown scaling.

- \(N \geq \frac{n - 1}{n - m}\), is much less demanding than the previous condition \(N \geq m\).
- Sample complexity is optimal – matches the necessary condition.
- If \(m \leq \frac{n}{2}\), i.e., the dimension of the subspace is less than half the ambient dimension, then \(N = 2\) signals are sufficient to recover \((\lambda_0, X_0)\) uniquely.
- Example: inverse rendering - \(m = 9 \ll n = 256 \times 256 = 2^{16}\) \(\Rightarrow N = 2\) images under different lighting conditions suffice.
BGPC with a Joint Sparsity Constraint - Tight Conditions

Find \((\lambda, X)\),

\[
\text{s.t. } \text{diag}(\lambda)AX = Y,
\]

\[
\lambda \in \mathbb{C}^n,
\]

\[
X \in \{X \in \mathbb{C}^{m \times N} : \text{ } X \text{ has at most } s \text{ nonzero rows}\}.
\]

Theorem (Y. Li, K. Lee, and Bresler, 2016c)

if \(n > 2s\) and \(N \geq \frac{n-1}{n-2s}\), then for almost all \(\lambda_0 \in \mathbb{C}^n\), almost all \(X_0 \in \mathbb{C}^{m \times N}\) with \(s\) nonzero rows, and almost all \(A \in \mathbb{C}^{n \times m}\), the pair \((\lambda_0, X_0)\) is identifiable up to an unknown scaling.

- \(N \geq \frac{n-1}{n-2s}\), is far superior to the previous \(N \geq s\).
- If \(s \leq \frac{n}{4}\), then \(N = 2\) signals are sufficient to recover \((\lambda_0, X_0)\) uniquely.
- Example: sensor array processing, with \(s\) sources fewer than a quarter of the number \(n\) of sensors, \(N = 2\) snapshots suffice.
Summary - Identifiability of BGPC

- **Tight** conditions for identifiability in BGPC under subspace or joint sparsity constraints

- Result true for “almost all” $\lambda_0$, $X_0$, and $A \Rightarrow$ it fails only on a degenerate set.

- For a given $\lambda_0$, $X_0$, and $A$, can test to determine whether the solution $(\lambda_0, X_0)$ is unique up to scaling.

- The degenerate set of $(\lambda_0, X_0, A)$ that fails the test, is an algebraic variety, which is not dense in the ambient space.
1 Bilinear Inverse Problems

2 Identifiability in Bilinear Inverse Problems

3 Blind Deconvolution
   • Identifiability
   • Data-Efficient Algorithm

4 Blind Gain and Phase Calibration (BGPC)
   • Identifiability in BGPC
   • Blind Gain and Phase Calibration via Power Iteration

5 Discussion
Blind Gain and Phase Calibration

\[ Y = \text{diag}(\lambda)AX + W \]

- \( \lambda \in \mathbb{C}^n \)
- \( A \in \mathbb{C}^{n \times m} \) known
- \( X \in \mathbb{C}^{m \times N} \)
- \( Y, W \in \mathbb{C}^{n \times N} \)

**Subspace Case:** \( n > m \), \( X \) is unconstrained

**Sparsity Case:** columns of \( X \) are \( s_0 \)-sparse
  - Joint Sparsity Case
BGPC as a Linear Inverse Problem

A popular trick:

\[ Y = \text{diag}(\lambda)AX \Rightarrow AX - \text{diag}(\gamma)Y = 0 \]

Prior work

- Application-specific
  - [Balzano and Nowak, 2007]: sensor network calibration
  - [Morrison et al., 2009]: SAR autofocus
  - [Nguyen et al., 2013]: inverse rendering
  - ...

  - Subspace case. No error bounds

- [Ling and Strohmer, 2016]: least squares

\[
\min_{X, \gamma} \|AX - \text{diag}(\gamma)Y\|_F^2 + \|e^*\gamma - 1\|_2^2
\]

  - Subspace case. Error sensitive to \(e\)

- [Wang and Chi, 2016]: \(\ell_1\) minimization, when \(A\) is the DFT matrix

\[
\min_{\gamma} \|A^{-1}\text{diag}(\gamma)Y\|_1 \quad \text{s.t. } 1^T\gamma = 1
\]

  - Sparsity case. Only works when \(\gamma_k \approx 1\) (gains \(\approx 1\), phases \(\approx 0\))
BGPC as an Eigenvector Problem

\[ Y = \text{diag}(\lambda)AX \implies AX - \text{diag}(\gamma)Y = 0 \]
\[ \implies Dx - E\gamma = 0 \]
\[ \implies [D, \alpha E][x^\top, -\gamma^\top/\alpha]^\top = 0 \]
\[ \implies B\eta = 0 \]
\[ \implies \eta \text{ is the dominant eigenvector of } G = \beta I - B \]

\[ \gamma_k = \frac{1}{\lambda_k} \]
BGPC as an Eigenvector Problem

\[ Y = \text{diag}(\lambda)AX \implies AX - \text{diag}(\gamma)Y = 0 \]

\[ \implies Dx - E\gamma = 0 \]

\[ \implies [D, \alpha E][x^\top, -\gamma^\top / \alpha]^\top = 0 \]

\[ \implies B\eta = 0 \]

\[ \implies \eta \text{ is the dominant eigenvector of } G = \beta I - B \]

\[
D = \begin{bmatrix}
I_N \otimes a_1^* \\
\vdots \\
I_N \otimes a_n^*
\end{bmatrix}, \quad E = \begin{bmatrix}
y_1 \\
\vdots \\
y_n
\end{bmatrix}, \quad x = \text{vec}(X)
\]
BGPC as an Eigenvector Problem

\[ Y = \text{diag}(\lambda)AX \implies AX - \text{diag}(\gamma)Y = 0 \]

\[ \implies Dx - E\gamma = 0 \]

\[ \implies [D, \alpha E][x^\top, -\gamma^\top/\alpha]^\top = 0 \]

\[ \implies B\eta = 0 \]

\[ \implies \eta \text{ is the dominant eigenvector of } G = \beta I - B \]

Parameter \( \alpha \): balance \( x \) and \( \gamma \)
BGPC as an Eigenvector Problem

\[ Y = \text{diag}(\lambda) AX \implies AX - \text{diag}(\gamma) Y = 0 \]
\[ \implies Dx - E\gamma = 0 \]
\[ \implies [D, \alpha E][x^\top, -\gamma^\top / \alpha]^\top = 0 \]
\[ \implies B\eta = 0 \]
\[ \implies \eta \text{ is the dominant eigenvector of } G = \beta I - B \]

\[ B = \begin{bmatrix}
D^* D & \alpha D^* E \\
\alpha E^* D & \alpha^2 E^* E
\end{bmatrix}, \quad \eta = [x^\top, -\gamma^\top / \alpha]^\top \]
BGPC as an Eigenvector Problem

\[ Y = \text{diag}(\lambda)AXAX - \text{diag}(\gamma)Y = 0 \]
\[ Dx - E\gamma = 0 \]
\[ [D, \alpha E][x^\top, -\gamma^\top/\alpha]^\top = 0 \]
\[ B\eta = 0 \]
\[ \eta \text{ is the dominant eigenvector of } G = \beta I - B \]

Parameter \( \beta \): sufficiently large

- Big data problems?
Power Iteration

\[
\max_{\eta} \quad \eta^* G\eta \\
\text{s.t.} \quad \eta = [x^\top, -\gamma^\top / \alpha]^\top \in \text{constraint set}
\]

Subspace Case – Power Iteration

\[ \eta_t = G\eta_{t-1} / \|G\eta_{t-1}\|_2 \]

Sparsity Case – Truncated Power Iteration [Yuan and Zhang, 2013]

\[ \tilde{\eta}_t = G\eta_{t-1} / \|G\eta_{t-1}\|_2 \]
\[ \eta_t = \Pi_{s_1}(\tilde{\eta}_t) / \|\Pi_{s_1}(\tilde{\eta}_t)\|_2 \]

- \(\Pi_{s_1}\) – projection s.t. columns of \(X\) are \(s_1\)-sparse

- \(s_1\) – threshold usually chosen larger than the true sparsity level, \(s_1 > s_0\).
Main Assumptions

To derive an error bound...

- \( A \): i.i.d. \( \mathcal{CN}(0, \frac{1}{n}) \)
- \( \lambda \): flat, \( 1 - \delta \leq |\lambda_k|^2 \leq 1 + \delta \)
- \( X \): normalized and well-conditioned
- \( W \): moderate noise

Parameters

- \( \alpha = \sqrt{n} \) (for which the norms of \( x \) and \( \gamma/\alpha \) are roughly the same)
- \( \beta = 3/2 \) (based on the eigenvalues. In practice, we use \( \beta = \|G\| \))
- \( s_1 = cs_0 \) (In practice, the choice of \( s_1/s_0 \) depends on application)
Power Iteration - Subspace Case

- $\hat{\eta}$: true $\eta$
- $\tilde{\eta}$: the largest eigenvector of noisy $G$
- $\eta_t$: iterates $\eta_0$: initial estimate
- $d(\eta, \eta') := \min_{\varphi} \left\| e^{\sqrt{-1} \varphi} \eta - \eta' \right\|_2 = \sqrt{2 - 2|\langle \eta, \eta' \rangle|}$

Theorem (Subspace Case, Y. Li, Lee, and Bresler, 2017b)

- $\lambda$-Flatness: $\delta < 1/4 \quad \leftarrow \quad 1 - \delta \leq |\lambda_k|^2 \leq 1 + \delta$
- Initialization: $\xi := |\langle \eta_0, \tilde{\eta} \rangle | > 0$ (trivial)
- Sample complexity condition:

$$\max \left\{ \frac{m \log^2(Nm + n)}{n}, \frac{\log(Nm + n)}{N}, \frac{\log(Nm + n)}{m} \right\} \leq C$$

Then w.h.p.

$$d(\eta_t, \hat{\eta}) \leq \rho^t d(\eta_0, \tilde{\eta}) + 2\Delta$$

where $\rho < 1$ depends on $\delta$ and $\xi$, and $\Delta$ is proportional to the noise level.

- Sample complexity condition: $n \gtrsim m \gtrsim 1, \quad N \gtrsim 1$
- Recall sufficient condition for uniqueness: $n > m, \quad N \geq \frac{n-1}{n-m}$
Truncated Power Iteration - Sparsity Case

- \( \hat{\eta} \): true \( \eta \)
- \( \hat{\eta} \): the largest eigenvector of noisy \( G \)
- \( \eta_t \): iterates \( \eta_0 \): initial estimate
- \( d(\eta, \eta') := \min_\phi \left\| e^{\sqrt{-1} \phi} \eta - \eta' \right\|_2 = \sqrt{2 - 2|\langle \eta, \eta' \rangle|} \)

Theorem (Joint Sparsity Case, Y. Li, Lee, and Bresler, 2017b)

- \( \lambda \)-Flatness: \( \delta < 1/4 \leftarrow 1 - \delta \leq |\lambda_k|^2 \leq 1 + \delta \)
- Initialization: \( |\langle \eta_0, \hat{\eta} \rangle| > \xi + \tilde{\Delta} \) for some \( \xi \in (0, 1) \)
- Sample complexity condition (assuming \( ss_0 + 2s_1 \)):
  \[
  \max \left\{ \frac{(s + N) \log^3 n \log(sN + m)}{n}, \frac{\sqrt{s} \log^3 n \log(sN + m)}{N}, \frac{\log n}{s_0} \right\} \leq C
  \]
  Then w.h.p.
  \[
  d(\eta_t, \hat{\eta}) \leq (\tilde{\rho})^t d(\eta_0, \hat{\eta}) + 2\sqrt{5} \tilde{\Delta} / (1 - \tilde{\rho})
  \]
  where \( \tilde{\rho} \) depends on \( \delta, s_1/s_0, \) and \( \xi \) (a good initialization is necessary for \( \tilde{\rho} < 1 \)). \( \tilde{\Delta} \) is proportional to the noise level.

- Sample complexity condition: \( n \gtrsim s \gtrsim 1, \quad N \gtrsim \sqrt{s} \)
- Recall sufficient condition for uniqueness: \( n > 2s, \quad N \geq \frac{n-1}{n-2s} \)
**Intuition**

- Key to the success of power iteration – a gap between the largest and the second largest (in terms of absolute value) eigenvalues of $G = \beta I - B$

  $$B = \begin{bmatrix} D^* D & \alpha D^* E \\ \alpha E^* D & \alpha^2 E^* E \end{bmatrix}$$

- Under our assumptions on $A, \lambda, X, W$, and $\alpha = \sqrt{n}$
  - Smallest eigenvalue of $B \approx 0$
  - Largest eigenvalue of $B \approx 2$
  - Other eigenvalues of $B \approx 1$

  Therefore, we choose $\beta = 3/2$

- Bound the deviation of $B$ from its expectation

- In the joint sparsity case, apply the same argument (and a union bound) to submatrices of $B$
Numerical Experiments - Subspace Case

- $n = 128$, $N = 16$, $m$ varying
- power iteration vs. least squares

Noiseless

20dB

14dB

6dB
Numerical Experiments - Sparsity Case

- $n = 128$, $N = 16$, $m = 256$, $s_0$ varying
- Truncated power iteration vs. $\ell_1$ minimization
- Access to an initial estimate $\lambda_0$
  - Truncated power iteration: initialize with $\gamma_{0,k} = 1/\lambda_{0,k}$
  - $\ell_1$ minimization: let $\gamma_0^*\gamma = 1$

Noiseless

$8$ $16$ $24$ $32$ $40$ $48$ $56$ $64$

Success rate

20dB

$8$ $16$ $24$ $32$ $40$ $48$ $56$ $64$

Success rate

14dB

$8$ $16$ $24$ $32$ $40$ $48$ $56$ $64$

Success rate

6dB

$8$ $16$ $24$ $32$ $40$ $48$ $56$ $64$

Success rate
**Numerical Experiments - Sparsity Case**

- **Truncated power iteration vs. $\ell_1$ minimization**
- **Some of the $\{\lambda_k\}_{k=1}^n$ are initialized randomly**

![Graph 1/2 random](image1)

![Graph 3/4 random](image2)

![Graph 1/2 random](image3)

![Graph 3/4 random](image4)
Deconvolution as BGPC

\[ Y = \text{diag}(\lambda)AX \]

- \( W X \) – Images
- \( F^{-1}\lambda \) – Point spread function
- \( F^{-1}Y \) – Measurements

- Support of \( X \) is known \(\implies\) Truncation step is simple
- Ignore rows where \(|\lambda_k|\) is small \(\implies\) \(\lambda\) is approximately flat \((0.01 \leq \lambda_i \leq 5)\)
Deconvolution as BGPC

\[ |\langle \dot{x}, x_t \rangle| \quad \frac{|\langle \dot{x}, x_t \rangle|}{\|\dot{x}\|_2 \|x_t\|_2} \quad |\langle \dot{\eta}, \eta_t \rangle| \quad \frac{|\langle \dot{\eta}, \eta_t \rangle|}{\|\dot{\eta}\|_2 \|\eta_t\|_2} \]
Sensor Array Processing as BGPC

\[ Y = \text{diag}(\lambda)AX + W \]

- Number of candidate DOAs \( m = 128 \)
- Number of sensors \( n = 32 \)
- Number of sources \( s = 4 \)
- Number of snapshots \( N = 64 \)

\[ A = [a(\theta_1), a(\theta_2), \ldots, a(\theta_m)] \]

- \( X \) – Gaussian random sources (support corresponds to DOA)
- \( W \) – Gaussian random noise (SNR = 10.5dB)
Sensor Array Processing as BGPC

- Number of DOAs $m = 128$
- Number of sensors $n = 32$
- Number of sources $s_0 = 4$
- $s_1 = s_0$

Space the sensors s.t.

\[
A = \begin{bmatrix}
m \\
n
\end{bmatrix}
\]

is a partial DFT matrix

Randomly spaced array

True DOAs:

Recovered DOAs:
Summary

- **BGPC as an eigenvector problem**
  - Subspace case: power iteration
  - Sparsity case: truncated power iteration

- **Theoretical guarantees**
  - Subspace case: near optimal numbers of sensors & snapshots
  - Joint sparsity case: near optimal number of sensors, suboptimal number of snapshots

- Robust against noise and bad initial estimate
Conclusions

- Bilinear Inverse Problems (BLIPS) are everywhere
- Results for Blind Deconvolution and Blind Gain & Phase Calibration
- Tight conditions for identifiability subject to subspace and/or sparsity constraints for “almost all” scenarios
- Data-efficient practical algorithms
- Challenges and Opportunities
Challenges

• Technical refinements
  – Optimal scaling of number of signals $N$ for BGPC with sparse signal models
  – Results for BGPC with general sparsity

• Guaranteed algorithms with optimum data requirements for general bilinear inverse problems

• Large scale, distributed, and online versions

• Theory and Algorithms for multi-linear problems

• Applications
Radio Telescope arrays

- Big Data
- Distributed across continents
- processed jointly to produce images

- Jansky Very Large Array (JVLA)
- Very Long Baseline Array (VLBA)
- Atacama Large Millimeter/submillimeter Array (ALMA)
- In 2024, the Square Kilometer Array (SKA)
From the Big Bang to Big Data

Daily raw data generated by the Square Kilometer Array:

15 million 64GB iPods
2x daily global internet traffic

Figure Credit: IBM Research. SKA to be completed 2024
https://youtu.be/w_q6kB2nCdw
Opportunities:
Blind SP to enable Inexpensive Imaging

MRI Cost ~ 1-3 $M
– strong & very uniform magnetic field (~50 ppm)

CT Cost ~ 0.5-2 $M – very precise, challenging design (>10g accelerations)

• Relaxed design, greatly reduced tolerances
→ low cost, but unknown time-varying characteristics
Thank you!
References I


References II


