Identifiability of Blind Deconvolution with Subspace or Sparsity Constraints

Yanjun Li

Joint work with Kiryung Lee and Yoram Bresler

Coordinated Science Laboratory
Department of Electrical and Computer Engineering
University of Illinois, Urbana-Champaign
Email: yli145@illinois.edu

SPARS 2015
July 6-9, 2015, Cambridge, UK
Blind Deconvolution

\[ u \ast v = z \]

Both \( u \) and \( v \) are unknown \(\implies\) 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. \(\implies\) The focus of this presentation
Problem Statement

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

\[ \text{find } (u, v) \]
\[ \text{s.t. } u \odot v = z, \]
\[ u \in \Omega_U, \ v \in \Omega_V. \]

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: $z = u_0 \odot v_0 \in \mathbb{C}^n$

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: $z = u_0 \otimes v_0 \in \mathbb{C}^n$

find $(u, v)$

s.t. $u \circledast v = z, \quad u \in \Omega_U, \; v \in \Omega_V$

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

- **Signal:** $u_0 = Dx_0$, 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:** $z = u_0 \ast v_0 = (Dx_0) \ast (Ey_0) \in \mathbb{C}^n$

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

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 \[\Omega_Y = \mathbb{C}^{m_2}\]
Identifiability up to Scaling, and Lifting

Definition (Identifiability up to scaling)
For (BD), the pair \((x_0, y_0)\) is **identifiable up to scaling** from the measurement \((Dx_0) \odot (Ey_0)\), if every solution \((x, y)\) satisfies \(x = \sigma x_0\) and \(y = \frac{1}{\sigma} y_0\) for some nonzero scalar \(\sigma\).

Lifting
Define \(G_{DE} : \mathbb{C}^{m_1 \times m_2} \to \mathbb{C}^n\) such that \(G_{DE}(xy^T) = (Dx) \odot (Ey)\), and \(M_0 = x_0 y_0^T \in \Omega_M = \{xy^T : x \in \Omega_X, y \in \Omega_Y\}\).

\((\text{BD})\) find \((x, y)\), s.t. \((Dx) \odot (Ey) = z\), \(x \in \Omega_X, y \in \Omega_Y\). \hspace{1cm} \Rightarrow \hspace{1cm} \text{(Lifted BD)}\) find \(M\), s.t. \(G_{DE}(M) = z\), \(M \in \Omega_M\).
Identifiability up to Scaling, and Lifting

Definition (Identifiability up to scaling)

For (BD), the pair \((x_0, y_0)\) is identifiable up to scaling from the measurement \((Dx_0) \circledast (Ey_0)\), if every solution \((x, y)\) satisfies \(x = \sigma x_0\) and \(y = \frac{1}{\sigma} y_0\) for some nonzero scalar \(\sigma\).

Lifting

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

(BD) find \((x, y)\),

s.t. \((Dx) \circledast (Ey) = z, \quad x \in \Omega_X, y \in \Omega_Y\).

\(\implies\) (Lifted BD) find \(M\),

s.t. \(G_{DE}(M) = z, \quad M \in \Omega_M\).
Previous Results (all based on a lifting formulation)

- Identifiability analysis
  - [Choudhary and Mitra, 2014]: canonical sparsity constraints
    - Lacks sample-complexity type interpretation

- Guaranteed recovery algorithms
  - [Ahmed, Recht, and Romberg, 2014]: nuclear norm minimization
  - [Ling and Strohmer, 2015]: $\ell_1$ norm minimization
  - [Lee, Y. Li, Junge, and Bresler, 2015]: alternating minimization
  - [Chi, 2015]: atomic norm minimization
  ✓ Constructive proof of uniqueness
    - Requires probabilistic assumptions and interpretations

Goal

- Identifiability in BD with more general bases or frames
- Algebraic analysis with minimal and deterministic assumptions
- Optimality in terms of sample complexities
Previous Results (all based on a lifting formulation)

- Identifiability analysis
  - [Choudhary and Mitra, 2014]: canonical sparsity constraints
    - Lacks sample-complexity type interpretation

- Guaranteed recovery algorithms
  - [Ahmed, Recht, and Romberg, 2014]: nuclear norm minimization
  - [Ling and Strohmer, 2015]: \( \ell_1 \) norm minimization
  - [Lee, Y. Li, Junge, and Bresler, 2015]: alternating minimization
  - [Chi, 2015]: atomic norm minimization
  ✓ Constructive proof of uniqueness
    - Requires probabilistic assumptions and interpretations

Goal

- Identifiability in BD with more general bases or frames
- Algebraic analysis with minimal and deterministic assumptions
- Optimality in terms of sample complexities
Sample Complexities for Uniqueness in BD

\[ z = (Dx_0) \odot (Ey_0) \]

**Theorem (Generic bases or frames)**

The pair \((x_0, y_0)\) is identifiable up to scaling from \((Dx_0) \odot (Ey_0)\) for almost all \(D \in \mathbb{C}^{n \times m_1}\) and \(E \in \mathbb{C}^{n \times m_2}\) if:

- **(subspace constraints)** \(n \geq m_1 m_2\)
- **(sparsity constraints)** \(n \geq 2s_1 s_2\)
- **(mixed constraints)** \(n \geq 2s_1 m_2\)
Proof Sketch (Subspace Constraints, Generic $D$ & $E$)

Lemma

If $n \geq m_1 m_2$, then for almost all $D \in \mathbb{C}^{n \times m_1}$ and $E \in \mathbb{C}^{n \times m_2}$, the following matrix $G_{DE}$ has full column rank:

$$G_{DE} \text{vec}(xy^T) = (Dx) \odot (Ey)$$

Lemma [Harikumar and Bresler, 1998] “Proof by Example”

- Suppose the entries of $G_{DE}$ are polynomials in the entries of $D$ and $E$.
- Suppose $G_{DE}$ has full column rank for at least one choice of $D$ and $E$.
- Then $G_{DE}$ has full column rank for almost all $D$ and $E$.

One good choice of $D$ & $E$ for $n \geq m_1 m_2$

$$F_n z = (F_n Dx) \odot (F_n Ey) = F_n G_{DE} \text{vec}(xy^T)$$ – In frequency domain

DFT matrix $\tilde{D}$ $\tilde{E}$ $\tilde{G}_{DE}$

$D$, $E$, and $G_{DE}$ have full rank submatrices.
Optimality?

Theorem (Generic bases or frames)

The pair \((x_0, y_0)\) is identifiable up to scaling from \((Dx_0) \otimes (Ey_0)\) for almost all \(D \in \mathbb{C}^{n \times m_1}\) and almost all \(E \in \mathbb{C}^{n \times m_2}\) if:

- (subspace constraints) \(n \geq m_1 m_2\)
- (sparsity constraints) \(n \geq 2s_1 s_2\)
- (mixed constraints) \(n \geq 2s_1 m_2\)

Suspect this is suboptimal (# df = \(m_1 + m_2 - 1\) for subspace constraints)

Q: Can we get optimal sample complexities?

A: Yes, if we consider more specialized scenarios.
Optimality?

**Theorem (Generic bases or frames)**

The pair \((x_0, y_0)\) is identifiable up to scaling from \((Dx_0) \odot (Ey_0)\) for almost all \(D \in \mathbb{C}^{n \times m_1}\) and almost all \(E \in \mathbb{C}^{n \times m_2}\) if:

- **(subspace constraints)** \(n \geq m_1 m_2\)
- **(sparsity constraints)** \(n \geq 2s_1 s_2\)
- **(mixed constraints)** \(n \geq 2s_1 m_2\)

Suspect this is suboptimal (# df = \(m_1 + m_2 - 1\) for subspace constraints)

Q: Can we get optimal sample complexities?

A: Yes, if we consider more specialized scenarios.
Optimality?

Theorem (Generic bases or frames)

The pair \((x_0, y_0)\) is identifiable up to scaling from \((Dx_0) \odot (Ey_0)\) for almost all \(D \in \mathbb{C}^{n \times m_1}\) and almost all \(E \in \mathbb{C}^{n \times m_2}\) if:

- (subspace constraints) \(n \geq m_1 m_2\)
- (sparsity constraints) \(n \geq 2s_1 s_2\)
- (mixed constraints) \(n \geq 2s_1 m_2\)

Suspect this is suboptimal (# df = \(m_1 + m_2 - 1\) for subspace constraints)

Q: Can we get optimal sample complexities?

A: Yes, if we consider more specialized scenarios.
Optimality?

Theorem (Generic bases or frames)

The pair \((x_0, y_0)\) is identifiable up to scaling from \((Dx_0) \circledast (Ey_0)\) for almost all \(D \in \mathbb{C}^{n \times m_1}\) and almost all \(E \in \mathbb{C}^{n \times m_2}\) if:

- (subspace constraints) \(n \geq m_1 m_2\)
- (sparsity constraints) \(n \geq 2s_1 s_2\)
- (mixed constraints) \(n \geq 2s_1 m_2\)

Suspect this is suboptimal (# df = \(m_1 + m_2 - 1\) for subspace constraints)

Q: Can we get optimal sample complexities?

A: Yes, if we consider more specialized scenarios.
Sub-band Structured Basis

**Definition**

- $\widetilde{E}(:,k) := F_n E(:,k)$ – the DFT of the $k$th atom (column) in $E$
- $J_k$ – the support of $\widetilde{E}(:,k)$
- $\hat{J}_k$ – passband
- $\ell_k := |\hat{J}_k|$ – bandwidth

DFTs of the atoms in $E$

- $\widetilde{E}(:,1)$
- $\widetilde{E}(:,2)$
- $\widetilde{E}(:,3)$

DFTs of some possible signals

- $\widetilde{E}y_1$
- $\widetilde{E}y_2$
- $\widetilde{E}y_3$
Sub-band Structured Basis

Definition

- $\tilde{E}(;k) := F_n E(;k)$ – the DFT of the $k$th atom (column) in $E$
- $J_k$ – the support of $\tilde{E}(;k)$
- $\hat{J}_k$ – passband
- $\ell_k := |\hat{J}_k|$ – bandwidth

DFTs of the atoms in $E$

$\tilde{E}(;1)$

$\tilde{E}(;2)$

$\tilde{E}(;3)$

DFTs of some possible signals

$\tilde{E}y_1$

$\tilde{E}y_2$

$\tilde{E}y_3$
Sub-band Structured Basis

Definition

- $\widehat{E}(:,k) := F_n E(:,k)$ – the DFT of the $k$th atom (column) in $E$
- $J_k$ – the support of $\widehat{E}(:,k)$
- $\widehat{J}_k$ – passband
- $\ell_k := |\widehat{J}_k|$ – bandwidth
BD with a Sub-band Structured Basis

Blind Deconvolution: given $D$, $E$, & $z$, find $x$ & $y$

\[ z = (Dx) \odot (Ey) \]

Blind Gain and Phase Calibration

\[ z_i = (\tilde{E}_\phi) \odot (Ax_i), \]

column of $A$ – array response

support of $x$ – DOA

structure of $\tilde{E}$ – sensor groups

entry of $\phi$ – gain and phase

\[ \tilde{E} = \vspace{1cm} \]
BD with a Sub-band Structured Basis

Blind Deconvolution: given $D$, $E$, & $z$, find $x$ & $y$

Blind Gain and Phase Calibration
BD with a Sub-band Structured Basis
Sufficient Conditions with (Essentially) Optimal Sample Complexities

Theorem (Sub-band structured basis)

Suppose $E$ forms a sub-band structured basis, $x_0 \in \mathbb{C}^{m_1}$ is nonzero, and all the entries of $y_0 \in \mathbb{C}^{m_2}$ are nonzero. If the sum of all the bandwidths satisfies

- **(subspace constraints)** \[ \sum_{k=1}^{m_2} \ell_k \geq m_1 + m_2 - 1 \]
- **(mixed constraints)** \[ \sum_{k=1}^{m_2} \ell_k \geq 2s_1 + m_2 - 1 \]

then for almost all $D \in \mathbb{C}^{n \times m_1}$, the pair $(x_0, y_0)$ is identifiable up to scaling.
Proof Sketch


In (BD), the pair \((x_0, y_0) \ (x_0 \neq 0, y_0 \neq 0)\) is identifiable up to scaling if and only if the following two conditions are met:

1. If there exists \((x, y) \in \Omega_X \times \Omega_Y\) such that \((Dx) \otimes (Ey) = (Dx_0) \otimes (Ey_0)\), then \(x = \sigma x_0\) for some nonzero \(\sigma \in \mathbb{C}\).

2. If there exists \(y \in \Omega_Y\) such that \((Dx_0) \otimes (Ey) = (Dx_0) \otimes (Ey_0)\), then \(y = y_0\).

Condition 2 is easy to verify.

Condition 1 relies on the following fact:
If \(D\) is generic, and \((x, y) \in \Omega_X \times \Omega_Y\) satisfies \((Dx) \otimes (Ey) = (Dx_0) \otimes (Ey_0)\), then

\[P_{x_0 \perp} x = 0.\]

Hence \(x = \sigma x_0\) for some scalar \(\sigma\).
Theorem (Necessary conditions)

If the supports $J_k$ ($1 \leq k \leq m_2$) partition the DFT frequency range, then $(x_0, y_0)$ is identifiable up to scaling only if:

- (subspace constraints) $n \geq m_1 + m_2 - 1$
- (mixed constraints) $n \geq s_1 + m_2 - 1$
BD with a Sub-band Structured Basis

Necessary Conditions with *Optimal Sample Complexities*

DFTs of the atoms in $E$

\[
\tilde{E}(::, 1) \quad \tilde{E}(::, 2) \quad \tilde{E}(::, 3)
\]

**Theorem (Necessary conditions)**

*If the supports $J_k$ ($1 \leq k \leq m_2$) partition the DFT frequency range, then $(x_0, y_0)$ is identifiable up to scaling only if*

- *(subspace constraints)* $n \geq m_1 + m_2 - 1$
- *(mixed constraints)* $n \geq s_1 + m_2 - 1$

$Necessary\ Conditions$  $Sufficient\ Conditions$

$n \geq m_1 + m_2 - 1$  $n \geq 2s_1 + m_2 - 1$
Conclusions

- The first algebraic sample complexities for **unique** blind deconvolution

<table>
<thead>
<tr>
<th>Constraints Type</th>
<th>Subspace Constraints</th>
<th>Sparsity Constraints</th>
<th>Mixed Constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Generic bases or frames</strong></td>
<td>$n \geq m_1 m_2$</td>
<td>$n \geq 2s_1 s_2$</td>
<td>$n \geq 2s_1 m_2$</td>
</tr>
<tr>
<td><strong>A sub-band structured basis</strong></td>
<td>$n \geq m_1 + m_2 - 1$ (optimal)</td>
<td>$n \geq 2s_1 + m_2 - 1$ (nearly optimal)</td>
<td></td>
</tr>
</tbody>
</table>

- **Generic bases or frames** $\Rightarrow$ violated on a set of Lebesgue measure zero

Thank you!
References


Proof Sketch


In (BD), the pair \((x_0, y_0) (x_0 \neq 0, y_0 \neq 0)\) is identifiable up to scaling if and only if the following two conditions are met:

1. If there exists \((x, y) \in \Omega_X \times \Omega_Y\) such that \((Dx) \circ (Ey) = (Dx_0) \circ (Ey_0)\), then \(x = \sigma x_0\) for some nonzero \(\sigma \in \mathbb{C}\).
2. If there exists \(y \in \Omega_Y\) such that \((Dx_0) \circ (Ey) = (Dx_0) \circ (Ey_0)\), then \(y = y_0\).

Condition 2 is easy to verify.

Condition 1 relies on the following fact:
If \(D\) is generic, and \((x, y) \in \Omega_X \times \Omega_Y\) satisfies \((Dx) \circ (Ey) = (Dx_0) \circ (Ey_0)\), then

\[
\text{diag}(\tilde{E}y)\tilde{D}x = (\tilde{D}x) \circ (\tilde{E}y) = (\tilde{D}x_0) \circ (\tilde{E}y_0) = \text{diag}(\tilde{E}y_0)\tilde{D}x_0.
\]

Consider the passband \(\hat{J}_k, k = 1, 2, \ldots, m_2\),

\[
P_{x_0^\perp} x \in x_0^\perp \cap \left( \mathcal{R}(\tilde{D}(\hat{J}_k;:)^*) \cap x_0^\perp \right)^\perp = x_0^\perp \cap \mathcal{V}_k^\perp.
\]

Hence

\[
P_{x_0^\perp} x \in x_0^\perp \cap \mathcal{V}_1^\perp \cap \mathcal{V}_2^\perp \cap \cdots \cap \mathcal{V}_{m_2}^\perp.
\]
Proof Sketch

\[ P_{x_0} x \in x_0^\perp \cap V_1^\perp \cap V_2^\perp \cap \cdots \cap V_{m_2}^\perp \]

For a generic matrix \( D \), the subspaces \( V_1, V_2, \ldots, V_{m_2} \) are generic subspaces of \( x_0^\perp \), with \( \dim(V_k) = \ell_k - 1 \). If \( \sum_{k=1}^{m_2} \ell_k \geq m_1 + m_2 - 1 \), i.e., \( \sum_{k=1}^{m_2} (\ell_k - 1) \geq m_1 - 1 \), then

\[ \sum_{k=1}^{m_2} V_k = x_0^\perp, \]

\[ \text{span}(x_0) + \sum_{k=1}^{m_2} V_k = \mathbb{C}^{m_1}, \]

\[ P_{x_0} x \in \left( \text{span}(x_0) + \sum_{k=1}^{m_2} V_k \right)^\perp = \{0\}. \]

Hence \( x = \sigma x_0 \) for some scalar \( \sigma \).