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 7, 2015, Cambridge, UK
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: \( z = u_0 \ast v_0 \in \mathbb{C}^n \)

\[
\begin{align*}
\text{find } (u, v) \\
\text{s.t. } u \ast v &= z, \\
&\quad u \in \Omega_U, v \in \Omega_V.
\end{align*}
\]

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

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

Three scenarios:

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

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

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

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:** \( z = u_0 \odot v_0 = (Dx_0) \odot (Ey_0) \in \mathbb{C}^n \)

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

Three scenarios:

1. **Subspace constraints:**
   \[ \Omega_X = \mathbb{C}^{m_1} \quad \text{and} \quad \Omega_Y = \mathbb{C}^{m_2} \]

2. **Sparsity constraints:**
   \[ \Omega_X = \{ x \in \mathbb{C}^{m_1} : \|x\|_0 \leq s_1 \} \quad \text{and} \quad \Omega_Y = \{ y \in \mathbb{C}^{m_2} : \|y\|_0 \leq s_2 \} \]

3. **Mixed constraints:**
   \[ \Omega_X = \{ x \in \mathbb{C}^{m_1} : \|x\|_0 \leq s_1 \} \quad \text{and} \quad \Omega_Y = \mathbb{C}^{m_2} \]
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 \((Dx_0) \otimes (Ey_0)\), if every solution \((x, y)\) satisfies \(x = \sigma x_0\) and \(y = \frac{1}{\sigma} y_0\) for some nonzero scalar \(\sigma\).

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

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

(Lifted BD) find \(M\),
\[\text{s.t.} \quad G_{DE}(M) = z, \quad M \in \Omega_M.\]
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 \((Dx_0) \odot (Ey_0)\), if every solution \((x, y)\) satisfies \(x = \sigma x_0\) and \(y = \frac{1}{\sigma} y_0\) for some nonzero scalar \(\sigma\).

Lifting

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

\[(\text{BD}) \quad \text{find} \quad (x, y), \quad \text{s.t.} \quad (Dx) \odot (Ey) = z, \quad x \in \Omega_X, \quad y \in \Omega_Y. \quad \quad \Rightarrow \quad \quad (\text{Lifted BD}) \quad \text{find} \quad M, \quad \text{s.t.} \quad G_{DE}(M) = z, \quad M \in \Omega_M.\]
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

\[ z = (D x_0) \circledast (E y_0) \]

**Theorem (Generic bases or frames)**

The pair \((x_0, y_0)\) is identifiable up to scaling from \((D x_0) \circledast (E y_0)\) 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:

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

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 $n \geq m_1 m_2$

$$\tilde{F}_n z = (\tilde{F}_n D x) \odot (\tilde{F}_n E y) = \tilde{F}_n G_{DE} \operatorname{vec}(xy^T) \quad \text{— In frequency domain}$$

DFT matrix $\tilde{D}$ $\tilde{E}$ $\tilde{G}_{DE}$
The pair \((x_0, y_0)\) is identifiable up to scaling from \((D x_0) \odot (E y_0)\) 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 \((Dx_0) \otimes (Ey_0)\) 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 \((Dx_0) \ast (Ey_0)\) 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

\[
\begin{align*}
\widetilde{E}(::,1) & \quad \text{DFTs of the atoms in } E \\
\widetilde{E}(::,2) & \quad \text{DFTs of some possible signals} \\
\widetilde{E}(::,3)
\end{align*}
\]

\[
\begin{align*}
\widetilde{E} y_1 \\
\widetilde{E} y_2 \\
\widetilde{E} y_3
\end{align*}
\]
Sub-band Structured Basis

Definition

- $\tilde{E}(::k) := F_nE(::k)$ – 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

- $\tilde{E}(::k) := F_n E(::k)$ – 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
BD with a Sub-band Structured Basis

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

$$z = (Dx) \circ (Ey)$$

Blind Gain and Phase Calibration

$$z_i = (\tilde{E}\phi) \circ (Ax_i),$$

| Column of $A$ | array response |
| Support of $x$ | DOA |
| Structure of $\tilde{E}$ | sensor groups |
| Entry of $\phi$ | gain and phase |
BD with a Sub-band Structured Basis

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

$z = (Dx) ∗ (Ey)$

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 \((x_0, y_0) \neq (0, 0)\) 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 $E$

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) $n \geq s_1 + m_2 - 1$  

**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**
Conclusions

- The first algebraic sample complexities for unique blind deconvolution

**Generic bases or frames:**
- Subspace constraints: \( n \geq m_1m_2 \)
- Sparsity constraints: \( n \geq 2s_1s_2 \)
- Mixed constraints: \( n \geq 2s_1m_2 \)

**A sub-band structured basis:**
- Subspace constraints: \( n \geq m_1 + m_2 - 1 \) (optimal)
- Mixed constraints: \( n \geq 2s_1 + m_2 - 1 \) (nearly optimal)

- Generic bases or frames \( \Rightarrow \) violated on a set of Lebesgue measure zero

**Journal version:** http://arxiv.org/abs/1505.03399

**Blind gain and phase calibration:** http://arxiv.org/abs/1501.06120
Thank you!


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

Consider the passband \(\tilde{J}_k, k = 1, 2, \ldots, m_2\),
\[
P_{x_0^\perp} x \in x_0^\perp \bigcap \left(\mathcal{R}(\tilde{D}^{(\tilde{J}_k,:)*}) \bigcap x_0^\perp \right)^\perp = x_0^\perp \bigcap V_k^\perp.
\]
Hence
\[
P_{x_0^\perp} x \in x_0^\perp \bigcap V_1^\perp \bigcap V_2^\perp \bigcap \cdots \bigcap V_{m_2}^\perp.
\]
Proof Sketch

\[ P_{x_0^\perp} x \in x_0^\perp \bigcap \mathcal{V}_1^\perp \bigcap \mathcal{V}_2^\perp \bigcap \cdots \bigcap \mathcal{V}_{m_2}^\perp \]

For a generic matrix \( D \), the subspaces \( \mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_{m_2} \) are generic subspaces of \( x_0^\perp \), with \( \dim(\mathcal{V}_k) = \ell_k - 1 \). 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

\[ \sum_{k=1}^{m_2} \mathcal{V}_k = x_0^\perp , \]

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

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

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