Identifiability of Blind Deconvolution with Subspace or Sparsity Constraints

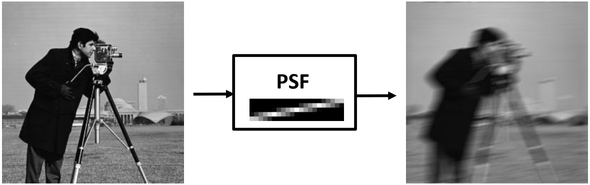
Yanjun Li

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



signal: u $\circledast$ filter: v = measurement: z

- Both u and v are unknown $\Longrightarrow$ 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. $\Longrightarrow$ The focus of this presentation

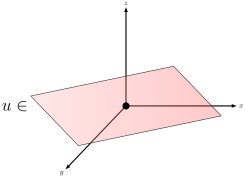
Problem Statement

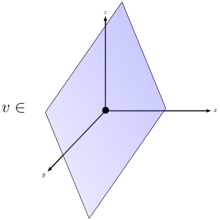
- Signal: $u_0 \in \mathbb{C}^n$
- Filter: $v_0 \in \mathbb{C}^n$
- Measurement: $z = u_0 \circledast v_0 \in \mathbb{C}^n$

$$\begin{aligned} \text{find} \quad & (u, v) \\ \text{s.t.} \quad & u \circledast v = z, \\ & u \in \Omega_{\mathcal{U}}, \; v \in \Omega_{\mathcal{V}}. \end{aligned}$$

Three scenarios:

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





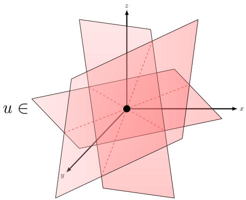
Problem Statement

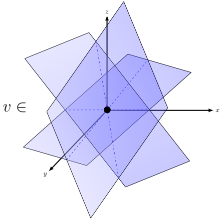
- Signal: $u_0 \in \mathbb{C}^n$
- Filter: $v_0 \in \mathbb{C}^n$
- Measurement: $z = u_0 \circledast v_0 \in \mathbb{C}^n$

$$\begin{aligned} \text{find} \quad & (u, v) \\ \text{s.t.} \quad & u \circledast v = z, \\ & u \in \Omega_{\mathcal{U}}, \ v \in \Omega_{\mathcal{V}}. \end{aligned}$$

Three scenarios:

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





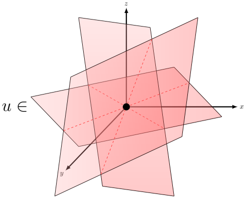
Problem Statement

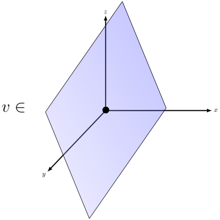
- Signal: $u_0 \in \mathbb{C}^n$
- Filter: $v_0 \in \mathbb{C}^n$
- Measurement: $z = u_0 \circledast v_0 \in \mathbb{C}^n$

$$\begin{aligned} \text{find} \quad & (u, v) \\ \text{s.t.} \quad & u \circledast v = z, \\ & u \in \Omega_{\mathcal{U}}, \; v \in \Omega_{\mathcal{V}}. \end{aligned}$$

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

$$\text{(BD)} \quad \begin{aligned} \text{find} \;\; & (x, y) \\ \text{s.t.} \;\; & (Dx) \circledast (Ey) = z, \\ & x \in \Omega_{\mathcal{X}}, \; y \in \Omega_{\mathcal{Y}}. \end{aligned}$$

Three scenarios:

1. Subspace constraints:
   $\Omega_{\mathcal{X}} = \mathbb{C}^{m_1}$ and $\Omega_{\mathcal{Y}} = \mathbb{C}^{m_2}$
2. Sparsity constraints:
   $\Omega_{\mathcal{X}} = \{x \in \mathbb{C}^{m_1} : \|x\|_0 \le s_1\}$ and $\Omega_{\mathcal{Y}} = \{y \in \mathbb{C}^{m_2} : \|y\|_0 \le s_2\}$
3. Mixed constraints:
   $\Omega_{\mathcal{X}} = \{x \in \mathbb{C}^{m_1} : \|x\|_0 \le s_1\}$ and $\Omega_{\mathcal{Y}} = \mathbb{C}^{m_2}$

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 $(Dx_0) \circledast (Ey_0)$, if every solution (x, y) satisfies $x = \sigma x_0$ and $y = \frac{1}{\sigma} y_0$ for some nonzero scalar σ.

Lifting

Define $\mathcal{G}_{DE} : \mathbb{C}^{m_1 \times m_2} \to \mathbb{C}^n$ such that $\mathcal{G}_{DE}(xy^T) = (Dx) \circledast (Ey)$, and $M_0 = x_0 y_0^T \in \Omega_{\mathcal{M}} = \{xy^T : x \in \Omega_{\mathcal{X}}, y \in \Omega_{\mathcal{Y}}\}$.

$$\text{(BD)} \quad \begin{aligned} \text{find} \quad & (x, y), \\ \text{s.t.} \quad & (Dx) \circledast (Ey) = z, \\ & x \in \Omega_{\mathcal{X}},\ y \in \Omega_{\mathcal{Y}}. \end{aligned} \qquad \Longrightarrow \qquad \text{(Lifted BD)} \quad \begin{aligned} \text{find} \quad & M, \\ \text{s.t.} \quad & \mathcal{G}_{DE}(M) = z, \\ & M \in \Omega_{\mathcal{M}}. \end{aligned}$$

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 σ.

Lifting

Define $\mathcal{G}_{DE} : \mathbb{C}^{m_1 \times m_2} \to \mathbb{C}^n$ such that $\mathcal{G}_{DE}(xy^T) = (Dx) \circledast (Ey)$, and $M_0 = x_0 y_0^T \in \Omega_{\mathcal{M}} = \{xy^T : x \in \Omega_{\mathcal{X}}, y \in \Omega_{\mathcal{Y}}\}$.

$$\text{(BD)} \quad \begin{aligned} \text{find} \quad & (x, y), \\ \text{s.t.} \quad & (Dx) \circledast (Ey) = z, \\ & x \in \Omega_{\mathcal{X}},\ y \in \Omega_{\mathcal{Y}}. \end{aligned} \qquad \Longrightarrow \qquad \text{(Lifted BD)} \quad \begin{aligned} \text{find} \quad & M, \\ \text{s.t.} \quad & \mathcal{G}_{DE}(M) = z, \\ & M \in \Omega_{\mathcal{M}}. \end{aligned}$$

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]: ℓ_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

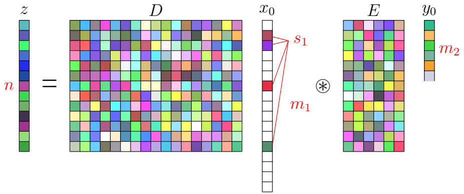
- Guaranteed recovery algorithms
  - [Ahmed, Recht, and Romberg, 2014]: nuclear norm minimization
  - [Ling and Strohmer, 2015]: ℓ_1 norm minimization
  - [Lee, Y. Li, Junge, and Bresler, 2015]: alternating minimization
  - [Chi, 2015]: atomic norm minimization

  ✓ 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



Mixed constraints

$$z = (Dx_0) \circledast (Ey_0)$$

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* $E \in \mathbb{C}^{n \times m_2}$ *if:*

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

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:

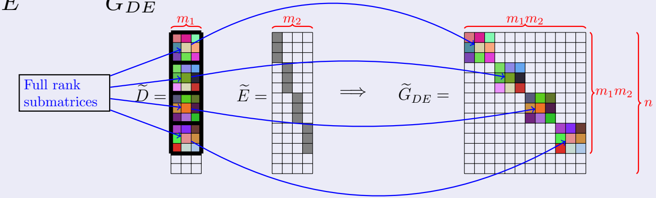
$$G_{DE}\,\mathrm{vec}(xy^T) = (Dx) \circledast (Ey)$$

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 $n \geq m_1 m_2$

$$\underbrace{F_n}_{\text{DFT matrix}} z = (\underbrace{F_n D}_{\widetilde{D}} x) \odot (\underbrace{F_n E}_{\widetilde{E}} y) = \underbrace{F_n G_{DE}}_{\widetilde{G}_{DE}} \mathrm{vec}(xy^T)$$ – In frequency domain



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)* $n \geq 2 s_1 s_2$
- *(mixed constraints)* $n \geq 2 s_1 m_2$

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)* $n \geq 2 s_1 s_2$
- *(mixed constraints)* $n \geq 2 s_1 m_2$

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)* $n \geq 2 s_1 s_2$
- *(mixed constraints)* $n \geq 2 s_1 m_2$

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)* $n \geq 2 s_1 s_2$
- *(mixed constraints)* $n \geq 2 s_1 m_2$

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

Q: Can we get optimal sample complexities?

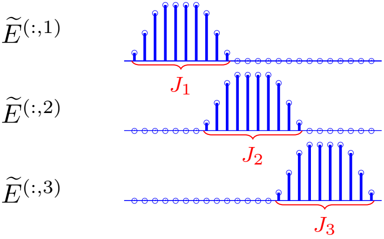
A: Yes, if we consider more specialized scenarios.

Sub-band Structured Basis

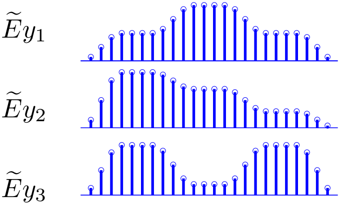
Definition

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

DFTs of the atoms in E



DFTs of some possible signals

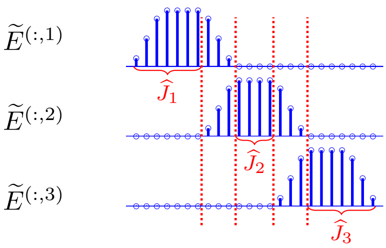


Sub-band Structured Basis

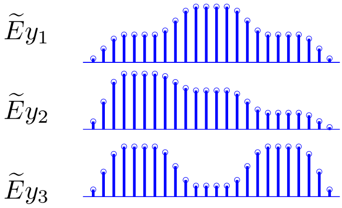
Definition

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

DFTs of the atoms in E



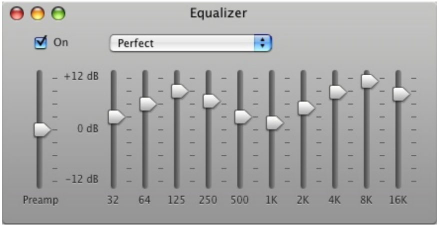
DFTs of some possible signals



Sub-band Structured Basis

Definition

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

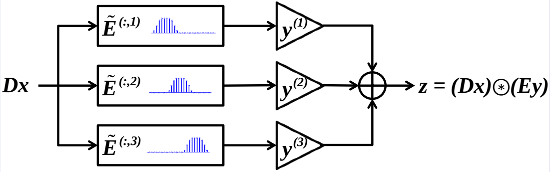


☑ On

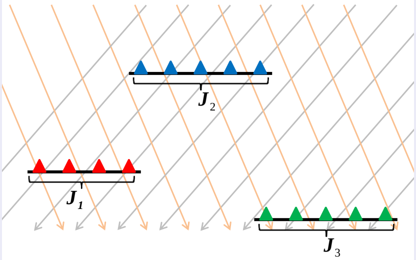
Perfect

BD with a Sub-band Structured Basis

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

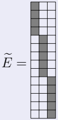


Blind Gain and Phase Calibration



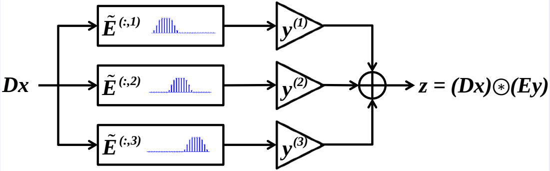
$$z_i = (\widetilde{E}\phi) \odot (Ax_i),$$

column of A – array response
support of x – DOA
structure of $\widetilde{E}$ – sensor groups
entry of ϕ – gain and phase

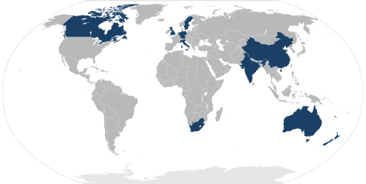


BD with a Sub-band Structured Basis

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

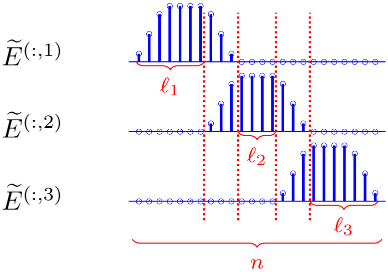


Blind Gain and Phase Calibration



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

Lemma [Y. Li, Lee, & Bresler, 2015] Identifiability in bilinear inverse problems: `http://arxiv.org/abs/1501.06120`

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 $(x, y) \in \Omega_{\mathcal{X}} \times \Omega_{\mathcal{Y}}$ such that $(Dx) \circledast (Ey) = (Dx_0) \circledast (Ey_0)$, then $x = \sigma x_0$ for some nonzero $\sigma \in \mathbb{C}$.
2. If there exists $y \in \Omega_{\mathcal{Y}}$ such that $(Dx_0) \circledast (Ey) = (Dx_0) \circledast (Ey_0)$, then $y = y_0$.

Condition 2 is easy to verify.

Condition 1 relies on the following fact:
If D is generic, and $(x, y) \in \Omega_{\mathcal{X}} \times \Omega_{\mathcal{Y}}$ satisfies $(Dx) \circledast (Ey) = (Dx_0) \circledast (Ey_0)$, then

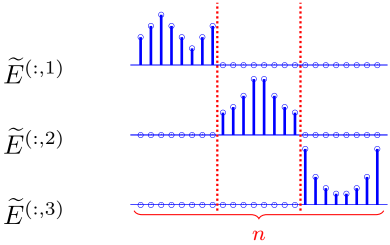
$$P_{x_0^\perp} x = 0.$$

Hence $x = \sigma x_0$ for some scalar σ.

BD with a Sub-band Structured Basis

Necessary Conditions with *Optimal Sample Complexities*

DFTs of the atoms in E



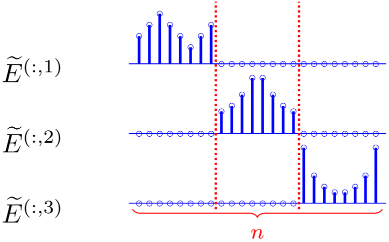
Theorem (Necessary conditions)

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

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)* $n \geq m_1 + m_2 - 1$ $n \geq m_1 + m_2 - 1$
- *(mixed constraints)* $n \geq s_1 + m_2 - 1$ $n \geq 2s_1 + m_2 - 1$

Necessary Conditions *Sufficient Conditions*

Conclusions

- The first algebraic sample complexities for unique blind deconvolution

Generic bases or frames:

  - Subspace constraints: $n \geq m_1 m_2$
  - Sparsity constraints: $n \geq 2 s_1 s_2$
  - Mixed constraints: $n \geq 2 s_1 m_2$

A sub-band structured basis:

  - Subspace constraints: $n \geq m_1 + m_2 - 1$ (optimal)
  - Mixed constraints: $n \geq 2 s_1 + m_2 - 1$ (nearly optimal)

- Generic bases or frames ⇒ 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!

References

A. Ahmed, B. Recht, and J. Romberg. Blind deconvolution using convex programming. *IEEE Trans. Inf. Theory*, 60(3):1711–1732, Mar 2014.

Y. Chi. Guaranteed blind sparse spikes deconvolution via lifting and convex optimization. *arXiv preprint arXiv:1506.02751*, 2015.

S. Choudhary and U. Mitra. Sparse blind deconvolution: What cannot be done. In *ISIT*, pages 3002–3006. IEEE, June 2014.

G. Harikumar and Y. Bresler. FIR perfect signal reconstruction from multiple convolutions: minimum deconvolver orders. *IEEE Trans. Signal Process.*, 46(1): 215–218, Jan 1998.

K. Lee, Y. Li, M. Junge, and Y. Bresler. Stability in blind deconvolution of sparse signals and reconstruction by alternating minimization. *SampTA*, 2015.

Y. Li, K. Lee, and Y. Bresler. A unified framework for identifiability analysis in bilinear inverse problems with applications to subspace and sparsity models. *arXiv preprint arXiv:1501.06120*, 2015.

S. Ling and T. Strohmer. Self-calibration and biconvex compressive sensing. *arXiv preprint arXiv:1501.06864*, 2015.

Proof Sketch

Lemma [Y. Li, Lee, & Bresler, 2015]

Identifiability in bilinear inverse problems: `http://arxiv.org/abs/1501.06120`

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 $(x, y) \in \Omega_{\mathcal{X}} \times \Omega_{\mathcal{Y}}$ such that $(Dx) \circledast (Ey) = (Dx_0) \circledast (Ey_0)$, then $x = \sigma x_0$ for some nonzero $\sigma \in \mathbb{C}$.
2. If there exists $y \in \Omega_{\mathcal{Y}}$ such that $(Dx_0) \circledast (Ey) = (Dx_0) \circledast (Ey_0)$, then $y = y_0$.

Condition 2 is easy to verify.

Condition 1 relies on the following fact:
If D is generic, and $(x, y) \in \Omega_{\mathcal{X}} \times \Omega_{\mathcal{Y}}$ satisfies $(Dx) \circledast (Ey) = (Dx_0) \circledast (Ey_0)$, then

$$\mathrm{diag}(\widetilde{E}y)\widetilde{D}x = (\widetilde{D}x) \odot (\widetilde{E}y) = (\widetilde{D}x_0) \odot (\widetilde{E}y_0) = \mathrm{diag}(\widetilde{E}y_0)\widetilde{D}x_0.$$

Consider the passband $\widehat{J}_k$, $k = 1, 2, \cdots, m_2$,

$$P_{x_0^\perp} x \in x_0^\perp \bigcap \left(\mathcal{R}(\widetilde{D}^{(\widehat{J}_k, :)*}) \bigcap x_0^\perp \right)^\perp = x_0^\perp \bigcap \mathcal{V}_k^\perp.$$

Hence

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

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 $\mathcal{V}_1, \mathcal{V}_2, \cdots, \mathcal{V}_{m_2}$ 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

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

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

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

Hence $x = \sigma x_0$ for some scalar σ.