ON IDENTIFIABILITY IN BILINEAR INVERSE PROBLEMS

Sunav Choudhary and Urbashi Mitra

Ming Hsieh Dept. of Electrical Engineering
University of Southern California
{sunavcho,ubli}@usc.edu

ABSTRACT

This paper considers identifiability and recoverability in bilinear inverse problems which is relevant to blind deconvolution and matrix factorization. It is shown that bilinear inverse problems can be posed as rank-1 matrix recovery problems subject to linear constraints. Sufficient conditions for identifiability are developed for the cases when rank-2 matrices are present in the null space of the linear operator. Signal recovery using the nuclear norm heuristic for rank-1 matrix recovery is considered and simple conditions for success are provided.

***Index Terms*—** identifiability, bilinear inverse problems, matrix recovery

1. INTRODUCTION

Blind deconvolution, matrix factorization, etc. are nonlinear inverse problems of tremendous significance in signal processing. A key question is when unique solutions to these bilinear inverse problems exist. In this paper, we consider a rank-1 matrix recovery formulation for these bilinear inverse problems, enabling the unification of these problems. We address identifiability and our results can be extended to incorporate further constraints such as sparsity or approximate bilinearity. Our characterization relies on the properties of the null space of the linear operator acting on the rank-1 matrix in question. For many cases of interest, this null space admits a simple characterization. Further, the effect of additional constraints can be readily inferred from the resulting change in the rank-2 null space of the resultant linear operator. The formulation of the bilinear inverse problem as low rank matrix recovery enables the use of nuclear norm minimization [1] to determine solutions via the existing efficient convex optimization solvers.

We do not attempt to review the vast prior art on blind deconvolution, dictionary learning and matrix factorization. Blind deconvolution with smoothness and statistical priors has been surveyed in [2, 3] and sparsity priors in [4]. While identifiability results for blind deconvolution are overviewed in [5], incorporating sparsity is not straightforward. Non-negative matrix factorization and identifiability was examined in [6] exploiting geometry. To the best of our knowledge, low-rank matrix recovery methods have not been applied to study identifiability in these previous problems.

A common strategy employed in bilinear inverse problems is based on alternatively holding one input fixed and solving for the other (*e.g.* [7, 8]). Such a heuristic converges to a fixed point [9] and works well for certain applications (*e.g.* image processing); however, convergence is not robust to arbitrary initialization. For the context of random channel coding, blind deconvolution with sparsity priors via a convex relaxation of rank-1 matrix recovery is developed in [10] and some recoverability results are proposed in [11] under the assumption of known support on one input. We exploit a similar strategy herein for the more general bilinear inverse problem, but do not assume any knowledge of sparse support. A necessary and sufficient condition for successful recovery of *all* low-rank matrices using nuclear norm minimization is developed in [12] which could be specialized to rank-1 matrix recovery. While [13] shows that this condition is likely to be satisfied with high probability for random linear operators drawn from a suitable Gaussian ensemble, the results do not carry over for bilinear inverse problems, as the measurement matrices defining the respective problems are fixed and not drawn from a distribution. We develop two sufficient conditions on the input signals which, when satisfied, guarantee recoverability by the nuclear norm minimization heuristic for the bilinear inverse problem in question. These results are applicable even in the presence of rank-2 matrices in the null space of the linear operator.

Compressibility of the output of a bilinear system is examined in [14, 15] and methods for original signal recovery from the compressed signals are also developed. Our problem is more challenging: input signal recovery from the output of the bilinear system. Finally, we note that identifiability in low-rank matrix completion is studied in [16] using algebraic and combinatorial conditions; the additional generality comes at the price of the complexity of the characterization.

This paper makes the following contributions:

1. casting the bilinear inverse problem as a rank-1 matrix recovery problem,
2. provision of sufficient conditions for identifiability based on the null-space of the linear matrix operator, and
3. provision of sufficient conditions for recoverability using the nuclear norm heuristic for rank-1 matrix recovery.

2. SYSTEM MODEL

2.1. Bilinear Maps

Definition 1 (Bilinear Map)**.** *A mapping $\boldsymbol{S}\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^p$ is called a bilinear map if $\boldsymbol{S}(\cdot, \boldsymbol{y}) : \mathbb{R}^m \to \mathbb{R}^p$ is a linear map $\forall \boldsymbol{y} \in \mathbb{R}^n$ and $\boldsymbol{S}(\boldsymbol{x}, \cdot) : \mathbb{R}^n \to \mathbb{R}^p$ is a linear map $\forall \boldsymbol{x} \in \mathbb{R}^m$. If $p = 1$, $\boldsymbol{S}$ is called a bilinear form.*

We shall consider a generic bilinear system model,

$$\boldsymbol{z} = \boldsymbol{S}(\boldsymbol{x}, \boldsymbol{y}), \tag{1}$$

This research has been funded in part by the following grants and organizations: ONR N00014-09-1-0700, NSF CNS-0832186, NSF CNS-0821750 (MRI), NSF CCF-0917343, NSF CCF-1117896 and DOT CA-26-7084-00.

where $\boldsymbol{z}$ is the vector of observations and $\boldsymbol{S}(\cdot,\cdot)$ is a bilinear map. Let $\phi_j\colon \mathbb{R}^p \to \mathbb{R}$ be the j^{th} coordinate projection operator. Clearly, ϕ_j is a linear operator and hence the composition $\phi_j \circ \boldsymbol{S}\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ is a bilinear form. Because $\boldsymbol{S}$ is a finite dimensional operator, it is a bounded operator. Hence $\exists \boldsymbol{S}_j \in \mathbb{R}^{m\times n}$ such that $\boldsymbol{S}_j$ is the unique linear operator satisfying,

$$\phi_j \circ \boldsymbol{S}(\boldsymbol{x},\boldsymbol{y}) = \langle \boldsymbol{x}, \boldsymbol{S}_j \boldsymbol{y}\rangle\,, \tag{2}$$

for all $\boldsymbol{x} \in \mathbb{R}^m, \boldsymbol{y} \in \mathbb{R}^n$. Here, $\langle\cdot,\cdot\rangle$ denotes an inner product operation. Specific examples of bilinear maps are very common in signal processing applications. Some of them are listed below.

1. Linear and circular convolutions when both input vectors are unknown,

$$\boldsymbol{S}(\boldsymbol{x},\boldsymbol{y}) = \boldsymbol{x} \star \boldsymbol{y} \qquad (\text{Linear Convolution})\,, \tag{3}$$
$$\boldsymbol{S}(\boldsymbol{x},\boldsymbol{y}) = \boldsymbol{x} \circledast \boldsymbol{y} \qquad (\text{Circular Convolution})\,. \tag{4}$$

2. Linear representation in an unknown dictionary,

$$\boldsymbol{S}(\boldsymbol{\Phi},\boldsymbol{y}) = \boldsymbol{\Phi}\boldsymbol{y}. \tag{5}$$

3. Matrix product when both matrices are unknown,

$$\boldsymbol{S}(\boldsymbol{X},\boldsymbol{Y}) = \boldsymbol{X}\boldsymbol{Y}. \tag{6}$$

We observe that recovering $\boldsymbol{x}$ and $\boldsymbol{y}$ (or $\boldsymbol{X}$ and $\boldsymbol{Y}$) above is often trivial or ill-posed without further constraints. Our objective herein is to endeavor to characterize the types of constraints that render these problems well-posed.

2.2. Bilinear Inverse Problems

Given a bilinear map $\boldsymbol{S}\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^p$, we call the problem of determining the vector pair $(\boldsymbol{x},\boldsymbol{y})$ from the observation $\boldsymbol{S}(\boldsymbol{x},\boldsymbol{y})$ as a bilinear inverse problem,

$$\begin{aligned} &\text{find} && (\boldsymbol{x},\boldsymbol{y}) \\ &\text{subject to} && \boldsymbol{S}(\boldsymbol{x},\boldsymbol{y}) = \boldsymbol{z}. \end{aligned} \tag{P$_1$}$$

In general, Problem (P$_1$) is an ill-posed problem as a unique solution is not guaranteed to exist and the recovery is not robust to noisy observations. A popular approach to make such problems well-posed is to restrict the set of values that can be assumed by the pair $(\boldsymbol{x},\boldsymbol{y})$, denoted by $(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{K}$ for some set $\mathcal{K}$ (not necessarily convex). For example, $\mathcal{K}$ might represent the set of pairs of sparse vectors up to a certain sparsity. Sensible choices of $\mathcal{K}$ should depend on the underlying application (see [17] for an example on blind deconvolution for cooperative underwater acoustic communications). In Sections 3 and 4 we provide conditions, based on the null space of $\boldsymbol{S}$, that the constraint $(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{K}$ should encourage, so as to make Problem (P$_1$) well-posed.

Using (2), we can convert the bilinear constraint in Problem (P$_1$) into a set of p linear constraints as follows,

$$\begin{aligned} \phi_j \circ \boldsymbol{S}(\boldsymbol{x},\boldsymbol{y}) &= \langle \boldsymbol{x}, \boldsymbol{S}_j\boldsymbol{y}\rangle = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{S}_j\boldsymbol{y} \\ &= \operatorname{Tr}\Big(\boldsymbol{y}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{S}_j\Big) = \Big\langle \boldsymbol{x}\boldsymbol{y}^{\mathrm{T}}, \boldsymbol{S}_j\Big\rangle\,, \end{aligned} \tag{7}$$

for each $1 \le j \le p$. Identifying $\boldsymbol{x}\boldsymbol{y}^{\mathrm{T}}$ as an unknown matrix $\boldsymbol{W}$ with $\operatorname{rank}(\boldsymbol{W}) = 1$, Problem (P$_1$) can be rewritten as,

$$\begin{aligned} &\text{find} && \boldsymbol{W} \\ &\text{subject to} && \operatorname{rank}(\boldsymbol{W}) \le 1, \\ &&& \langle \boldsymbol{W}, \boldsymbol{S}_j\rangle = z_j, \quad \forall 1 \le j \le p. \end{aligned} \tag{P$_2$}$$

We have assumed that $\boldsymbol{z} \neq \boldsymbol{0}$ so that $\boldsymbol{x}\boldsymbol{y}^{\mathrm{T}} = \boldsymbol{0}$ is not a feasible solution to Problem (P$_1$). Under the assumption $\boldsymbol{z} \neq \boldsymbol{0}$, the constraints $\operatorname{rank}(\boldsymbol{W}) = 1$ and $\operatorname{rank}(\boldsymbol{W}) \le 1$ are equivalent. Finding $\boldsymbol{W}$ is equivalent to finding the vector pair $(\boldsymbol{x},\boldsymbol{y})$ up to multiplicative constants. Conversion of Problem (P$_1$) to Problem (P$_2$) provides several advantages,

1. Problem (P$_2$) has linear equality constraints versus the bilinear equality constraints of Problem (P$_1$) facilitating optimization. In particular, the tightest convex relaxation for the non-convex rank constraint in Problem (P$_2$) is well known [1].
2. The bilinear map is completely determined by the set of matrices $\boldsymbol{S}_j$ and is separated from the variable $\boldsymbol{W}$ in Problem (P$_2$). Thus, Problem (P$_2$) can be used to study general bilinear inverse problems.
3. For every bilinear inverse problem there is an inherent scaling ambiguity, $\forall\alpha \neq 0$, $\boldsymbol{S}(\boldsymbol{x},\boldsymbol{y}) = \boldsymbol{S}\big(\alpha\boldsymbol{x}, \frac{1}{\alpha}\boldsymbol{y}\big)$. Problem (P$_2$) is invariant to this ambiguity when $\boldsymbol{W} = \boldsymbol{x}\boldsymbol{y}^{\mathrm{T}}$ is the variable to be determined. Additional norm constraints on $\boldsymbol{x}$ or $\boldsymbol{y}$ can be used to recover $\boldsymbol{x}$ and $\boldsymbol{y}$ from $\boldsymbol{W}$, without affecting Problem (P$_2$).

2.3. Convex Relaxation

In the absence of any additional information about $\boldsymbol{x}$ and $\boldsymbol{y}$, the best convex relaxation of Problem (P$_2$) is given by [1],

$$\begin{aligned} &\underset{\boldsymbol{W}}{\text{minimize}} && \|\boldsymbol{W}\|_* \\ &\text{subject to} && \langle \boldsymbol{W}, \boldsymbol{S}_j\rangle = z_j \quad \forall 1 \le j \le p. \end{aligned} \tag{P$_3$}$$

Without loss of generality, we can assume that the mapping matrices $\boldsymbol{S}_j$ are orthonormal [1]. Problem (P$_3$) is the convex relaxation heuristic used for the low-rank matrix recovery problem in [18].

3. IDENTIFIABILITY RESULTS

We state conditions under which Problem (P$_2$) yields a unique solution, *i.e.* the input is identifiable. Let $\mathcal{S}\colon \mathbb{R}^{m\times n} \to \mathbb{R}^p$ denote the linear operator acting on the optimization variable $\boldsymbol{W}$ to generate the p linear constraints in Problem (P$_2$). So we have,

$$\mathcal{S}\Big(\boldsymbol{x}\boldsymbol{y}^{\mathrm{T}}\Big) = \boldsymbol{S}(\boldsymbol{x},\boldsymbol{y})\,, \quad \forall \boldsymbol{x} \in \mathbb{R}^m, \boldsymbol{y} \in \mathbb{R}^n. \tag{8}$$

We shall denote the set of all matrices of rank at most k in the null space of $\mathcal{S}$ by $\mathcal{N}(\mathcal{S},k)$. Thus,

$$\mathcal{N}(\mathcal{S},k) \triangleq \left\{\boldsymbol{X} \in \mathbb{R}^{m\times n} \,\middle|\, \operatorname{rank}(\boldsymbol{X}) \le k,\ \mathcal{S}(\boldsymbol{X}) = \boldsymbol{0}\right\}. \tag{9}$$

In particular, $\mathcal{N}(\mathcal{S},m)$ represents the null space of $\mathcal{S}$. For any matrix $\boldsymbol{M}$, we denote the row and column spaces by $\mathcal{R}(\boldsymbol{M})$ and $\mathcal{C}(\boldsymbol{M})$ and the j^{th} singular value in non-increasing order of magnitude by $\sigma_j(\boldsymbol{M})$.

We have the following necessary and sufficient condition for Problem (P$_2$) to succeed for *all* $\boldsymbol{z} = \boldsymbol{S}(\boldsymbol{x},\boldsymbol{y})$,

Theorem 1. *Problem* (P$_2$) *will find the correct solution for every observation* $\boldsymbol{z} = \boldsymbol{S}(\boldsymbol{x},\boldsymbol{y})$ *if and only if* $\mathcal{N}(\mathcal{S},2) = \{\boldsymbol{0}\}$.

[1] If not, then the mapping can always be made orthonormal with an accompanying change in $\boldsymbol{z}$.

Proof. Problem (P$_2$) fails if and only if more than one matrix satisfies the desired constraints.

Assume $\mathcal{N}(\mathcal{S},2)=\{\mathbf{0}\}$ and suppose that $\boldsymbol{W}_1$ and $\boldsymbol{W}_2$ denote two solutions to Problem (P$_2$). Then, $\mathcal{S}(\boldsymbol{W}_1)=\mathcal{S}(\boldsymbol{W}_2)$ so that $(\boldsymbol{W}_1-\boldsymbol{W}_2)\in\mathcal{N}(\mathcal{S},m)$. But, $\operatorname{rank}(\boldsymbol{W}_1-\boldsymbol{W}_2)\le\operatorname{rank}(\boldsymbol{W}_1)+\operatorname{rank}(\boldsymbol{W}_2)\le 2$ so that $\boldsymbol{W}_1-\boldsymbol{W}_2=\mathbf{0}$ and Problem (P$_2$) has a unique solution.

Conversely, let Problem (P$_2$) have a unique solution for every observation $\boldsymbol{z}=\boldsymbol{S}(\boldsymbol{x},\boldsymbol{y})$. First we argue that $\mathcal{N}(\mathcal{S},1)=\{\mathbf{0}\}$. Suppose that $\mathbf{0}\neq\boldsymbol{X}\in\mathcal{N}(\mathcal{S},1)$. Then we would have $\mathcal{S}(\boldsymbol{X})=\mathbf{0}=\mathcal{S}(\mathbf{0})$ so that the observation $\boldsymbol{z}=\mathbf{0}$ gives both $\boldsymbol{X}$ and $\mathbf{0}$ as valid solutions to Problem (P$_2$) leading to a contradiction. Thus, $\mathcal{N}(\mathcal{S},1)=\{\mathbf{0}\}$. Next, we argue that $\mathcal{N}(\mathcal{S},2)=\{\mathbf{0}\}$. Suppose that $\mathbf{0}\neq\boldsymbol{Y}\in\mathcal{N}(\mathcal{S},2)$ with the decomposition $\boldsymbol{Y}=\boldsymbol{Y}_1-\boldsymbol{Y}_2$ where $\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$ are rank-1 matrices. If we observe $\boldsymbol{z}=\mathcal{S}(\boldsymbol{Y}_1)$ then $\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$ are both valid solutions to Problem (P$_2$) owing to the equality $\mathcal{S}(\boldsymbol{Y}_1)=\mathcal{S}(\boldsymbol{Y}_2)$ thus leading to a contradiction. Hence, $\mathcal{N}(\mathcal{S},2)=\{\mathbf{0}\}$. $\square$

Unfortunately, for certain bilinear inverse problems of interest (like blind deconvolution), $\mathcal{N}(\mathcal{S},2)\neq\{\mathbf{0}\}$ (although we still have $\mathcal{N}(\mathcal{S},1)=\{\mathbf{0}\}$). Thus, we cannot hope to recover the correct solution for every observation $\boldsymbol{z}=\boldsymbol{S}(\boldsymbol{x},\boldsymbol{y})$. We shall endeavor to characterize which rank-1 matrices $\boldsymbol{W}$ result in observations $\boldsymbol{z}=\mathcal{S}(\boldsymbol{W})$ such that Problem (P$_2$) finds the correct solution, and when the correct solution can be found efficiently by solving Problem (P$_3$) instead of Problem (P$_2$).

First, we write Problem (P$_2$) in a different (but equivalent) form as below,

$$\begin{aligned}&\underset{\boldsymbol{W}}{\text{minimize}} && \operatorname{rank}(\boldsymbol{W})\\ &\text{subject to} && \mathcal{S}(\boldsymbol{W})=\boldsymbol{z}.\end{aligned}\tag{P$_4$}$$

Let $\boldsymbol{M}$ be a rank-1 solution to Problem (P$_2$) given the observation $\boldsymbol{z}=\boldsymbol{S}(\boldsymbol{x},\boldsymbol{y})$. It is clear that $\boldsymbol{M}$ is a solution of Problem (P$_4$) as well. All other solutions to Problem (P$_4$) must be of the form $(\boldsymbol{M}+\boldsymbol{X})$ where $\boldsymbol{X}$ is in the null space of the linear operator $\mathcal{S}$. If $\boldsymbol{M}+\boldsymbol{X}$ is another solution to Problem (P$_4$) then we have $\operatorname{rank}(\boldsymbol{M}+\boldsymbol{X})=\operatorname{rank}(\boldsymbol{M})=1$. Using the rank inequality $\operatorname{rank}(\boldsymbol{M}+\boldsymbol{X})\ge\operatorname{rank}(\boldsymbol{X})-\operatorname{rank}(\boldsymbol{M})$, we get $\operatorname{rank}(\boldsymbol{X})\le 2$. Hence, it is sufficient to consider $\boldsymbol{X}\in\mathcal{N}(\mathcal{S},2)$ and both Problems (P$_2$) and (P$_4$) will have a unique solution if and only if the following problem admits $\boldsymbol{X}=\mathbf{0}$ as the only solution,

$$\begin{aligned}&\underset{\boldsymbol{X}}{\text{minimize}} && \operatorname{rank}(\boldsymbol{M}+\boldsymbol{X})\\ &\text{subject to} && \boldsymbol{X}\in\mathcal{N}(\mathcal{S},2)\,.\end{aligned}\tag{P$_5$}$$

Theorem 2. *Let $\boldsymbol{M}=\sigma\boldsymbol{u}\boldsymbol{v}^T$ be a rank-1 matrix in $\mathbb{R}^{m\times n}$ such that for every $\boldsymbol{X}\in\mathcal{N}(\mathcal{S},2)$ either $\boldsymbol{u}\notin\mathcal{C}(\boldsymbol{X})$ or $\boldsymbol{v}^T\notin\mathcal{R}(\boldsymbol{X})$ is true. Given the observation $\boldsymbol{z}=\mathcal{S}(\boldsymbol{M})$, Problem* (P$_4$) *successfully recovers $\boldsymbol{M}$.*

Proof. Let $\boldsymbol{M}^*$ be a solution to Problem (P$_4$) such that $\boldsymbol{M}^*\neq\boldsymbol{M}$. Since $\boldsymbol{M}$ is a valid solution to Problem (P$_4$), we have $\operatorname{rank}(\boldsymbol{M}^*)=\operatorname{rank}(\boldsymbol{M})=1$ and $\boldsymbol{X}=\boldsymbol{M}-\boldsymbol{M}^*\in\mathcal{N}(\mathcal{S},2)$. If $\boldsymbol{M}^*=\sigma_*\boldsymbol{u}_*\boldsymbol{v}_*^{\mathrm{T}}$, then $\mathcal{R}(\boldsymbol{X})=\operatorname{Span}\big(\boldsymbol{v}^{\mathrm{T}},\boldsymbol{v}_*^{\mathrm{T}}\big)$ and $\mathcal{C}(\boldsymbol{X})=\operatorname{Span}(\boldsymbol{u},\boldsymbol{u}_*)$. This contradicts the assumption that at least one of $\boldsymbol{u}\notin\mathcal{C}(\boldsymbol{X})$ or $\boldsymbol{v}^{\mathrm{T}}\notin\mathcal{R}(\boldsymbol{X})$ is true. This completes the proof. $\square$

We consider an important special case when $\sigma_1(\boldsymbol{X})=\sigma_2(\boldsymbol{X})$ for all $\boldsymbol{X}\in\mathcal{N}(\mathcal{S},2)$. This special case is relevant to blind deconvolution. We have the following necessary and sufficient condition for the success of Problem (P$_4$). As the proof is similar to that of Theorem 2, we omit it here.

Corollary 3. *Let $\boldsymbol{M}=\sigma\boldsymbol{u}\boldsymbol{v}^T$ be a rank-1 matrix in $\mathbb{R}^{m\times n}$. Suppose that every matrix $\boldsymbol{X}\in\mathcal{N}(\mathcal{S},2)$ satisfying $\boldsymbol{u}\in\mathcal{C}(\boldsymbol{X})$ and $\boldsymbol{v}^T\in\mathcal{R}(\boldsymbol{X})$ admits a singular value decomposition with $\sigma_1(\boldsymbol{X})=\sigma_2(\boldsymbol{X})$. Let us denote such a decomposition as $\boldsymbol{X}=\sigma_*\boldsymbol{u}_1\boldsymbol{v}_1^T+\sigma_*\boldsymbol{u}_2\boldsymbol{v}_2^T$, and let $\boldsymbol{u}=\alpha_1\boldsymbol{u}_1+\alpha_2\boldsymbol{u}_2$ and $\boldsymbol{v}=\alpha_3\boldsymbol{v}_1+\alpha_4\boldsymbol{v}_2$ for some $\alpha_1,\alpha_2,\alpha_3,\alpha_4\in\mathbb{R}$ with $\alpha_1^2+\alpha_2^2=\alpha_3^2+\alpha_4^2=1$. Given the observation $\boldsymbol{z}=\mathcal{S}(\boldsymbol{M})$, Problem* (P$_4$) *successfully recovers $\boldsymbol{M}$ if and only if for every $\boldsymbol{X}\in\mathcal{N}(\mathcal{S},2)$ satisfying $\boldsymbol{u}\in\mathcal{C}(\boldsymbol{X})$ and $\boldsymbol{v}^T\in\mathcal{R}(\boldsymbol{X})$, $\alpha_1\alpha_3+\alpha_2\alpha_4\le 0$.*

4. RECOVERABILITY RESULTS

If we strengthen the sufficient conditions for identifiability assumed in Theorem 2 then we can obtain a similar result for recoverability using Problem (P$_3$). For the proof, we will need a result for norms of partitioned matrices from [19] (Proposition 3). We state it below for the special case of the nuclear norm of a 2×2 block partioned matrix.

Theorem 4 (adapted from [19])**.** *Let $\boldsymbol{X}$ be a 2×2 block partitioned matrix given by,*

$$\boldsymbol{X}=\begin{bmatrix}\boldsymbol{X}_{11} & \boldsymbol{X}_{12}\\ \boldsymbol{X}_{21} & \boldsymbol{X}_{22}\end{bmatrix}.\tag{10}$$

Then, the following inequality is true,

$$\|\boldsymbol{X}\|_*^2\ge\|\boldsymbol{X}_{11}\|_*^2+\|\boldsymbol{X}_{12}\|_*^2+\|\boldsymbol{X}_{21}\|_*^2+\|\boldsymbol{X}_{22}\|_*^2\,.\tag{11}$$

Theorem 5. *Let $\boldsymbol{M}$ be a rank-1 matrix in $\mathbb{R}^{m\times n}$ such that for every $\boldsymbol{X}$ in the null space of $\mathcal{S}$ at least one of $\boldsymbol{M}\boldsymbol{X}^T=\mathbf{0}$ or $\boldsymbol{M}^T\boldsymbol{X}=\mathbf{0}$ is true. Given the observation $\boldsymbol{z}=\mathcal{S}(\boldsymbol{M})$, Problem* (P$_3$) *successfully recovers $\boldsymbol{M}$.*

Proof. Problem (P$_3$) will recover $\boldsymbol{M}$ from the observation $\boldsymbol{z}=\mathcal{S}(\boldsymbol{M})$ if for all $\boldsymbol{X}\in\mathcal{N}(\mathcal{S},m)\setminus\{\mathbf{0}\}$, the strict inequality $\|\boldsymbol{M}+\boldsymbol{X}\|_*>\|\boldsymbol{M}\|_*$ is satisfied. We define some notation. Let $\boldsymbol{M}=\sigma\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$. Let the columns of $\boldsymbol{U}_\perp$ (respectively $\boldsymbol{V}_\perp$) represent an orthonormal basis for the orthogonal complement of $\boldsymbol{u}$ (respectively $\boldsymbol{v}$). Let $\boldsymbol{P}_{\boldsymbol{u}}=\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}$ and $\boldsymbol{P}_{\boldsymbol{U}_\perp}=\boldsymbol{U}_\perp\boldsymbol{U}_\perp^{\mathrm{T}}$ respectively denote the orthogonal projection matrices onto the spaces spanned by columns of $\boldsymbol{u}$ and $\boldsymbol{U}_\perp$. Let $\boldsymbol{P}_{\boldsymbol{v}}=\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}$ and $\boldsymbol{P}_{\boldsymbol{V}_\perp}=\boldsymbol{V}_\perp\boldsymbol{V}_\perp^{\mathrm{T}}$ respectively denote the orthogonal projection matrices onto the spaces spanned by rows of $\boldsymbol{v}^{\mathrm{T}}$ and $\boldsymbol{V}_\perp^{\mathrm{T}}$. For any matrix $\boldsymbol{Y}\in\mathbb{R}^{m\times n}$ we have the identity,

$$\boldsymbol{Y}=\boldsymbol{P}_{\boldsymbol{u}}\boldsymbol{Y}\boldsymbol{P}_{\boldsymbol{v}}+\boldsymbol{P}_{\boldsymbol{u}}\boldsymbol{Y}\boldsymbol{P}_{\boldsymbol{V}_\perp}+\boldsymbol{P}_{\boldsymbol{U}_\perp}\boldsymbol{Y}\boldsymbol{P}_{\boldsymbol{v}}+\boldsymbol{P}_{\boldsymbol{U}_\perp}\boldsymbol{Y}\boldsymbol{P}_{\boldsymbol{V}_\perp}.\tag{12}$$

Let $\boldsymbol{X}\in\mathcal{N}(\mathcal{S},m)\setminus\{\mathbf{0}\}$ and assume that $\boldsymbol{M}^{\mathrm{T}}\boldsymbol{X}=\mathbf{0}$. Using (12), we have

$$\boldsymbol{X}=\boldsymbol{P}_{\boldsymbol{U}_\perp}\boldsymbol{X}\boldsymbol{P}_{\boldsymbol{v}}+\boldsymbol{P}_{\boldsymbol{U}_\perp}\boldsymbol{X}\boldsymbol{P}_{\boldsymbol{V}_\perp}.\tag{13}$$

Because the nuclear norm is unitarily invariant, the nuclear norm of $(\boldsymbol{M}+\boldsymbol{X})$ does not change if $(\boldsymbol{M}+\boldsymbol{X})$ is expressed in a different basis. Using $(\boldsymbol{u},\boldsymbol{U}_\perp)$ and $\big(\boldsymbol{v}^{\mathrm{T}},\boldsymbol{V}_\perp^{\mathrm{T}}\big)$ as the column and row bases respectively for $(\boldsymbol{M}+\boldsymbol{X})$, we have

$$\begin{aligned}\|\boldsymbol{M}+\boldsymbol{X}\|_* &\overset{(a)}{=}\left\|\begin{bmatrix}\sigma & \mathbf{0}^{\mathrm{T}}\\ \boldsymbol{U}_\perp^{\mathrm{T}}\boldsymbol{X}\boldsymbol{v} & \boldsymbol{U}_\perp^{\mathrm{T}}\boldsymbol{X}\boldsymbol{V}_\perp\end{bmatrix}\right\|_*\\ &\overset{(b)}{\ge}\sqrt{\sigma^2+\|\boldsymbol{U}_\perp^{\mathrm{T}}\boldsymbol{X}\boldsymbol{v}\|_2^2+\|\boldsymbol{U}_\perp^{\mathrm{T}}\boldsymbol{X}\boldsymbol{V}_\perp\|_*^2}\\ &\overset{(c)}{>}\sigma=\|\boldsymbol{M}\|_*\,.\end{aligned}\tag{14}$$

where (a) holds due to change in the representation basis for $(\boldsymbol{M}+\boldsymbol{X})$, (b) holds due to Theorem 4, and (c) is due to the fact that if both $\boldsymbol{U}_\perp^{\mathrm{T}}\boldsymbol{X}\boldsymbol{v}$ and $\boldsymbol{U}_\perp^{\mathrm{T}}\boldsymbol{X}\boldsymbol{V}_\perp$ are zero then $\boldsymbol{X}$ has to be the zero matrix (from (13)), an impossibility as $\boldsymbol{X}\in\mathcal{N}(\mathcal{S},m)\setminus\{\mathbf{0}\}$. $\square$

A somewhat weaker sufficient condition can be derived if we consider the subdifferential of the nuclear norm.

Theorem 6. *Let $\boldsymbol{M}$ be a rank-1 matrix in $\mathbb{R}^{m\times n}$. Given the observation $\boldsymbol{z} = \mathcal{S}(\boldsymbol{M})$, Problem* (P$_3$) *successfully recovers $\boldsymbol{M}$ if $\left\|\boldsymbol{U}_\perp^T \boldsymbol{X} \boldsymbol{V}_\perp\right\|_* > \left|\boldsymbol{u}^T \boldsymbol{X} \boldsymbol{v}\right|$, for all $\boldsymbol{X} \in \mathcal{N}(\mathcal{S}, m) \setminus \{\boldsymbol{0}\}$.*

Proof. The proof is quite straightforward and uses the concept of subgradients from convex optimization. It is well known [20] that the subdifferential $\partial \left\|\boldsymbol{M}\right\|_*$ is given by,

$$\partial \left\|\boldsymbol{M}\right\|_* = \left\{\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}} + \boldsymbol{Z} \;\middle|\; \|\boldsymbol{Z}\| \le 1,\, \mathcal{C}(\boldsymbol{Z}) \subseteq \mathcal{C}(\boldsymbol{U}_\perp)\,, \right. \\ \left. \mathcal{R}(\boldsymbol{Z}) \subseteq \mathcal{R}\left(\boldsymbol{V}_\perp^{\mathrm{T}}\right)\right\} \tag{15}$$

and we have $\forall \boldsymbol{G} \in \partial \left\|\boldsymbol{M}\right\|_*$ and $\forall \boldsymbol{X} \in \mathbb{R}^{m\times n}$,

$$\left\|\boldsymbol{M} + \boldsymbol{X}\right\|_* - \left\|\boldsymbol{M}\right\|_* \ge \langle \boldsymbol{G}, \boldsymbol{X}\rangle \tag{16}$$

Problem (P$_3$) recovers $\boldsymbol{M}$ if and only if $\boldsymbol{M}$ is the unique minimizer of Problem (P$_3$). For all $\boldsymbol{X} \in \mathcal{N}(\mathcal{S}, m) \setminus \{\boldsymbol{0}\}$, we have

$$\begin{aligned} 0 &< \left\|\boldsymbol{U}_\perp^{\mathrm{T}} \boldsymbol{X} \boldsymbol{V}_\perp\right\|_* - \left|\boldsymbol{u}^{\mathrm{T}} \boldsymbol{X} \boldsymbol{v}\right| \\ &\overset{(a)}{=} \sup_{\substack{\|\boldsymbol{Z}\|\le 1 \\ \mathcal{C}(\boldsymbol{Z})\subseteq\mathcal{C}(\boldsymbol{U}_\perp) \\ \mathcal{R}(\boldsymbol{Z})\subseteq\mathcal{R}\left(\boldsymbol{V}_\perp^{\mathrm{T}}\right)}} \langle \boldsymbol{Z}, \boldsymbol{X}\rangle - \left|\boldsymbol{u}^{\mathrm{T}} \boldsymbol{X} \boldsymbol{v}\right| \\ &\le \sup_{\substack{\|\boldsymbol{Z}\|\le 1 \\ \mathcal{C}(\boldsymbol{Z})\subseteq\mathcal{C}(\boldsymbol{U}_\perp) \\ \mathcal{R}(\boldsymbol{Z})\subseteq\mathcal{R}\left(\boldsymbol{V}_\perp^{\mathrm{T}}\right)}} \langle \boldsymbol{Z}, \boldsymbol{X}\rangle + \left\langle \boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}, \boldsymbol{X}\right\rangle \\ &= \sup_{\substack{\|\boldsymbol{Z}\|\le 1 \\ \mathcal{C}(\boldsymbol{Z})\subseteq\mathcal{C}(\boldsymbol{U}_\perp) \\ \mathcal{R}(\boldsymbol{Z})\subseteq\mathcal{R}\left(\boldsymbol{V}_\perp^{\mathrm{T}}\right)}} \left\langle \boldsymbol{u}\boldsymbol{v}^{\mathrm{T}} + \boldsymbol{Z}, \boldsymbol{X}\right\rangle \\ &\overset{(b)}{=} \langle \boldsymbol{G}, \boldsymbol{X}\rangle \overset{(c)}{\le} \left\|\boldsymbol{M} + \boldsymbol{X}\right\|_* - \left\|\boldsymbol{M}\right\|_* \end{aligned} \tag{17}$$

for some $\boldsymbol{G} \in \partial \left\|\boldsymbol{M}\right\|_*$. In (17), (a) is due to the spectral norm being dual to the nuclear norm, (b) is due to the fact that the subdifferential for a convex function is a nonempty compact set and the supremum over a compact set is achieved at some point in that set, and (c) holds from (16). □

5. A NUMERICAL EXAMPLE

We consider a constrained blind linear deconvolution example, motivated by cooperative communications in sparse channels [21], *i.e.* $\boldsymbol{z} = \boldsymbol{x} \star \boldsymbol{y}$, where $\boldsymbol{x}$ and $\boldsymbol{y}$ are both sparse vectors. While we have not explicitly treated sparsity in the theoretical results, we shall see that sparsity in the appropriate domain is more likely to satisfy our needed null space conditions. It is NP-hard to check whether a given realization of $(\boldsymbol{x}, \boldsymbol{y})$ satisfies the required null space conditions, so we shall restrict ourselves to results averaged over independent realizations of the pair $(\boldsymbol{x}, \boldsymbol{y})$. For this case, the matrices $\boldsymbol{S}_k$ are given by,

$$\left(\boldsymbol{S}_k\right)_{ij} = \begin{cases} 1, & i + j = k + 1 \\ 0, & \text{otherwise} \end{cases} \tag{18}$$

with $p = m + n - 1$. Sparse signals are generated and their linear convolution is fed to the recovery Problem (P$_3$) (solved using CVX). We consider sparsity in two different bases, *viz.* the standard Euclidean and Hadamard bases. As the solution to Problem (P$_3$) is not, in general, guaranteed to be a rank-1 matrix, we use the best rank-1 approximation of this matrix, and study the error performance versus sparsity. In Figure 1, we have density denoting the fraction of representation coefficients being non-zero and $m = n = 8$, implying $p = 15$. For each sparsity level, in either representation, the support set is chosen via uniform random sampling without replacement and the nonzero representation coefficients are chosen independently from the distribution of $|Y|$ where Y has a standard normal distribution. We study the Average Relative Approximation Error (ARAE) metric given by,

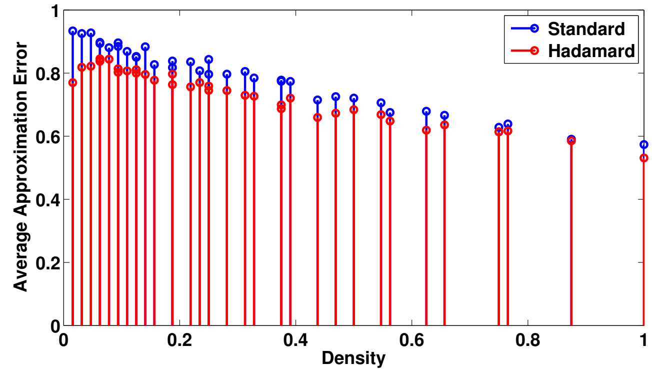

Fig. 1. Average approximation error vs. density for blind deconvolution. The sparsity level is normalized w.r.t. mn.

$$\text{ARAE} = \left\langle \frac{\left\|\boldsymbol{M} - \widehat{\boldsymbol{M}}\right\|_{\mathrm{F}}}{\left\|\boldsymbol{M}\right\|_{\mathrm{F}}} \right\rangle \tag{19}$$

where $\boldsymbol{M}$ is the true solution, $\widehat{\boldsymbol{M}}$ is the best rank-1 estimate of $\boldsymbol{M}$ from solving Problem (P$_3$), $\|\cdot\|_{\mathrm{F}}$ denotes Frobenius norm, and $\langle\cdot\rangle$ denotes averaging over 100 independent realizations.

We make the following observations. The ARAE steadily decreases with increasing density for both representation bases. For the standard Euclidean basis, this trend has a simple interpretation in terms of the null space of the linear convolution operator. This null space has a basis comprised of sparse rank-2 matrices, and hence it is very likely that the true sparse rank-1 solution is not even identifiable in most instances. Sparsity in the Hadamard basis, on the other hand, ensures a dense representation in the standard Euclidean basis via the uncertainty principle. These matrices are quite likely to satisfy the sufficient conditions of Theorem 6. Hence, it is not surprising that the ARAE for the Hadamard basis is consistently lower than that for the standard Euclidean basis and this difference decreases as the normalized density increases, i.e. representation in the standard Euclidean basis becomes more dense.

6. CONCLUSIONS

In this paper, we have studied identifiability and recoverability for bilinear inverse problems from a rank-1 matrix recovery perspective. This treatment yields simple sufficient conditions for instance identifiability and recoverability based on the rank-2 null space of the linear matrix operator, when universal identifiability is impossible. Future work shall consider identifiability and recoverability for more specific constraints like sparsity, and approximately rank-1 matrix recovery problems for the cases of model mismatch and observation noise.

7. REFERENCES

[1] B. Recht, M. Fazel, and P. A. Parrilo, "Guaranteed Minimum-Rank Solutions of Linear Matrix Equations via Nuclear Norm Minimization," *SIAM Rev.*, vol. 52, no. 3, pp. 471–501, 2010.

[2] D. Kundur and D. Hatzinakos, "Blind Image Deconvolution," *IEEE Signal Process. Mag.*, vol. 13, no. 3, pp. 43–64, May 1996.

[3] C. R. Johnson, Jr., P. Schniter, T. J. Endres, J. D. Behm, D. R. Brown, and R. A. Casas, "Blind Equalization Using the Constant Modulus Criterion: A Review," *Proc. IEEE*, vol. 86, no. 10, pp. 1927–1950, Oct. 1998.

[4] M. Elad, M. A. T. Figueiredo, and Y. Ma, "On the Role of Sparse and Redundant Representations in Image Processing," *Proc. IEEE*, vol. 98, no. 6, pp. 972–982, Jun. 2010.

[5] K. Abed-Meraim, W. Qiu, and Y. Hua, "Blind System Identification," *Proc. IEEE*, vol. 85, no. 8, pp. 1310–1322, Aug. 1997.

[6] D. L. Donoho and V. Stodden, "When Does Non-Negative Matrix Factorization Give a Correct Decomposition into Parts?" in *NIPS*, 2003.

[7] K. Herrity, R. Raich, and A. O. Hero, III, "Blind Reconstruction of Sparse Images with Unknown Point Spread Function," C. A. Bouman, E. L. Miller, and I. Pollak, Eds., vol. 6814, no. 1. SPIE, 2008, p. 68140K.

[8] D. Barchiesi and M. D. Plumbley, "Dictionary Learning of Convolved Signals," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2011*, May 2011, pp. 5812–5815.

[9] M. Burger and O. Scherzer, "Regularization Methods for Blind Deconvolution and Blind Source Separation Problems," *Math. Control Signals Systems*, vol. 14, no. 4, pp. 358–383, 2001.

[10] M. S. Asif, W. Mantzel, and J. K. Romberg, "Random Channel Coding and Blind Deconvolution," in *47th Annual Allerton Conference on Communication, Control, and Computing, 2009. Allerton 2009.*, Oct. 2009, pp. 1021–1025.

[11] A. Ahmed, B. Recht, and J. Romberg, "Blind deconvolution using convex programming," *CoRR*, vol. abs/1211.5608, 2012.

[12] B. Recht, W. Xu, and B. Hassibi, "Necessary and Sufficient Conditions for Success of the Nuclear Norm Heuristic for Rank Minimization," in *47th IEEE Conference on Decision and Control, 2008. CDC 2008.*, Dec. 2008, pp. 3065–3070.

[13] ——, "Null space conditions and thresholds for rank minimization," *Math. Program.*, vol. 127, no. 1, Ser. B, pp. 175–202, 2011.

[14] C. Hegde and R. G. Baraniuk, "Sampling and Recovery of Pulse Streams," *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1505–1517, 2011.

[15] P. Walk and P. Jung, "Compressed Sensing on the Image of Bilinear Maps," in *IEEE International Symposium on Information Theory Proceedings (ISIT), 2012*, Jul. 2012, pp. 1291–1295.

[16] F. Kiraly and R. Tomioka, "A Combinatorial Algebraic Approach for the Identifiability of Low-Rank Matrix Completion," *ArXiv e-prints*, Jun. 2012.

[17] S. Choudhary and U. Mitra, "Sparse recovery from convolved output in underwater acoustic relay networks," in *2012 Asia-Pacific Signal Information Processing Association Annual Summit and Conference (APSIPA ASC)*, Dec. 2012, pp. 1–8.

[18] D. Gross, "Recovering Low-Rank Matrices From Few Coefficients in Any Basis," *IEEE Trans. Inf. Theory*, vol. 57, no. 3, pp. 1548–1566, 2011.

[19] R. Bhatia and F. Kittaneh, "Norm inequalities for partitioned operators and an application," *Math. Ann.*, vol. 287, no. 4, pp. 719–726, 1990.

[20] G. A. Watson, "Characterization of the subdifferential of some matrix norms," *Linear Algebra Appl.*, vol. 170, pp. 33–45, 1992.

[21] N. Richard and U. Mitra, "Sparse Channel Estimation for Cooperative Underwater Communications: A Structured Multichannel Approach," in *IEEE International Conference on Acoustics, Speech and Signal Processing, 2008.*, 2008, pp. 5300–5303.