# On the Implementation of Semi-Blind Block-Term Decomposition-based Receivers for Uniform Rectangular Arrays

Eleftherios Kofidis
Dept of Statistics and Insurance Science
University of Piraeus
Piraeus 185 34, Greece
Email: kofidis@unipi.gr

***Abstract*—Using the commonly employed uniform rectangular array as an illustrating example, we have recently shown that the block-term tensor decomposition model provides a natural representation for the output signal of a multi-dimensional sensor array structure in the presence of channel multipath. This allowed us to develop effective and efficient semi-blind joint channel estimation/data detection algorithms, and, more recently, we have studied the role of the discrete nature of the symbol constellation in this context. In this paper, we build on that work by endowing receivers with the ability to estimate the number of impinging signals, previously assumed to be *a-priori* known. A palette of alternatives in model selection and detection performance, as well as in computational requirements, is considered, which includes schemes based on expectation maximization with different choices of the symbol prior. The schemes considered are evaluated through simulations at various noise levels and different constellation sizes and are favorably compared with the training-only-based solution.**

## I. Introduction

Semi-blind *joint* channel estimation/data detection (JCD) (see, e.g., [1]) has regained interest in the array signal processing (ASP) community as a means of meeting the requirement for resource-efficient receiver realization, which is particularly relevant in massively large multi-user multiple-input multiple-output (MIMO) systems. Tensor models and methods [2] provide a suitable solution approach to this problem, given the inherently multidimensional (m-D) nature of such systems and the ability of tensor decomposition (TD)-based methods to recover latent information from the received (Rx) signals in a sample-efficient manner and with little or no training overhead. Block-term decomposition (BTD) [3] has seen relatively fewer applications in ASP compared to other TD models, notably the canonical polyadic decomposition (CPD) and, to a smaller extent, the Tucker decomposition (TKD) (see [4] for a brief overview of the related literature). In BTD, the tensor is decomposed into a sum of low multilinear rank terms (referred to as blocks), and this can be seen as an intermediate between CPD (sum of rank-1 terms) and TKD (only one term of low multilinear rank). This gives BTD its flexibility and justifies its wide applicability [5]. Its most common variant comprises rank-$(L, L, 1)$ block terms and is hence referred to as the LL1 model [3].

This work has been partly supported by the University of Piraeus Research Center.

It was recently shown [4] that an LL1 representation *naturally* results when the (base station) receiver is equipped with a uniform rectangular array (URA) and a number, $R$, of user signals arrive through *multiple* paths, say $L_r$, $r = 1, 2, \ldots, R$. Iterative algorithms that wed existing JCD schemes with BTD approximation were considered and thoroughly evaluated. That scheme singled out for its simplicity, estimation/detection performance, and computational and data efficiency combines a classical ASP algorithm, namely iterative least squares with projection (ILSP) [6], with what is called in [4] singular value projection (SVP) and consists of projecting the channel matrix estimate onto the set of matrices with low-rank columns to respect the LL1 model. For simplicity, the vector symbol estimates in ILSP are projected onto the input *discrete* constellation in an *entry-wise* manner, which renders the algorithm suboptimal and not provenly monotonically convergent (although such difficulties are rarely encountered in practice; see [4] for details). Our work in [4] was revisited in [7] from a Bayesian standpoint to shed light on the role of the constellation's discrete nature in such algorithms. ILSP was given a new interpretation as a heuristic version of an expectation-maximization (EM) procedure [8], [9] that places a Gaussian prior on the symbols. Per-user minimum mean squared error (MMSE) symbol estimation, aided by a minimum power distortionless response (MPDR) filter [10], [11], was also considered and shown to have the best detection performance at *all* signal-to-noise ratio (SNR) levels.

Both $R$ and the $L_r$s were, however, assumed to be *a-priori* known, an assumption that is generally not valid in practice. Algebraic methods for LL1 model selection and computation (e.g., [12], [13]) tend to be computationally expensive (starting, in the present context of an $N_x \times N_y$ URA with inputs of length $N_s$, with the computation of the null space of an $N_x N_y N_s \times (N_x^2 + N_y^2)$ matrix) and only work safely at high enough signal-to-noise ratio (SNR) levels, which may not be the case in practical communication systems [13]. Rank-revealing optimization looks more promising in this respect, and a preliminary study of such methods in the present context is the theme of this paper. A palette of alternatives in model selection and detection performance, as well as computational cost, is considered. This includes the ILSP-SVP scheme, assisted by a singular value decomposition (SVD)-based rank estimation heuristic, and two EM-based algorithms with rank-revealing channel prior and two choices of the data prior, differing in their proximity to the known discrete constellation. The schemes considered are evaluated in rank recovery and detection performance and compared in their computational requirements through simulations at various noise levels and different constellation sizes.

### A. Notation

The superscript $^*$ stands for complex conjugation. $\circ$ and $\odot, \oslash$ denote the outer product and entry-wise multiplication and division, respectively. $|\mathbf{A}|^2$ means $\mathbf{A} \odot \mathbf{A}^*$. $\mathrm{Diag}(\mathbf{x})$ is the diagonal matrix with the vector $\mathbf{x}$ on its main diagonal, while $\mathrm{diag}(\mathbf{A})$ yields the main diagonal of $\mathbf{A}$ as a (column) vector. The trace operator is denoted by $\mathrm{tr}(\cdot)$. $\Re\{\cdot\}$ takes the real part of a complex-valued quantity.

## II. System Model and Problem Statement

Let $R$ far-field sources emit uncorrelated narrowband signals, which impinge on a uniform xy planar antenna array of dimensions $N_x \times N_y$. The $r$th signal arrives through $L_r$ paths that are assumed to be sufficiently closely spaced compared to the duration of the symbol so that they all have the same delay. If the impinging signals, $\mathbf{s}_r \in \mathbb{C}^{N_\mathrm{s}\times 1}$, $r = 1, 2, \ldots, R$, are of length $N_\mathrm{s}$, then the array output can be represented by an $N_x \times N_y \times N_\mathrm{s}$ tensor, expressed as [4]

$$\boldsymbol{\mathcal{Y}} = \sum_{r=1}^{R} \mathbf{H}_r \circ \mathbf{s}_r + \boldsymbol{\mathcal{W}}, \qquad (1)$$

where $\boldsymbol{\mathcal{W}}$ stands for the zero-mean Gaussian noise, assumed spatially and temporally white with variance $\sigma_w^2$. The $r$th channel coefficient can be written as $\mathbf{H}_r = \mathbf{A}_{x,r}\mathbf{D}_r\mathbf{A}_{y,r}^{\mathrm{T}}$, where $\mathbf{A}_{x,r}$ and $\mathbf{A}_{y,r}$ are the corresponding $N_x \times L_r$ and $N_y \times L_r$ steering matrices, respectively, and the path gains are on the main diagonal of the diagonal matrix $\mathbf{D}_r$. It is clear that, especially for large-scale arrays, the steering matrices are generically full column rank, $L_r$, which implies that (1) is a noisy decomposition into rank-$(L_r, L_r, 1)$ terms. Collect all $\mathbf{A}_{x,r}$'s in $\mathbf{A}_x \in \mathbb{C}^{N_x\times \bar{L}}$, with $\bar{L} \triangleq \sum_{r=1}^{R} L_r$, and similarly for $\mathbf{A}_{y,r}$'s. Let $\mathbf{S} = \begin{bmatrix} \mathbf{s}_1 & \cdots & \mathbf{s}_R \end{bmatrix} \in \mathbb{C}^{N_s\times R}$ be the transmitted (Tx) symbol matrix. Then the following factorization for the transposed mode-3 unfolding holds [4]

$$\mathbf{Y} \triangleq \mathbf{Y}_{(3)}^{\mathrm{T}} = \mathbf{H}\mathbf{S}^{\mathrm{T}} + \mathbf{W}_{(3)}^{\mathrm{T}}, \qquad (2)$$

where $\mathbf{H} \in \mathbb{C}^{N_xN_y\times R}$ represents the combined channel matrix, with columns $\mathrm{vec}(\mathbf{H}_r)$, $r = 1, 2, \ldots, R$. With all $L_r = 1$, i.e., only line-of-sight transmission, the above reduces to a rank-$R$ CPD approximation problem.

With known $R$ and $L_r$s, the discrete-valued nature of $\mathbf{S}$ and the low-rankness of the columns of $\mathbf{H}$ can be exploited to solve (2) for the Tx symbols and the channel [4], [7]. $\mathbf{S}$ should belong to the set $\mathcal{F}$ of $N_s \times R$ matrices with entries in the symbol constellation employed, say $\mathcal{A} = \{a_1, a_2, \ldots, a_M\}$ with order $M = |\mathcal{A}|$, while the $r$th column of $\mathbf{H}$, when unfolded into an $N_x \times N_y$ matrix, should be of rank (at most) $L_r$, $r = 1, 2, \ldots, R$. Let $\mathcal{L}$ denote the set of $N_xN_y \times R$ matrices that enjoy the latter property. Then, the above JCD task can be stated as the following maximum likelihood (ML) optimization problem

$$\min_{\mathbf{H}\in\mathcal{L},\mathbf{S}\in\mathcal{F}} h(\mathbf{H},\mathbf{S}) \triangleq \frac{1}{2}\|\mathbf{Y}-\mathbf{H}\mathbf{S}^{\mathrm{T}}\|_{\mathrm{F}}^2 + \frac{\eta}{2}(\|\mathbf{H}\|_{\mathrm{F}}^2+\|\mathbf{S}\|_{\mathrm{F}}^2), \qquad (3)$$

where a ridge regularizer, with a regularization parameter $\eta > 0$, has been included for ensuring stability. Alternatively, an EM approach can be taken, as detailed in [7], with an appropriate prior chosen for the hidden variable $\mathbf{S}$. The above solution approaches are revisited in this paper in the realistic scenario of *a-priori* unknown $R$ and $L_r$s. Given the relative robustness of the approximated LL1 model to an overestimation of the $L_r$s, observed in several application contexts (see, e.g., [14] and references therein), here only the number, $R$, of input signals is estimated, with the $L_r$s being overestimated at the same value, $L_{\mathrm{ini}}$.

## III. Rank-revealing BTD-based JCD

The factors in (3) can be identified by alternatingly solving for $\mathbf{S}$ and $\mathbf{H}$ while projecting $\mathbf{S}$ onto $\mathcal{F}$ and $\mathbf{H}$ onto $\mathcal{L}$ in each iteration [4]. The former step is generally non-trivial and can be performed in an exhaustive search manner [6], [8] at a cost exponential in the number of sources. Instead, relying simply on entry-wise projection was demonstrated in [4] as a computationally efficient yet well-performing option. Regarding the projection onto $\mathcal{L}$, this can be done via an SVP step[1], namely replacing each column of $\mathbf{H}$ by the vectorization of the rank-$L_{\mathrm{ini}}$ SVD of its $N_x \times N_y$ matricization. Before that, $R$ can be estimated from the (assumed in non-increasing order) singular values[2] of $\mathbf{Y}$ as follows[3] [17, Section 3.2]

$$\hat{R} = \arg\min_{i=1,2,\ldots,N_xN_y} \frac{\sigma_{i+1}}{\sigma_i} \qquad (4)$$

The resulting algorithm is tabulated here as Algorithm 1.

---

**Algorithm 1:** ILSP-SVP

**Data:** $\mathbf{Y}, L_{\mathrm{ini}}, \eta$
**Result:** Estimates of $R$, $\mathbf{H} \in \mathbb{C}^{N_xN_y\times R}$, $\mathbf{S} \in \mathbb{C}^{N_s\times R}$

1 Estimate $R$ from (4);
2 Initialize $\mathbf{H} \in \mathbb{C}^{N_xN_y\times R}$;
3 **repeat**
4 $\quad \mathbf{S} \leftarrow \mathbf{Y}^{\mathrm{T}}\mathbf{H}^*(\mathbf{H}^{\mathrm{T}}\mathbf{H}^* + \eta\mathbf{I}_R)^{-1}$;
5 $\quad$ Project $\mathbf{S}$ entry-wise onto $\mathcal{F}$;
6 $\quad \mathbf{H} \leftarrow \mathbf{Y}\mathbf{S}^*(\mathbf{S}^{\mathrm{T}}\mathbf{S}^* + \eta\mathbf{I}_R)^{-1}$;
7 $\quad$ Project $\mathbf{H}$ onto $\mathcal{L}$ via (rank-$L_{\mathrm{ini}}$) SVP per column;
8 **until** *convergence*;

---

To mitigate the exponential complexity of the $\mathcal{F}$-constrained ML problem, [8] opted to model Tx symbols in the general

[1]This is somewhat abusing the use of the acronym SVP, which is, strictly speaking, used in the literature to refer to projected (onto $\mathcal{L}$) gradient descent iterations aiming at a low-rank matrix that satisfies affine constraints (as in, e.g., matrix completion) [15].

[2]Alternatively, from its QR decomposition, as in, e.g., [16].

[3]The cost function is given by $\frac{\sigma_{i+1}+\sigma_1\sqrt{i/\epsilon}}{\sigma_i}$ in [17], for a large value of $\epsilon$, which when set to infinity yields (4).

MIMO system setting as zero-mean Gaussian random variables and resort to EM for developing a semi-blind receiver of affordable complexity and good performance at moderate to low SNR, high-order constellations or large-scale systems. Performance improvements were observed in [9] when *entry-wise* projections onto $\mathcal{F}$, called "heuristic de-mapping" therein, were intertwined with the iterations. Interestingly, this yields ILSP when replacing $\sigma_w^2$ by $\eta$ and assuming that all symbols are also *a-posteriori* uncorrelated of the same power. [8] also proposed a version of its EM-based channel estimator that uses a prior for the channel coefficients as well. It is not difficult to see that this amounts to the well-known Bayesian principal component analysis (BPCA) method [18], originally developed as a means of performing probabilistic PCA (PPCA) that also has the ability of dimensionality selection. In the present context, the channel prior takes the form

$$p(\mathbf{H}|\mathbf{\Gamma_H}) = \prod_{r=1}^{R_{\text{ini}}} \mathcal{CN}(\mathbf{H}(:,r)|\mathbf{0}, \gamma_{\mathbf{H}}(r)^{-1}\mathbf{I}_{N_x N_y}), \quad (5)$$

where $\mathbf{\Gamma_H} = \text{Diag}(\{\gamma_{\mathbf{H}}(r)\}_{r=1}^{R_{\text{ini}}})$ and the channel column precisions $\gamma_{\mathbf{H}}$ are to be learned from the data. Those among them with large (ideally infinite) values will then correspond to columns that should be dropped, allowing an estimate of $R$ to be obtained. Furthermore, incorporating an SVP step at each iteration results in one more candidate scheme for our JCD task, named here *BPCA-SVP* and summarized in Algorithm 2. The diagonal matrix $\mathbf{\Gamma_S}$ has on its diagonal the inverse powers

**Algorithm 2:** BPCA-SVP

**Data:** $\mathbf{Y}, R_{\text{ini}}, L_{\text{ini}}, \mathbf{\Gamma_S} \in \mathbb{R}^{R_{\text{ini}} \times R_{\text{ini}}}$
**Result:** Estimates of $R$, $\mathbf{H} \in \mathbb{C}^{N_x N_y \times R}$, $\mathbf{S} \in \mathbb{C}^{N_s \times R}$
1 Initialize $\mathbf{H} \in \mathbb{C}^{N_x N_y \times R_{\text{ini}}}$, $\mathbf{\Gamma_H} \in \mathbb{R}^{R_{\text{ini}} \times R_{\text{ini}}}$, and $\sigma_w^2$;
2 **repeat**
3 $\quad \mathbf{\Sigma} \leftarrow (\mathbf{H}^{\text{T}}\mathbf{H}^* + \sigma_w^2 \mathbf{\Gamma_S})^{-1}$;
4 $\quad \mathbf{S} \leftarrow \mathbf{Y}^{\text{T}}\mathbf{H}^*\mathbf{\Sigma}$;
5 $\quad \mathbf{\Sigma_S} \leftarrow \sigma_w^2 \mathbf{\Sigma}$;
6 $\quad \langle\mathbf{SS}\rangle \leftarrow \mathbf{S}^{\text{T}}\mathbf{S}^* + N_s \mathbf{\Sigma_S}$;
7 $\quad \mathbf{\Sigma_H} \leftarrow (\langle\mathbf{SS}\rangle + \sigma_w^2 \mathbf{\Gamma_H})^{-1}$;
8 $\quad \mathbf{H} \leftarrow \mathbf{Y}\mathbf{S}^*\mathbf{\Sigma_H}$;
9 $\quad \mathbf{\Gamma_H} \leftarrow \text{Diag}((\mathbf{1}_{R_{\text{ini}}} N_x N_y) \oslash (|\mathbf{H}^{\text{T}}|^2 \mathbf{1}_{N_x N_y}))$;
10 $\quad$ Project $\mathbf{H}$ onto $\mathcal{L}$ via (rank-$L_{\text{ini}}$) SVP per column; $\hat{\mathbf{X}} \leftarrow \mathbf{H}\mathbf{S}^{\text{T}}$;
11 $\quad \sigma_w^2 \leftarrow \frac{\|\mathbf{Y}\|_{\text{F}}^2 - 2\Re\{\text{tr}(\mathbf{Y}^{\text{H}}\hat{\mathbf{X}})\} + \text{tr}(\mathbf{H}^{\text{H}}\mathbf{H}\langle\mathbf{SS}\rangle)}{N_x N_y N_s}$;
12 **until** *convergence*;
13 Use $\mathbf{\Gamma_H}$ to drop redundant columns of $\mathbf{H}$ and corresponding columns of $\mathbf{S}$;
14 Project $\mathbf{S}$ entry-wise onto $\mathcal{F}$;

of the $\mathrm{s}_r$s, and can in practice be set to $\frac{1}{\text{var}(\mathcal{A})}\mathbf{I}_{R_{\text{ini}}}$. $\langle\mathbf{SS}\rangle$ contains the posterior 2nd-order moments of the symbols. A first estimate of the noise power can be computed from the training data and subsequently updated as in Line 11. Note that *soft* symbol estimates are used throughout, and *hard* estimates are only computed at the end.

To come up with a semi-blind MIMO receiver that both takes into account the discrete constellation and is computationally tractable, [11] proposed the adoption of an MPDR filter [10] to first *unmix* the sources before performing MMSE estimation in their noisy versions. The MPDR principle from ASP [19] suggests filtering the Rx signal to minimize the power at the output of the filter while preserving the Tx signal of interest. The $r$th filter output, say $\mathbf{R}(:,r)$, can then be written as the sum of the corresponding Tx signal, $\mathbf{s}_r^{\text{T}}$, and noise plus interference, $\mathbf{n}_r^{\text{T}}$, which, for a large number of users, is well approximated by zero-mean Gaussian:

$$\mathbf{R}(:,r)^{\text{T}} = \frac{\mathbf{H}(:,r)^{\text{H}}\mathbf{\Sigma_Y}\mathbf{Y}}{\mathbf{H}(:,r)^{\text{H}}\mathbf{\Sigma_Y}\mathbf{H}(:,r)} = \mathbf{s}_r^{\text{T}} + \mathbf{n}_r^{\text{T}}, \quad (6)$$

where $\mathbf{\Sigma_Y} = (\mathbf{H}\mathbf{\Gamma}\mathbf{H}^{\text{H}} + \sigma_w^2 \mathbf{I}_{N_x N_y})^{-1}$ is the inverse of the Rx signal covariance and $\frac{1}{\mathbf{H}(:,r)^{\text{H}}\mathbf{\Sigma_Y}\mathbf{H}(:,r)}$ equals the power of $\mathbf{R}(:,r)$ [10], [11]. Hence, one can compute the conditional means of the Tx symbols *separately* for each signal and assume their known discrete distribution, which would otherwise lead to an intractable exponential complexity detection problem [8]. Including the channel prior from (5) and inserting the LL1-induced SVP step gives rise to an additional JCD scheme, presented as Algorithm 3 and henceforth called *MPDR-SVP*. $\boldsymbol{\rho}$ contains the probabilities of occurrence of the

**Algorithm 3:** MPDR-SVP

**Data:** $\mathbf{Y}, R_{\text{ini}}, L_{\text{ini}}, \boldsymbol{\rho}$
**Result:** Estimates of $R$, $\mathbf{H} \in \mathbb{C}^{N_x N_y \times R}$, $\mathbf{S} \in \mathbb{C}^{N_s \times R}$
1 Initialize $\mathbf{H} \in \mathbb{C}^{N_x N_y \times R_{\text{ini}}}$, $\mathbf{\Gamma_S} \in \mathbb{R}^{R_{\text{ini}} \times R_{\text{ini}}}$, $\mathbf{\Gamma_H} \in \mathbb{R}^{R_{\text{ini}} \times R_{\text{ini}}}$, and $\sigma_w^2$;
2 **repeat**
3 $\quad \mathbf{\Sigma_Y} \leftarrow (\mathbf{H}\mathbf{\Gamma_S}\mathbf{H}^{\text{H}} + \sigma_w^2 \mathbf{I}_{N_x N_y})^{-1}$;
4 $\quad \mathbf{H_Y} \leftarrow \mathbf{\Sigma_Y}\mathbf{H}$;
5 $\quad \mathbf{G} \leftarrow \mathbf{H}^{\text{H}}\mathbf{H_Y}$;
6 $\quad \mathbf{\Sigma_r} \leftarrow \text{Diag}(\mathbf{1}_{R_{\text{ini}}} \oslash \text{diag}(\mathbf{G}))$;
7 $\quad \mathbf{\Sigma_n} \leftarrow \mathbf{\Sigma_r} - \mathbf{\Gamma_S}$;
8 $\quad \mathbf{R} \leftarrow \mathbf{Y}^{\text{T}}\mathbf{H}_{\mathbf{Y}}^*\mathbf{\Sigma_r}$;
9 $\quad$ **for** $m = 1, 2, \ldots, M$ **do**
10 $\quad\quad \boldsymbol{\mathcal{Z}}(:,m,:) \leftarrow \exp\left(\left|\mathbf{R} - a_m \mathbf{1}_{N_s \times R_{\text{ini}}}\right|^2 \mathbf{\Sigma_n}^{-1}\right)$;
11 $\quad$ **end**
12 $\quad \mathbf{S} \leftarrow (\sum_{m=1}^{M} \rho_m a_m \boldsymbol{\mathcal{Z}}(:,m,:)) \oslash (\sum_{m=1}^{M} \rho_m \boldsymbol{\mathcal{Z}}(:,m,:))$;
13 $\quad \mathbf{S}_2 \leftarrow |\mathbf{S}|^2$;
14 $\quad \mathbf{\Sigma_s} \leftarrow (\sum_{m=1}^{M} \rho_m |a_m|^2 \boldsymbol{\mathcal{Z}}(:,m,:)) \oslash (\sum_{m=1}^{M} \rho_m \boldsymbol{\mathcal{Z}}(:,m,:)) - \mathbf{S}_2$;
15 $\quad \mathbf{\Gamma_S} \leftarrow \text{Diag}(\frac{1}{N_s}\mathbf{S}_2^{\text{T}}\mathbf{1}_{N_s})$;
16 $\quad \langle\mathbf{SS}\rangle \leftarrow \mathbf{S}^{\text{T}}\mathbf{S}^* + \text{Diag}(\frac{1}{N_s}\mathbf{\Sigma}_{\mathbf{s}}^{\text{T}}\mathbf{1}_{N_s})$;
17 $\quad \mathbf{H} \leftarrow \mathbf{Y}\mathbf{S}^*(\langle\mathbf{SS}\rangle + \sigma_w^2 \mathbf{\Gamma_H})^{-1}$;
18 $\quad \mathbf{\Gamma_H} \leftarrow \text{Diag}((\mathbf{1}_{R_{\text{ini}}} N_x N_y) \oslash (|\mathbf{H}^{\text{T}}|^2 \mathbf{1}_{N_x N_y}))$;
19 $\quad$ Project $\mathbf{H}$ onto $\mathcal{L}$ via (rank-$L_{\text{ini}}$) SVP per column; $\hat{\mathbf{X}} \leftarrow \mathbf{H}\mathbf{S}^{\text{T}}$;
20 $\quad \sigma_w^2 \leftarrow \frac{\|\mathbf{Y}\|_{\text{F}}^2 - 2\Re\{\text{tr}(\mathbf{Y}^{\text{H}}\hat{\mathbf{X}})\} + \text{tr}(\mathbf{H}^{\text{H}}\mathbf{H}\langle\mathbf{SS}\rangle)}{N_x N_y N_s}$;
21 **until** *convergence*;
22 Use $\mathbf{\Gamma_H}$ to drop redundant columns of $\mathbf{H}$ and corresponding columns of $\mathbf{S}$;
23 Project $\mathbf{S}$ entry-wise onto $\mathcal{F}$;

$a_m$s. In practice, these can be set to $\frac{1}{M}$. It should be noted that soft symbol estimates are computed and used throughout, as in BPCA-SVP, however, the knowledge of the input constellation is incorporated here in the form of the symbol prior. The posterior probabilities (conditioned on the MPDR output) are computed and stored in the $N_s \times M \times R$ array $\boldsymbol{\mathcal{Z}}$.

More details on the algorithms and their development are deferred to the full version of the paper.

## IV. Simulation Results

$R = 3$ sources emit at the frequency of 1 GHz $M$-QAM signals, which arrive through $L = 3$ paths each to a $6 \times 6$ URA with half-wavelength inter-sensor spacings. The azimuth and zenith angles of arrival are chosen uniformly at random from $(0, \pi)$ and $(0, \pi/2)$, respectively (i.e., at the "front" of the URA), and the complex path gains are drawn from the complex zero-mean, unit-variance Gaussian distribution. $R$ and the $L_r$s are over-estimated as $R_{\text{ini}} = 6$ and $L_{\text{ini}} = 6$, respectively. Information symbols are preceded by orthogonal training preambles[4] of (the smallest) length $R_{\text{ini}}$, which allows to account for scaling and permutation ambiguities and compute a first estimate of $\mathbf{H}$ and $\sigma_w^2$.

Figs. 1(a), (b) show the signal copying performance, in terms of the bit error rate (BER) as a function of the bit SNR, of the training-only-based MMSE detector and the detectors presented previously, for different constellation orders and for (a large number of) realizations in which *all* rank-selection criteria were successful. The superiority of the semi-blind over the training-only-based approach can be observed in both cases as the semi-blind receiver extracts channel and hence Tx signal information also from the Rx signals' information part. The deterministic scheme outclasses with low constellation order (Fig. 1(a)), while the situation reverses with higher-order (Fig. 1(b)) constellations. This is because, in the latter case, the distribution of the Tx symbols is closer to being continuous (Gaussian for BPCA-SVP), which favors their soft estimation. ILSP-SVP performs better than BPCA-SVP at moderate to high SNRs, where hard decisions are more reliable.

As Fig. 1(c) demonstrates, it is only at high SNR levels that the rank selection criterion of (4) can be practical. The EM-based schemes appear to enable the accurate estimation of the number of signals even in the presence of strong noise.

The average run-times[5] on a computer employing i5-8500 CPU@3.00 GHz and 8 GB RAM, and using the R2024b release of MATLAB©, are plotted versus the SNR level in Fig. 1(d) for the experiment of Fig. 1(b). The time required to estimate $R$ in ILSP-SVP is included. ILSP-SVP has the lowest computational cost, followed by MPDR-SVP. There are a few ways the computational complexity of the latter can be reduced. These include the use of the matrix inversion lemma on $\mathbf{\Sigma}_{\mathbf{Y}}$ and the compression of $\mathbf{Y}$ as detailed in [11].

In summary, ILSP-SVP appears to be an effective and efficient solution with the additional advantage of simplicity. However, the EM-based schemes are, by construction, able to *jointly* select and compute the Rx signal model, which renders them better suited for real-time implementation.

[4]Known to be MSE-optimal (see, e.g., [20] and references therein).

[5]MATLAB's© `svds` was used in the SVP steps. More efficient implementations, including randomized SVD, are, of course, possible.

## V. Future Work

Considering fully Bayesian extensions of the EM-based schemes presented here, in the vein of variational BPCA [21] or its equivalent in [22], can be the next step in this work, possibly accompanied with the adoption of a discretization-enforcing prior (cf. [23]) as an intermediate between the Gaussian (BPCA) and the ideal discrete (MPDR) priors adopted herein. The rather heuristic way (via the SVP step) the LL1 structure was enforced can be replaced by a more principled approach [24], [25] based on a channel prior that explicitly incorporates the spatial correlation present in practical arrays, not taken into account in (5). Message-passing variants of BPCA [26] and MPDR [10] are worth investigating as more realistic solutions to the large-scale JCD problem.

Orthogonal training sequences were employed in the simulations for the sake of convenience. This can, however, be very costly in terms of spectral efficiency, especially when the number of input signals, and hence the duration of the training session, is largely overestimated. Shorter preambles, designed to be "close-to-orthogonal" (e.g., with maximum incoherence [27]), would thus be preferable in practice and also need to be tested.

Scenarios involving missing/failing sensors should also be considered, as well as incremental/online extensions that can cope with large-scale and/or time-varying (possibly also with a changing number of signals [28]) systems, reducing the delay incurred from batch processing of a long sequence of snapshots [4]. Going from planar to higher-dimensional (e.g., 3-D) arrays [2] also seems feasible and rewarding.

## References


[1] C. Rizogiannis *et al.*, "Semi-blind maximum-likelihood joint channel/data estimation for correlated channels in multiuser MIMO networks," *Signal Process.*, vol. 90, pp. 1209—-1224, 2010.

[2] S. Miron *et al.*, "Tensor methods for multisensor signal processing," *IET Signal Process.*, vol. 14, no. 10, pp. 693–709, 2020.

[3] L. De Lathauwer, "Decompositions of a higher-order tensor in block terms — Part II: Definitions and uniqueness," *SIAM J. Matrix Anal. Appl.*, vol. 30, no. 3, pp. 1033–1066, 2008.

[4] E. Kofidis, "Block-term tensor decomposition in array signal processing," *Digital Signal Process.*, vol. 156, Jan. 2025, part A.

[5] A. A. Rontogiannis, E. Kofidis, and P. V. Giampouras, "Block-term tensor decomposition: Model selection and computation," *IEEE J. Sel. Topics Signal Process.*, vol. 15, no. 3, pp. 464–475, Apr. 2021.

[6] S. Talwar and A. Paulraj, "Blind separation of synchronous co-channel digital signals using an antenna array—–Part II: Performance analysis," *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 706–718, Mar. 1997.

[7] E. Kofidis, "Revisiting semi-blind block-term decomposition-based receivers for uniform rectangular arrays," in *Proc. SSP*, Edinburgh, UK, Jun. 2025.

[8] E. Nayebi and B. D. Rao, "Semi-blind channel estimation for multiuser massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 66, no. 2, pp. 540–553, Jan. 2018.

[9] ——, "Semi-blind channel estimation in massive MIMO systems with different priors on data symbols," in *Proc. ICASSP*, Calgary, AB, Canada, Apr. 2018.

[10] M. Al-Shoukairi and B. D. Rao, "Sparse signal recovery using MPDR estimation," in *Proc. ICASSP*, Brighton, UK, May 2019.

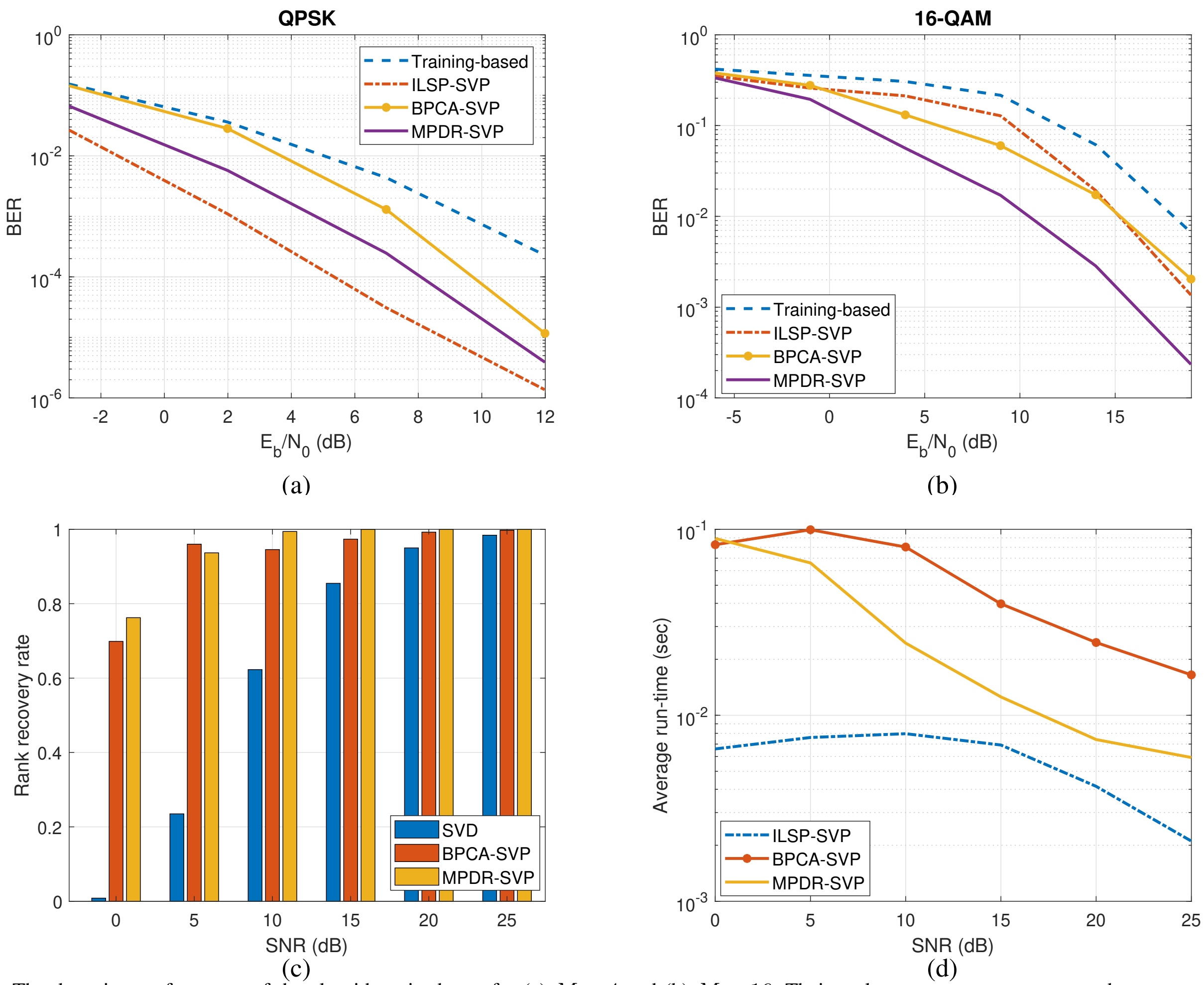


Fig. 1. The detection performance of the algorithms is shown for (a) $M = 4$ and (b) $M = 16$. Their rank recovery success rates and average run-times are depicted in (c) and (d), respectively, for the experiment of (b).


[11] ——, "Semi-blind channel estimation in MIMO systems with discrete priors on data symbols," *IEEE Signal Process. Lett.*, vol. 29, pp. 51–54, Nov. 2022.

[12] I. Domanov and L. De Lathauwer, "On uniqueness and computation of the decomposition of a tensor into multilinear rank-$(1, L_r, L_r)$ terms," *SIAM J. Matrix Anal. Appl.*, vol. 41, no. 2, pp. 747–803, 2020.

[13] E. Evert *et al.*, "Uniqueness result and algebraic algorithm for decomposition into multilinear rank-$(M_r, N_r, \cdot)$ terms and joint block diagonalization," in *Proc. CAMSAP-2023*, Los Sueños, Costa Rica, Dec. 2023.

[14] C. Chatzichristos *et al.*, "Blind fMRI source unmixing via higher-order tensor decompositions," *J. Neurosci. Methods*, vol. 315, pp. 17–47, Mar. 2019.

[15] P. Jain, R. Meka, and I. Dhillon, "Guaranteed rank minimization via singular value projection," in *Proc. NIPS*, Vancouver, BC, Canada, Dec. 2010.

[16] Z. Wen, W. Yin, and Y. Zhang, "Solving a low-rank factorization model for matrix completion by a nonlinear successive over-relaxation algorithm," *Math. Progr. Comput.*, vol. 4, pp. 333–361, 2012.

[17] R. H. Keshavan and S. Oh, "OptSpace: A gradient descent algorithm on the Grassman manifold for matrix completion," arXiv:0910.5260v2 [cs.NA], Nov. 2009.

[18] C. M. Bishop, "Bayesian PCA," in *Proc. NIPS*, Denver, CO, 30 Nov.–5 Dec. 1998.

[19] B. D. Rao, "Sparse Bayesian learning: A beamforming and Toeplitz approximation perspective," 2019. [Online]. Available: https://www.irit.fr/spars2019/en/program/Rao_SPARS_2019.pdf

[20] D. Katselis, E. Kofidis, and S. Theodoridis, "On training optimization for estimation of correlated MIMO channels in the presence of multiuser interference," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 4892–4904, Oct. 2008.

[21] C. M. Bishop, "Variational principal components," in *Proc. ICANN*, Edinburgh, UK, Sep. 1999.

[22] S. D. Babacan *et al.*, "Sparse Bayesian methods for low-rank matrix estimation," *IEEE Trans. Signal Process.*, vol. 60, no. 8, pp. 3964–3977, Aug. 2012.

[23] E. Kofidis, "Revisiting semi-blind block-term decomposition-based receivers for uniform rectangular arrays," Nov. 2024. [Online]. Available: https://www.techrxiv.org/doi/full/10.36227/techrxiv.173195536.67219802/v1

[24] P. Alquier, "Bayesian methods for low-rank matrix estimation: Short survey and theoretical study," in *Proc. ALT*, Singapore, Oct. 2013, pp. 309–323.

[25] P. V. Giampouras, A. A. Rontogiannis, and E. Kofidis, "Block-term tensor decomposition model selection and computation: The Bayesian way," *IEEE Trans. Signal Process.*, vol. 70, pp. 1704–1717, Mar. 2022.

[26] X. Zhong, C. Su, and Z. Fan, "Empirical Bayes PCA in high dimensions," *J. R. Stat. Soc. Series B*, vol. 84, pp. 853–878, 2022.

[27] Z. Zhou *et al.*, "Channel estimation for millimeter-wave multiuser MIMO systems via PARAFAC decomposition," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7501–7516, Nov. 2016.

[28] A. A. Rontogiannis, E. Kofidis, and P. V. Giampouras, "Online rank-revealing block-term tensor decomposition," *Signal Process.*, vol. 212, Nov. 2023.