A Khatri-Rao Subspace Approach to Blind Identification of Mixtures of Quasi-Stationary Sources

Abstract

Overall, I read and almost understood the Khatri-Rao framework for blind identification of mixtures of quasi-stationary sources in (Lee, 2013). In this report, the detail problem statement and the Khatri-Rao subspace-based methods are presented. Especially, I design new experiments with the synthetic quasi-stationary source signals (while the original paper used real speech recordings as the source signals). Numerical results of performance of the blind identification algorithm in terms of average mean square error are analyzed.

1. Introduction

Blind identification of a linear instantaneous mixture of quasi-stationary sources (QSS) has received much attention by its application to blind separation of speech and audio sources in microphone arrays.

The idea behind QSS-based blind identification is to exploit statistically time-varying characteristics of QSS to retrieve the mixing system. There are two main classes of formulations for QSS-based blind identification. The first one is based on parallel factor analysis (PARAFAC) which the blind identification problem is formulated as a three-way data array fitting problem (Sidiropoulos, 2000; Rong, 2005; Yeredor, 2002). The second one is joint diagonalization (JD) which the problem is formulated as a problem of diagonalizing multiple matrices (Pham, 2001; Ziehe, 2004).

From these formulations, effective blind identification algorithms have been developed. Under PARAFAC, there are trilinear alternating least squares (TALS) (Rong, 2005) and alternating-columns diagonal-centers (ACDC) (Yeredor, 2002). Under JD, including Pham's JD (Pham, 2001), fast Frobenius diagonalization (FFDiag) (Ziehe, 2004).

In (Lee, 2013), an alternative formulation for QSS-based blind identification using Khatri-Rao (KR) subspace is proposed. It should be noted that KR subspace was also proposed for DOA estimation of QSS (Ma, 2010). The key insight is the KR subspace formulation decomposes the blind identification problem into a number of less complex per-source. This is the significant different from PARAFAC and JD, which are inherently joint-source formulations.

Then, a specialized alternating projections (AP) algorithm for dealing with the simpler per-source problems is devised. A distinguishing result with this AP algorithm is its theoretically provable convergence. Using this important insight, a new blind identification algorithm is demonstrated to have good performance analysis.

The rest of the report is organized as follows. Section 2 describes the problem statement. In Section 3, the Khatri-Rao subspace-based methods are carried out. Numerical results are given in Section 4. Finally, Section 5 concludes the report.

Notation: Diag(x) denotes a diagonal matrix whose diagonal elements are $x_1, ..., x_n$; vec$(.)$ is a vectorization operator, where for $X = [x_1, ..., x_m] \in \mathbb{C}^{n\times m}$, we have vec$(X) = [x_1^T, ..., x_m^T] \in \mathbb{C}^{nm}$; vec$^{-1}(.)$ represents the inverse operation of vec$(.)$; $\otimes$ is the Kronecker product; $\odot$ is the Khatri-Rao product, where, given $A = [a_1, ..., a_k]$ and $B = [b_1, ..., b_k]$, we have $A \odot B = [a_1 \otimes b_1, ..., a_k \otimes b_k]$; $\mathcal{R}(X)$ denotes the range space of X; $\lambda_{min}(X)$ and $\lambda_{max}(X)$ denotes the magnitude-wise smallest and largest eigenvalues of X, respectively; $\|x\|_2$ and $\|x\|_F$ are the vector 2-norm and matrix Frobenius norm, respectively; $X^\dagger$ denotes the Moore-Penrose psedo-inverse of X; $X_{1:k}$ denotes a submatrix of X that consists of the first k columns of X.

2. Problem Statement

2.1. Signal Model and Assumptions

Consider a standard BID-QSS formulation as follows

$$x(t) = As(t) + v(t), t = 1, 2, ... \tag{1}$$

where $x(t) = [x_1(t), ..., x_N(t)]^T \in \mathbb{C}^N$ to be an N-sensor received signal vector, $s(t) = [s_1(t), ..., s_K(t)]^T \in \mathbb{C}^K$ to be a source signal vector, with K being the number of sources, $A = [a_1, ..., a_K] \in \mathbb{C}^{N\times K}$ to be a mixing matrix,

. **AUTHORERR: Missing \icmlcorrespondingauthor.**

Proceedings of the 38^{th} International Conference on Machine Learning, PMLR 139, 2021.

and $v(t) \in \mathbb{C}^N$ to be noise. It is assumed that

(A1) The source signals, $s_k(t), k = 1, ..., K$, are statistically independent of each other.

(A2) Each $s_k(t)$ is zero-mean wide-sense quasi-stationary, i.e., $E[|s_k(t)|^2]$ changes with t, but it is fixed within every local time interval $[(m-1)L+1, mL]$, $m \in [1, L]$. L is window length.

(A3) The noise vector, $v(t)$, is wide-sense stationary with mean zero and covariance $\sigma^2 I$. And, it is statistically independent of $s(t)$.

The local covariances of $x(t)$ is defined as

$$R_m = E[x(t)x(t)^H], t \in [(m-1)L+1, mL] \quad (2)$$

Note that, in practice, R_m can be estimated by local covariance sampling, e.g., $R_m \simeq (1/L)\sum_{t=(m-1)L+1}^{mL} x(t)x(t)^H$. From (1), it is shown that R_m follows to the model

$$R_m = AD_mA^H + \sigma^2 I, \quad (3)$$

where D_m are the local covariances of $s(t)$ and are given by $D_m = \text{Diag}(d_m)$, in which $d_m = [d_{m,1}, d_{m,2}, ..., d_{m,K}]^T$, $d_{m,k} = E[|s_k(t)|^2]$ for any $t \in [(m-1)L+1, mL]$.

2.2. Local Covariances Model and Khaitri-Rao Subspace

Suppose that we have measured a number of M local covariances of $x(t)$, or $R_1, ..., R_M$. We need to exploit the subspace characteristics of $R_1, ..., R_M$ for blind identification of A.

Assume a noise covariance-free scenario as follows

$$R_m = AD_mA^H, m = 1, ..., M \quad (4)$$

Consider the vectorization of R_m in (4)

$$y_m \triangleq \text{vec}(R_m) = (A^* \odot A)d_m \in \mathbb{C}^{N^2}, \quad (5)$$

where $A^* \odot A$ is a self-Khatri-Rao product of A and takes the form $A^* \odot A = [a_1^* \otimes a_1, ..., a_K^* \otimes a_K] \in \mathbb{C}^{N^2 \times K}$

Eq. (5) is virtually identical to a linear instantaneous mixture signal model, with a mixing matrix $A^* \odot A$ and a source vector d_m. Hence, the insight is that exploiting the self-Khatri-Rao product structure of the virtual mixing matrix $A^* \odot A$ to identify its physical counterpart, A.

The following assumptions are made.

(A4) The mixing matrix A has full Kruskal rank, i.e., any $\min\{K, N\}$ columns of A are linearly independent.

(A5) Let $\Psi = [d_1, ..., d_M]^T \in \mathbb{C}^{M \times K}$. The matrix Ψ has full column rank.

Fact 1. Assume (A4). The matrix $A^* \odot A$ has full column rank if $K \leq 2N - 1$ (Sidiropoulos, 2000).

Moreover, (A5) means that the source local variances are assumed to be sufficiently time-varying and different in their variations, thereby satisfying the full column rank assumption on Ψ.

$$Y \triangleq [y_1, ..., y_M] = (A^* \odot A)\Psi^T \in \mathbb{C}^{N^2 \times M} \quad (6)$$

Since $A^* \odot A$ and Ψ are of full column rank (assuming $K \leq 2N-1$), so Y has rank K and admits a singular value decomposition (SVD)

$$Y = U_s \Sigma_s V_s^H, \quad (7)$$

where $\Sigma_s \in \mathbb{R}^{K \times K}$ is the nonzero singular value matrix, $U_s \in \mathbb{C}^{N^2 \times K}$ and $V_s \in \mathbb{C}^{M \times K}$ are the associated left and right singular matrices, respectively. And, we have

$$\mathcal{R}(U_s) = \mathcal{R}(A^* \odot A), \quad (8)$$

The subspace $\mathcal{R}(U_s)$ or $\mathcal{R}(A^* \odot A)$ will be called the *Khatri-Rao (KR) subspace*.

2.3. Preprocessing Method

2.3.1. Noise Covariance Removal

It is well known that $\lambda_{min}(R_m) = \lambda_{min}(AD_mA^H) + \sigma^2$ and $\lambda_{min}(AD_mA^H) \geq 0$ (Stoica, 2005). For $N > K$, we have $\lambda_{min}(AD_mA^H) = 0$. Hence, for this strictly overdetermined case, we can estimate σ^2 by

$$\hat{\sigma}^2 = \min_{m=1,...,M} \lambda_{min}(R_m), \quad (9)$$

and then subtract $\lambda^2 I$ from R_m. It should be note that this procedure has been previously suggested in (Rahbar, 2005).

2.3.2. Prewhitening Procedure

The prewhitening procedure helps transform the problem A becomes unitary (Belouchrani, 1997). Consider the time-averaged global covariance as follows

$$\bar{R} = \frac{1}{M}\sum_{m=1}^{M} R_m = A\bar{D}A^H, \quad (10)$$

where $\bar{D} = (1/M)\sum_{m=1}^{M} D_m$. Since $\bar{D}$ is a positive definite matrix, we can apply a square-root factorization $\bar{R} = BB^H$, where $B \in \mathbb{C}^{N \times K}$ has full column rank (for $N \geq K$). The prewhitening operation is given by

$$\tilde{R}_m = B^\dagger R_m \left(B^\dagger\right)^H, m = 1, ..., M \quad (11)$$

The prewhitened local covariances $\tilde{R}_m$ can be written as

$$\tilde{R}_m = \tilde{A}\tilde{D}_m\tilde{A}^H, m = 1, ..., M, \tag{12}$$

where $\tilde{A} = B^\dagger A\bar{D}^{1/2} \in \mathbb{C}^{K\times K}$ is the transformed mixing matrix, and $\tilde{D}_m = \bar{D}^{-1}D_m$ the transformed source local covariances.

3. Method

3.1. Column-Decoupled Blind Identification

3.1.1. KR Subspace Criteria 1

BID-QSS aims at estimating the mixing matrix A from the observed local covariances $R_1, ..., R_M$, given the number of sources K. From the KR subspace identity (8), we see that any column a_k of the true mixing matrix A satisfies $a_k^* \otimes a_k \in \mathcal{R}(U_s)$. This observation leads to the following criterion for blind identification of A:

Criterion 1:

$$\begin{aligned} &\text{find} \quad a \in \mathbb{C}^N \\ &\text{such that} \quad a^* \otimes a \in \mathcal{R}(U_s). \end{aligned}$$

Criterion 1 suggests a column-decoupled BID approach to find one of the a_k's, assuming unique identifiability. Consider the following theorem:

Theorem 1. *Assume (4), (A4), and (A5). A sufficient and necessary condition for a satisfies Criterion $1 \Leftrightarrow a = ca_k$ for some k and constant $c \in \mathbb{C}$ is when $K \leq 2N - 2$.*

Theorem 1 confirms that Criterion 1 can operate in the underdetermined case.

3.1.2. Alternating Projections Algorithm

To implement Criterion 1, it is natural to formulate it as an optimization problem as follows

$$\begin{aligned} \min_{a\in\mathbb{C}^N} \ & (a^* \otimes a)^H P_s^\perp (a^* \otimes a) \\ s.t. \quad & \|a\|_2^2 = 1, \end{aligned} \tag{13}$$

where $P_s^\perp = I - U_sU_s^H$ denotes the orthogonal complement projector of the KR subspace $\mathcal{R}(U_s)$. In the other words, minimizing the projection residual of $a^* \otimes a$ on $\mathcal{R}(U_s)$. The approach is based on an alternative formulation of (13) that will lead to an iterative algorithm. The problem (13) is equivalent to

$$\begin{aligned} \min_{\alpha\in\mathbb{R}, a\in\mathbb{C}^N, h\in\mathbb{C}^{N^2}} \ & \|\alpha a^* \otimes a - h\|_2^2 \\ s.t. \quad & \alpha \in \{\pm 1\}, \|a\|_2^2 = 1, h \in \mathcal{R}(U_s), \end{aligned} \tag{14}$$

The equivalence of problems (13) and (14) is shown as follows. Fixing (α, a), the optimization of (14) over h is a linear projection problem, whose solution is

$$h = U_sU_s^H(\alpha a^* \otimes a), \tag{15}$$

By substituting (15) into (14), problem (14) can be shown as

$$\begin{aligned} &\min_{a\in\{\pm 1\}, \|a\|_2^2=1} \left\|(I - U_sU_s^H)(\alpha a^* \otimes a)\right\|_2^2 \\ &= \min_{\|a\|_2^2=1} (a^* \otimes a)^H P_s^\perp (a^* \otimes a), \end{aligned} \tag{16}$$

which is exactly the same as problem (13).

Problem (14) has an interpretation of finding a pair of closets points in two sets, namely, $h \in \mathcal{R}(U_s)$ and $(\alpha, a) \in \{\pm 1\} \times \mathcal{U}^N$, where $\mathcal{U}^n = \{x \in \mathbb{C}^n | \, \|x\|_2^2 = 1\}$. In additions, the formulation in (14) can be solved by applying the alternating projection (APs) (Boyd, 2003).

Essentially, the idea of AP is to fix (α, a) and solve (14) with respect to (w.r.t.) h at one time, and then fix h and solve (14) w.r.t. (α, a) at another time. For the partial optimization of (14) over h, we have seen that the solution is (15). Let examine the partial optimization of (14) over (α, a). By denoting $H = \text{vec}^{-1}(h) \in \mathbb{C}^{N\times N}$, problem (14) can be expressed as

$$\begin{aligned} &\min_{\alpha,a,H} \left\|\alpha aa^H - H\right\|_F^2 \\ &s.t. \quad \alpha \in \{\pm 1\}, \|a\|_2^2 = 1, \text{vec}(H) \in \mathcal{R}(U_s), \end{aligned} \tag{17}$$

For any $\alpha \in \{\pm 1\}$ and $\|a\|_2^2 = 1$, the objective function of (17) yields

$$\begin{aligned} \left\|\alpha aa^H - H\right\|_F^2 &= 1 - 2\alpha \text{Re}\{a^H Ha\} + \|H\|_F^2 \\ &\geq 1 - 2|\text{Re}\{a^H Ha\}| + \|H\|_F^2, \end{aligned} \tag{18}$$

where equality in (18) holds when $\alpha = \text{Re}\{a^H Ha\}/|\text{Re}\{a^H Ha\}|$. Moreover, the second term in (18) is minimized when $|\text{Re}\{a^H Ha\}|$ is maximized, and the latter is achieved when a aligns to a magnitude-wise most significant eigenvector of $\left(H + H^H\right)/2$ (note that $\text{Re}\{a^H Ha\} = \frac{1}{2}a^H(H + H^H)a$). Hence, the partial optimization of (17) w.r.t. (α, a) has a closed-form solution given by

$$a = q_{\max}\left(\frac{1}{2}(H + H^H)\right), \alpha = \frac{\lambda_{\max}\left(\frac{1}{2}(H + H^H)\right)}{\left|\lambda_{\max}\left(\frac{1}{2}(H + H^H)\right)\right|}, \tag{19}$$

where $\lambda_{\max}(X)$ denotes the largest eigenvalue of X (magnitude-wise), and $q_{\max}(X)$ denotes a unit-2-norm eigenvector of X associated with $\lambda_{\max}(X)$. The implementation of the AP method is shown in Algorithm 1.

3.2. Complete Blind Identification Using The Column-Decoupled Solutions

3.2.1. KR Subspace Criteria 2

A significant advantage of the column-decoupled BID criterion (in subsection 3.1) is that an efficient algorithm can be

Algorithm 1 AP algorithm for problem (14).

Input: the KR subspace matrix U_s;
1: $H := \text{vec}^{-1}(U_s\zeta), \zeta \sim \mathcal{CN}(0, I)$ (random initialization);
repeat
2: $a := q_{\max}\left(\frac{1}{2}(H + H^H)\right)$;
3: $\alpha := \frac{\lambda_{\max}\left(\frac{1}{2}(H+H^H)\right)}{\left|\lambda_{\max}\left(\frac{1}{2}(H+H^H)\right)\right|}$;
4: $H := \text{vec}^{-1}(U_s U_s^H(\alpha a^* \otimes a))$;
until a stopping criterion is satisfied.
Output: a as an estimate of a column of the mixing matrix.

developed. However, Criterion 1 does not tell how all the columns of A can simultaneously be identified.

This subsection will study how to estimate the whole A based on the column-decoupled BID solutions. Consider the following criterion:

Criterion 2:

$$\begin{aligned} \text{find} \quad & A \in \mathbb{C}^{N\times K}, \Theta \in \mathbb{C}^{K\times K} \\ \text{such that} \quad & U_s = (A^* \odot A)\Theta. \end{aligned}$$

From (8), it holds true that any column of U_s lies in $\mathcal{R}(A^* \odot A)$.

3.2.2. Successive Optimization For Unitary A

We can use an optimization formulation, i.e., least-squares fitting, to implement the Criterion 2 as follows

$$\min_{A\in\mathbb{C}^{N\times K},\Theta\in\mathbb{C}^{K\times K}} \|U_s - (A^* \odot A)\Theta\|_F^2 . \tag{20}$$

It should be noted that there exists pragmatic algorithms, e.g., ACDC and TALS, that have been found to produce reasonable estimate for problems (20) (Boyd, 2003). However, ACDC and TALS require initialization of A. In fact, poor initializations are likely to slow down convergence or leading to some unsatisfactory estimates. Hence, a two-state approach can be considered in which running AP multiple times to find some columns of A (or all) and then using them to initialize for ACDC or TALS.

Especially, for the overdetermined case, we can develop an efficient algorithm in which transforming A to a unitary matrix by prewhitening procedure.

When A is unitary, Θ is also unitary. Hence, we consider a modified form of the KR subspace fitting formulation (20) as follows:

$$\begin{aligned} \min_{A\in\mathbb{C}^{K\times K},\Theta\in\mathbb{C}^{K\times K}} & \|U_s - (A^* \odot A)\Theta\|_F^2 \\ s.t. \quad & \Theta\Theta^H = I, \end{aligned} \tag{21}$$

Let $Q = \Theta^H$. By the rotational invariance of $\|\|_F$, we can rewrite (21) as

$$\begin{aligned} \min_{A,Q} & \|U_S Q - (A^* \odot A)\|_F^2 \\ s.t. \quad & Q^H Q = I. \end{aligned} \tag{22}$$

In addition, by substituting $h_k = U_s q_k \in \mathcal{R}(U_s)$, where q_k is the kth column of Q, problem (22) can be equivalently expressed as

$$\begin{aligned} \min_{A,h_1,...,h_K} & \sum_{k=1}^{K} \|h_k - a_k^* \otimes a_k\|_2^2 \\ s.t. \quad & h_k \in \mathcal{R}(U_s), \quad k = 1, ..., K, \\ & h_k^H h_l = 0, \quad \forall k \neq l, \\ & \|h_k\|_2^2 = 1, \quad k = 1, ..., K. \end{aligned} \tag{23}$$

Let modify problem (23) by replacing the constrains $\|h_k\|_2^2 = 1$ by $\|a_k\|_2^2 = 1$ as follows

$$\begin{aligned} \min_{A,h_1,...,h_K} & \sum_{k=1}^{K} \|h_k - a_k^* \otimes a_k\|_2^2 \\ s.t. \quad & h_k \in \mathcal{R}(U_s), \quad k = 1, ..., K, \\ & h_k^H h_l = 0, \quad \forall k \neq l, \\ & \|a_k\|_2^2 = 1, \quad k = 1, ..., K. \end{aligned} \tag{24}$$

A key observation is that problem (24) can be expressed as (25) where

$$\mathcal{H}_k(h_1, ..., h_{k-1}) = \{h \in \mathcal{R}(U_s) | h^H h_l = 0, l = 1, ..., k-1\}. \tag{26}$$

The express in (25) suggests that we can apply a successive optimization strategy. To be specific, we decouple (25) into K sequentially processed stages. At stage k, we aim at solving

$$\begin{aligned} \left(\hat{a}_k, \hat{h}_k\right) = & \arg\min_{a_k,h_k} \|h_k - a_k^* \otimes a_k\|_2^2 \\ s.t. \quad & \|a_k\|_2^2 = 1, h_k \in \mathcal{H}_k(\hat{h}_1, ..., \hat{h}_{k-1}) \end{aligned} \tag{27}$$

where $\hat{h}_1, ..., \hat{h}_{k-1}$ are the decisions in the previous stages $1, ..., k-1$. Moreover, it can be shown that since $\hat{h}_1, ..., \hat{h}_{k-1} \in \mathcal{R}(U_s)$, the subspace $\mathcal{H}_k(\hat{h}_1, ..., \hat{h}_{k-1})$ takes an explicit form

$$\mathcal{H}_k(\hat{h}_1, ..., \hat{h}_{k-1}) = \mathcal{R}(P^{\perp}_{\hat{H}_{1:k-1}} U_s) \tag{28}$$

where $\hat{H}_{1:k-1} = [\hat{h}_1, ..., \hat{h}_{k-1}]$, $P^{\perp}_{\hat{H}_{1:k-1}} = I - \hat{H}_{1:k-1}(\hat{H}^H_{1:k-1}\hat{H}_{1:k-1})^{-1}\hat{H}^H_{1:k-1}$. Now, we can see the interesting connection: *problem (27) is equivalently the AP problem (14), with the original subspace matrix U_s being replaced by $P^{\perp}_{\hat{H}_{1:k-1}} U_s$*. As a result, problem (27) can be solved by applying AP algorithm (Algorithm 1). If the previous stages $1, ..., k-1$ have perfectly identified some

$$\min_{\substack{\|a_1\|_2^2=1,\\ h_1\in\mathcal{R}(U_s)}} \left\{ \|h_1 - a_1^* \otimes a_1\|_2^2 + \left[\min_{\substack{\|a_2\|_2^2=1,\\ h_2\in\mathcal{H}_2(h_1)}} \|h_2 - a_2^* \otimes a_2\|_2^2 + ... + \left(\min_{\substack{\|a_K\|_2^2=1,\\ h_K\in\mathcal{H}_K(h_1,...,h_{K-1})}} \|h_K - a_K^* \otimes a_K\|_2^2 \right) \right] \right\} \tag{25}$$

Algorithm 2 Prewhitened alternating projection algorithm.

Input: local covariance matrices $R_1, ..., R_M$;
1: $\bar{R} = \frac{1}{M}\sum_{m=1}^{M} R_m$;
2: compute a square-root factorization $\bar{R} = BB^H$;
3: $R_m := B^\dagger R_m (B^\dagger)^H, m = 1, ..., M$; (prewhitening)
4: compute the SVD (U, Σ, V) of $Y = [\text{vec}(R_1), ..., \text{vec}(R_m))]$, $U_s := U_{1:K,}$, and set $k = 1$;
5: run Algorithm 1 with $P^{\perp}_{\hat{H}_{1:k-1}} U_s$ as the input to obtain $(\hat{a}_k, \hat{h}_k)$, where $\hat{H}_{1:k-1} = [\hat{h}_1, ..., \hat{h}_{k-1}]$;
6: set $k := k + 1$ and goto step 5 until $k > K$;
7: $\hat{A} := B[\hat{a}_1, ..., \hat{a}_K]$; (post-dewhitening)
Output: $\hat{A}$ as an estimate of the mixing matrix

Table 1. Simulation Parameters

Name	Symbol	Value
Sequence length	T	79800
Number of sensors	N	5
Number of sources	K	4
Fixed time frame	L	200
Error bound	ϵ	10^{-6}
Maximal number of iteration	iter_{max}	100
Minimal local time frame	L_{low}	100
Maximal local time frame	L_{upp}	300

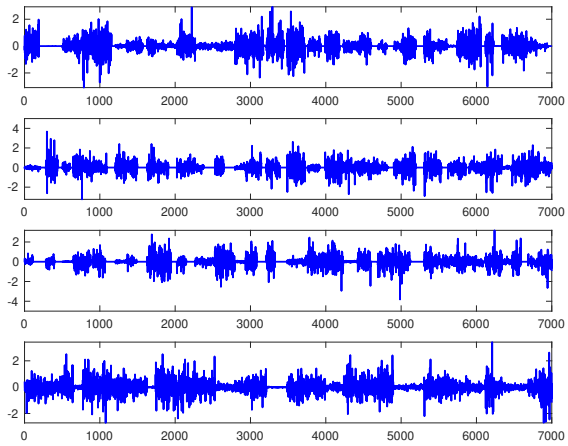

Figure 1. Illustration of a segment of synthetic quasi-stationary source signals $s_k(t)$. The only the real parts of the signals are plotted.

of the mixing matrix columns, $\hat{h}_1 = a_1^* \otimes a_1, ..., \hat{h}_{k-1} = a_{k-1}^* \otimes a_{k-1}$, then we can show that

$$\mathcal{H}_k(\hat{h}_1, ..., \hat{h}_{k-1}) = \mathcal{R}([a_k^* \otimes a_k, ..., a_K^* \otimes a_K]), \tag{29}$$

in which the previously found columns are removed from the subspace.

The AP-based successive optimization method proposed above is summarized in Algorithm 2. The algorithm is called as the *prewhitened alternating projection algorithm (PAPA)* for convenience. Overall, there are two major advantages of PAPA: (1) PAPA deals with K AP problems only; (2) Since A is unitary, the AP convergence is expected to be fast.

4. Experiment

In this section, some simulations are provided to analyze the advantages of the proposed algorithms. The simulation parameters are given in Table 1. The simulation settings are described as follows. The mixing matrix $A \in \mathbb{R}^{N\times K}$ is randomly generated at each trial with columns being normalized to unit 2-norm.

The quasi-stationary source signals, $s_k(t)$, were synthetically generated by a random generation procedure given in Table 2 (Ma, 2010). This procedure generates a locally stationary zero-mean complex Laplacian process. It's variances randomly varying from one frame to another. In additions, the duaration of each local time frame, L, is randomly drawn following a uniform distribution on $[L_{low}, L_{upp}]$. The purpose of this process is to simulate a more realistic situation in which the local stationary periods are uncertain and varying; e.g., in speech.

In order to obtain more local covariances under limited signal length, 50% overlapping frames is employed in acquiring R_m's, i.e., $R_m = (1/L)\sum_{t=0.5(m-1)L+1}^{0.5(m-1)L+L} x(t)x(t)^H$. Moreover, the noise covariance removal procedure (in subsection 2.3.1) is applied. For the proposed algorithms, a standard stopping criterion is adopted. Specifically, $|f^{(n)} - f^{(n-1)}| < \epsilon = 10^{-6}$, where $f^{(n)}$ is the objective value of the algorithm at the nth iteration.

The performance measure is the average mean square error

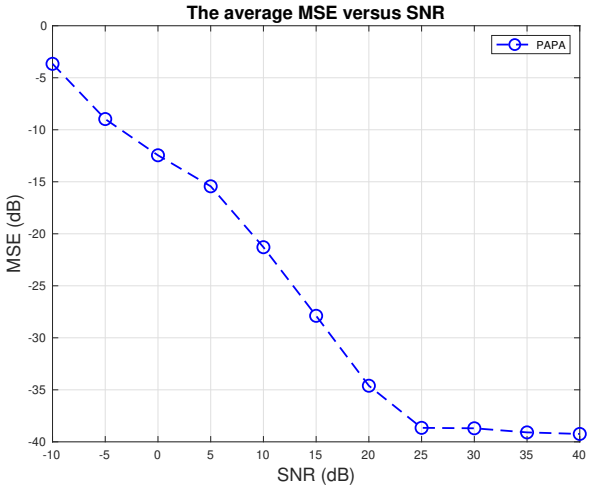

Figure 2. The average MSE of the PAPA algorithm versus the SNR.

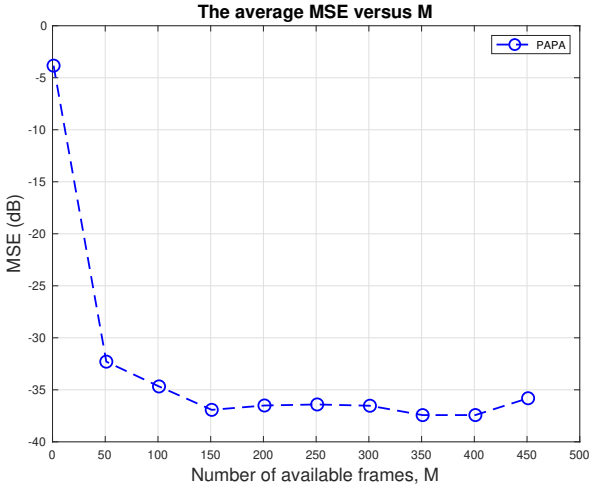

Figure 3. The average MSE of the PAPA algorithm versus the number of frames (M) with SNR $= 25$ dB.

Table 2. Generation of Synthetic Quasi-Stationary Signals (Ma, 2010)

Given	Bounds L_{low}, L_{upp}, and a sequence length T.
Step 1.	$T_{cur} := 0$.
Step 2.	Randomly generate L following a uniform distribution on $[L_{low}, L_{upp}]$.
Step 3.	Randomly generate σ_s following a uniform distribution on $[0, 1]$.
Step 4.	For $t = T_{cur}, T_{cur}+1, ..., T_{cur}+L-1$, randomly generate $s(t) = S_R(t) + jS_I(t)$ where $S_R(t)$ and $S_I(t)$ are i.i.d Laplacian distributed with zero mean and variance $\sigma_s^2/2$.
Step 5.	$T_{cur} := T_{cur} + L$. If $T_{cur} < T$ then go to Step 2.
Output	$s(t)_{t=0}^{T-1}$.

(MSE), defined as

$$\text{MSE} = \min_{\substack{\pi \in \Pi, \\ c_1,...,c_K \in \pm 1}} \frac{1}{K} \sum_{k=1}^{K} \left\| \frac{a_k}{\|a_k\|_2} - c_k \frac{\hat{a}_{\pi(k)}}{\|\hat{a}_{\pi(k)}\|_2} \right\|, \tag{30}$$

where Π is the set of all bijections $\pi : \{1, ..., K\} \to \{1, ..., K\}$; A and $\hat{A}$ are the true and estimated mixing matrices, respectively.

The signal-to-noise ratio (SNR) is defined as SNR $= (\frac{1}{T}\sum_{t=0}^{T-1} E[\|As(t)\|_2^2])/E[\|v(t)\|_2^2]$, where $T = LM$ is the total number of samples.

In Fig. 1, a segment of the synthetic quasi-stationary signals is illustrated. As we can see, the frame intervals of the 4 source signals are not uniform and not synchronized. I apply the KR subspace-based methods by choosing a fixed frame period of $L = 200$ (as Table 1).

Fig. 2 shows the average MSE of the PAPA algorithm versus the SNR. The increase of SNR results in the reduction of MSE since the decrease of the effect of noise. For example, when SNR runs from -10 dB to 25 dB, MSE gets the values between -5 dB to -38 dB.

Fig. 3 investigates the average MSE of the PAPA algorithm versus the number of frames or the number of local covariances (N) with SNR $= 25$ dB. We can see that PAPA works better for larger number of frames. When $M \geq 150$, MSE gets the minimal value, e.g., -37 dB and MSE is quite stable when $M \in [150, 450]$.

5. Conclusion

Maximum Ratio Transmission and Zero-Forcing Beamforming methods were investigated to solve the optimal multiuser transmit beamforming problems. Specially, performance analysis of these methods were shown in terms of the average sum rate and the feasibility rate.

By doing this project, I have learned something: (1) Maximum Ratio Transmission and Zero-Forcing Beamforming methods; (2) Have an understanding of some multiuser transmit beamforming optimization algorithms (3) Try to understand how authors manipulate the optimization problems (e.g., convexity, KKT conditions); (4) Get some hand-on experiences on implementing some algorithms on Matlab.

References

Ka-Kit Lee, Wing-Kin Ma, Xiao Fu, Tsung-Han Chan, Chon-Yung Chi. A Khatri–Rao subspace approach to blind identification of mixtures of quasi-stationary sources. *Signal Processing*, vol. 93, pp. 3515-3527, Dec.

2013.

N. D. Sidiropoulos, R. Bro, and G. B. Giannakis. Parallel factor analysis in sensor array processing. *IEEE Transactions on Signal Processing*, vol. 48, no. 8, pp. 2377-2388, Aug. 2000.

Y. Rong, S. A. Vorobyov, A. B. Gershman, and N. D. Sidiropoulos. Blind spatial signature estimation via time-varying user power loading and parallel factor analysis. *IEEE Transactions on Signal Processing*, vol. 53, no. 5, pp. 1697-1710, May. 2005.

A. Yeredor. Non-orthogonal joint diagonalization in the least-squares sense with application in blind source separation. *IEEE Transactions on Signal Processing*, vol. 50, no. 7, pp. 1545-1553, Jul. 2002.

D.-T. Pham and J.-F. Cardoso. Blind separation of instantaneous mixtures of nonstationary sources. *IEEE Transactions on Signal Processing*, vol. 49, no. 9, pp. 1837-1848, Sep. 2011.

A. Ziehe, P. Laskov, and G. Nolte. A fast algorithm for joint diagonalization with non-orthogonal transformations and its application to blind source separation. *Journal of Machine Learning Research*, vol. 5, pp. 801-818, 2004.

W.-K. Ma, T.-H. Hsieh, and C.-Y. Chi. DOA estimation of quasistationary signals with less sensors than sources and unknown spatial noise covariance: A Khatri-Rao subspace approach. *IEEE Transactions on Signal Processing*, vol. 58, no. 4, pp. 2168-2180, Apr. 2010.

P. Stoica, R. Moses. Spectral Analysis of Signals. Pearson Prentice Hall, 2005.

K. Rahbar, J. Reill. A frequency domain method for blind source separation of convolutive audio mixtures. *IEEE Transactions on Speech and Audio Processing*, pp. 823-844, Sep. 2005.

A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, E. Moulines. A blind source separation technique using second-order statistic. *IEEE Transactions on Signal Processing*, pp. 434-444, Feb. 1997.

S. Boyd, J. Dattoro. Alternating Projections. EE392: Optimization Projects, Stanford University, 2003.