SUBSPACE TRACKING IN COLORED NOISE BASED ON OBLIQUE PROJECTION

Minhua Chen, Zuoying Wang

Department of Electronic Engineering
Tsinghua University, Beijing, 100084, P.R.China
chmh00@mails.tsinghua.edu.cn

ABSTRACT

Projection Approximation Subspace Tracking (PAST) algorithm gives biased subspace estimation when the received signal is corrupted by colored noise. In this paper, an unbiased version of PAST is proposed for the colored noise scenario. Firstly, a maximum likelihood (ML) and minimum variance unbiased (MVUB) estimator for the clean signal is derived using simultaneous diagonalization and oblique projection. Then, we provide a recursive algorithm, named Oblique PAST (obPAST), to track the signal subspace and update the estimator in colored noise. Experimental results show the effectiveness of the obPAST algorithm.

1. INTRODUCTION

Subspace based methods play a very important role in signal processing. In the subspace approach, the noisy signal is decomposed into signal-plus-noise subspace and noise subspace using eigenvalue decomposition (EVD) or singular value decomposition (SVD). Then the noisy signal is enhanced by removing the noise subspace and recovering the clean signal from the remaining signal subspace.

Generally speaking, the subspace based algorithms require repeated matrix decomposition once a new signal sample is received, rendering the algorithms computationally inefficient. Yang [1] proposed Projection Approximation Subspace Tracking (PAST) algorithm to track the signal subspace efficiently, with EVD or SVD avoided. Direct extensions of PAST include OPAST[2], NIC[3], etc. However, PAST and most of its extensions are based on the white noise assumption. When the additive noise is colored, they all give biased estimations. The colored noise case is only considered in an instrumental variable(IV) approach so far [4].The IV method requires that an IV vector that is uncorrelated with the noise vector can be found. In real applications, however, this requirement is not always fulfilled.

In this paper, we consider another approach for subspace tracking in colored noise. Firstly we apply simultaneous diagonalization for the signal and noise covariance matrix. After that, an estimator for the clean signal is derived using oblique projection. A recursive method for tracking the signal subspace and updating the projector, named Oblique Projection Approximation Subspace Tracking (obPAST) is subsequently proposed. Experimental results show the effectiveness of the proposed algorithm.

2. OBLIQUE PROJECTION BASED ON SIMULTANEOUS DIAGOLIZATION

2.1. Derivation of the Oblique Projector in Colored Noise

Consider the linear signal model for the $m \times 1$ dimensional clean signal $\boldsymbol{x}$

$$\boldsymbol{x} = \boldsymbol{H} \cdot \boldsymbol{s} \tag{1}$$

where $\boldsymbol{H}$ is a $m \times r$ matrix($r < m$) with rank r. $\boldsymbol{s}$ is the source signal with dimension $r \times 1$. The covariance matrix of $\boldsymbol{x}$ is then written as

$$\boldsymbol{R_x} = \boldsymbol{H} \cdot \boldsymbol{R_s} \cdot \boldsymbol{H}^H. \tag{2}$$

$\boldsymbol{R_s}$ is the covariance matrix of $\boldsymbol{s}$ and is assumed to be full rank; thus the rank of $\boldsymbol{R_x}$ is r. $\boldsymbol{H}^H$ denotes the Hermitian transpose of $\boldsymbol{H}$. With this signal model, the corrupted signal(with dimension $m \times 1$) is given by

$$\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{n} = \boldsymbol{H} \cdot \boldsymbol{s} + \boldsymbol{n}. \tag{3}$$

We assume that the colored Gaussian noise $\boldsymbol{n}$(with dimension $m \times 1$) is uncorrelated with the clean signal $\boldsymbol{x}$ and has a full-rank covariance matrix $\boldsymbol{R_n}$. The covariance matrix of the observed corrupted signal $\boldsymbol{y}$ is given by $\boldsymbol{R_y} = \boldsymbol{R_x} + \boldsymbol{R_n} = \boldsymbol{H} \cdot \boldsymbol{R_s} \cdot \boldsymbol{H}^H + \boldsymbol{R_n}$.

According to matrix analysis theory, there exists a *non-singular* matrix $\boldsymbol{V}$ with dimension $m \times m$ such that the covariance matrix $\boldsymbol{R_x}$ and $\boldsymbol{R_n}$ are simultaneously diagonalized as [5], [6]

$$\boldsymbol{V}^H \boldsymbol{R_x} \boldsymbol{V} = \boldsymbol{\Lambda_x}; \quad \boldsymbol{V}^H \boldsymbol{R_n} \boldsymbol{V} = \boldsymbol{I} \tag{4}$$

Since the rank of $\boldsymbol{R_x}$ is r, the diagonal matrix $\boldsymbol{\Lambda_x}$ can be expressed as $\boldsymbol{\Lambda_x} = diag\left(\lambda_{\boldsymbol{x}}^{(1)}, \lambda_{\boldsymbol{x}}^{(2)}, \cdots, \lambda_{\boldsymbol{x}}^{(r)}, 0, 0, \cdots, 0\right)$ with $\lambda_{\boldsymbol{x}}^{(1)} \geq \lambda_{\boldsymbol{x}}^{(2)} \geq \cdots \lambda_{\boldsymbol{x}}^{(r)} > 0$. Thus (4) is equivalent to

$$\boldsymbol{V}^H \boldsymbol{R_y} \boldsymbol{V} = \boldsymbol{\Lambda_y}; \quad \boldsymbol{V}^H \boldsymbol{R_n} \boldsymbol{V} = \boldsymbol{I} \tag{5}$$

with $\boldsymbol{\Lambda_y} = diag\left(\lambda_{\boldsymbol{x}}^{(1)}+1,\cdots,\lambda_{\boldsymbol{x}}^{(r)}+1,1,1,\cdots,1\right)$. Equation (5) also implies $\boldsymbol{V}$ and $\boldsymbol{\Lambda_y}$ are eigenvector matrix and eigenvalue matrix of $\boldsymbol{R_n^{-1}R_y}$ [6].

We express $\boldsymbol{V}$ as $\boldsymbol{V} = [\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_m] = [\boldsymbol{V}_r, \boldsymbol{V}_{m-r}]$ where $\boldsymbol{V}_r = [\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_r]$ denotes the first r columns of $\boldsymbol{V}$ and $\boldsymbol{V}_{m-r}$ denotes the remaining columns. Since $\boldsymbol{V}$ is non-singular, $\boldsymbol{R_y}$ can be expressed as follows:

$$\begin{aligned}\boldsymbol{R_y} &= \boldsymbol{V}^{-H}\boldsymbol{\Lambda_y}\boldsymbol{V}^{-1} = (\boldsymbol{R_n V})\,\boldsymbol{\Lambda_y}(\boldsymbol{R_n V})^H \\ &= [\boldsymbol{R_n V}_r, \boldsymbol{R_n V}_{m-r}]\cdot\boldsymbol{\Lambda_y}\cdot[\boldsymbol{R_n V}_r, \boldsymbol{R_n V}_{m-r}]^H\end{aligned}$$

As both $\boldsymbol{R_n}$ and $\boldsymbol{V}$ are full rank matrices, we have $rank\,(\boldsymbol{R_n V}_r) = r$ and $rank\,(\boldsymbol{R_n V}_{m-r}) = m-r$. Thus $span\,\{\boldsymbol{R_n V}_r\}$, which means the space spanned by columns of $\boldsymbol{R_n V}_r$, is disjoint with $span\,\{\boldsymbol{R_n V}_{m-r}\}$. According to the expression for $\boldsymbol{R_y}$, $\boldsymbol{R_n V}_r$ corresponds to the r largest eigenvalues in $\boldsymbol{\Lambda_y}$ and $\boldsymbol{R_n V}_{m-r}$ the remaining $m-r$ small eigenvalues. Thus $span\,\{\boldsymbol{R_n V}_r\}$ can be regarded as the signal dominant subspace and $span\,\{\boldsymbol{R_n V}_{m-r}\}$ the noise subspace(A rigorous proof will be given in section 2.2). Naturally, we consider an estimator for the clean signal projecting the noisy signal onto the signal subspace while eliminating the components in the noise subspace. As the signal subspace is not necessarily orthogonal with the noise subspace, we apply oblique projection onto $span\,\{\boldsymbol{R_n V}_r\}$ along $span\,\{\boldsymbol{R_n V}_{m-r}\}$, which can be calculated as [7]

$$\begin{aligned}\boldsymbol{E} &= \left[\boldsymbol{R_n V}_r, \boldsymbol{O}_{m(m-r)}\right]\left((\boldsymbol{R_n V})^H(\boldsymbol{R_n V})\right)^{-1}(\boldsymbol{R_n V})^H \\ &= \left[\boldsymbol{R_n V}_r, \boldsymbol{O}_{m(m-r)}\right]\boldsymbol{V}^H = \boldsymbol{R_n V}_r\boldsymbol{V}_r^H\end{aligned} \tag{6}$$

where $\boldsymbol{O}_{m(m-r)}$ denotes $m\times(m-r)$ zero matrix. It is easy to verify that $\boldsymbol{E}$ is idempotent (equal to its own square), since $\boldsymbol{V}_r^H\boldsymbol{R_n V}_r = \boldsymbol{I}$ according to (5).

2.2. Properties of the Oblique Projector

Theorem 1: Projection $\hat{\boldsymbol{x}}$=$\boldsymbol{Ey}$=$\boldsymbol{R_n V}_r\boldsymbol{V}_r^H\boldsymbol{y}$ is the maximum likelihood (ML) and minimum variance unbiased (MVUB) estimation of clean signal $\boldsymbol{x}$; Moreover, $span\,\{\boldsymbol{R_n V}_r\}$ is identical to the signal subspace $span\,\{\boldsymbol{H}\}$.

Proof: It is given in [7] that the ML and MVUB estimation for signal $\boldsymbol{x}$ is

$$\hat{\boldsymbol{x}_0} = \boldsymbol{E}_0\boldsymbol{y} = \boldsymbol{H}\left(\boldsymbol{H}^H\boldsymbol{R_n^{-1}}\boldsymbol{H}\right)^{-1}\boldsymbol{H}^H\boldsymbol{R_n^{-1}}\boldsymbol{y} \tag{7}$$

We will show $\boldsymbol{E}$ in (6) and $\boldsymbol{E}_0$ in (7) are exactly the same.

From (2) and (4) we have $\boldsymbol{V}^H(\boldsymbol{H}\cdot\boldsymbol{R_s}\cdot\boldsymbol{H})\,\boldsymbol{V} = \boldsymbol{\Lambda_x}$. Rewrite $\boldsymbol{V}^H\boldsymbol{H}$ as the upper $r\times r$ block matrix and lower $(m-r)\times r$ block matrix, that is $\boldsymbol{V}^H\boldsymbol{H}$=$\begin{pmatrix}\boldsymbol{A}\\ \boldsymbol{B}\end{pmatrix}$ and we have

$$\begin{pmatrix}\boldsymbol{AR_sA}^H & \boldsymbol{AR_sB}^H\\ \boldsymbol{BR_sA}^H & \boldsymbol{BR_sB}^H\end{pmatrix} = \boldsymbol{\Lambda_x}.$$

This indicates that $\boldsymbol{BR_sB}^H$=$\boldsymbol{O}_{(m-r)(m-r)}$; so $\boldsymbol{B}$=$\boldsymbol{O}_{(m-r)r}$ since $\boldsymbol{R_s}$ is positive definite. This implies $\boldsymbol{V}^H\boldsymbol{H}$=$\begin{pmatrix}\boldsymbol{A}\\ \boldsymbol{O}\end{pmatrix}$ and $\boldsymbol{A}$ is full rank. Thus $\boldsymbol{H} = \boldsymbol{V}^{-H}\begin{pmatrix}\boldsymbol{A}\\ \boldsymbol{O}\end{pmatrix} = \boldsymbol{R_n V}\begin{pmatrix}\boldsymbol{A}\\ \boldsymbol{O}\end{pmatrix} = \boldsymbol{R_n V}_r\boldsymbol{A}$, which shows $span\,\{\boldsymbol{R_n V}_r\}$=$span\,\{\boldsymbol{H}\}$. Replace $\boldsymbol{H}$ with $\boldsymbol{R_n V}_r\boldsymbol{A}$ in (7) and it is straightforward to see that $\boldsymbol{E}_0 = \boldsymbol{E}$, which proves that $\hat{\boldsymbol{x}}$=$\boldsymbol{R_n V}_r\boldsymbol{V}_r^H\boldsymbol{y}$ is also the ML and MVUB estimation of $\boldsymbol{x}$.

Theorem 1 reveals that the oblique projector $\boldsymbol{E}$ proposed in section 2.1 from a simultaneous diagonalization perspective, is optimal in the sense of ML and MVUB; the signal subspace can also be identified.

3. OBPAST FOR COMPUTING THE OBLIQUE PROJECTOR

3.1. Subspace Tracking By Unconstrained Minimization

The computation of $\boldsymbol{E}$ in (6) can be carried out directly by solving the problem of simultaneous diagonalization in (5) using EVD (detailed computational procedure can be found in [5]), which is computationally expensive in real time applications. Therefore, we propose obPAST in the following to compute $\boldsymbol{E}$ approximately but more efficiently, with EVD avoided. The algorithm is similar with PAST proposed in [1], while it is able to deal with colored noise directly.

Consider the following minimization problem:

$$\min_{\boldsymbol{W}} J(\boldsymbol{W}) = \mathsf{E}\left(\|\boldsymbol{y} - \boldsymbol{R_n W W}^H\boldsymbol{y}\|^2_{\boldsymbol{R_n^{-1}}}\right) \tag{8}$$

where $\boldsymbol{W}$ is a matrix with dimension $m\times r$. Here square of Mahalanobis distance is employed as $\|\boldsymbol{y}\|^2_{\boldsymbol{R_n^{-1}}} = \boldsymbol{y}^H\boldsymbol{R_n^{-1}}\boldsymbol{y}$.

Theorem 2: $\boldsymbol{W}$ is a stationary point of $J(\boldsymbol{W})$ if and only if $\boldsymbol{W}$=$\boldsymbol{U}_r\,\boldsymbol{Q}$ where $\boldsymbol{U}_r$ contains any r columns of $\boldsymbol{V}$ defined in (5) and $\boldsymbol{Q}$ is an unitary matrix with dimension $r\times r$. At each stationary point, $J(\boldsymbol{W})$ equals the sum of eigenvalues not corresponding with column vectors in $\boldsymbol{U}_r$.

Proof: $J(\boldsymbol{W}) = tr(\boldsymbol{R_n^{-1}R_y}) - 2\cdot tr(\boldsymbol{W}^H\boldsymbol{R_y W})$ $+tr(\boldsymbol{W}^H\boldsymbol{R_y W W}^H\boldsymbol{R_n W})$ where $tr(\cdot)$ denotes trace of the matrix. $\partial J(\boldsymbol{W})/\partial\boldsymbol{W} = 0$ implies $-2\boldsymbol{R_y W} + (\boldsymbol{R_y W W}^H\boldsymbol{R_n} + \boldsymbol{R_n W W}^H\boldsymbol{R_y})\boldsymbol{W} = 0$. Multiplying $\boldsymbol{W}^H$ to both sides yields $\boldsymbol{W}^H\boldsymbol{R_y W}(\boldsymbol{W}^H\boldsymbol{R_n W}-\boldsymbol{I})+(\boldsymbol{W}^H\boldsymbol{R_n W}-\boldsymbol{I})\boldsymbol{W}^H\boldsymbol{R_y W}$=0. As $\boldsymbol{W}^H\boldsymbol{R_n W}-\boldsymbol{I}$ is Hermitian, and $\boldsymbol{W}^H\boldsymbol{R_y W}$ is Hermitian and positive definite, matrix theory guarantees $\boldsymbol{W}^H\boldsymbol{R_n W}-\boldsymbol{I} = 0$. Accordingly, $\partial J(\boldsymbol{W})/\partial\boldsymbol{W} = 0$ is equivalent to $\boldsymbol{R_y W}$=$\boldsymbol{R_n W}(\boldsymbol{W}^H\boldsymbol{R_y W})$. Applying EVD to $\boldsymbol{W}^H\boldsymbol{R_y W}$, we get $\boldsymbol{W}^H\boldsymbol{R_y W}$= $\boldsymbol{Q}^H\boldsymbol{\Sigma}_r\boldsymbol{Q}$. We then attain $\boldsymbol{R_y U}_r = \boldsymbol{R_n U}_r\boldsymbol{\Sigma}_r$ with $\boldsymbol{U}_r$=$\boldsymbol{W}\boldsymbol{Q}^H$. Subsequently we have $\boldsymbol{U}_r^H \boldsymbol{R_n} \boldsymbol{U}_r = \boldsymbol{I}$ and $\boldsymbol{U}_r^H \boldsymbol{R_y} \boldsymbol{U}_r = \boldsymbol{\Sigma}_r$. Compared with (5), we can derive that the r columns of $\boldsymbol{U}_r$ come from the columns of $\boldsymbol{V}$, and the diagonal elements of $\boldsymbol{\Sigma}_r$ come from the diagonal elements of $\boldsymbol{\Lambda_y}$ corresponding to columns of $\boldsymbol{U}_r$. Finally, $J(\boldsymbol{W}) = tr(\boldsymbol{R_n^{-1}} \boldsymbol{R_y}) - tr(\boldsymbol{W}^H \boldsymbol{R_y} \boldsymbol{W}) = tr(\boldsymbol{V} \boldsymbol{\Lambda_y} \boldsymbol{V}^{-1}) - tr(\boldsymbol{Q}^H \boldsymbol{\Sigma}_r \boldsymbol{Q}) = tr(\boldsymbol{\Lambda_y}) - tr(\boldsymbol{\Sigma}_r)$. This means $J(\boldsymbol{W})$ equals the sum of eigenvalues not corresponding to column vectors in $\boldsymbol{U}_r$. This ends the proof.

Theorem 3: All stationary points are saddle points except for $\boldsymbol{U}_r$ containing the r columns of $\boldsymbol{V}$ corresponding with the r dominant eigenvalues. In this case, $\boldsymbol{U}_r = \boldsymbol{V}_r$(thus $\boldsymbol{W} = \boldsymbol{V}_r \boldsymbol{Q}$)and $J(\boldsymbol{W})$ reaches the global minimum.

Proof: Let $\widetilde{H}$ to be the $mr \times mr$ Hessian matrix of $J(\boldsymbol{W})$ with respect to the vectorization of $\boldsymbol{W}$. $\widetilde{H}$ is nonnegative definite when $J(\boldsymbol{W})$ reaches the global minimum. We point out without detailed derivation that this is achieved only if the diagonal elements of $\boldsymbol{\Sigma}_r$ are the r largest eigenvalues in $\boldsymbol{\Lambda_y}$ and $\boldsymbol{U}_r = \boldsymbol{V}_r$. Otherwise, $\widetilde{H}$ is indefinite and the corresponding stationary point is saddle point.

Theorems 2 and 3 implies that when $J(\boldsymbol{W})$ is minimized, the oblique projector can be expressed using $\boldsymbol{R_n}$ and $\boldsymbol{W}$ as $\boldsymbol{E} = \boldsymbol{R_n} \boldsymbol{V}_r \boldsymbol{V}_r^H = \boldsymbol{R_n} \boldsymbol{V}_r \boldsymbol{Q} \boldsymbol{Q}^H \boldsymbol{V}_r^H = \boldsymbol{R_n} \boldsymbol{W} \boldsymbol{W}^H$. Meanwhile, $span\{\boldsymbol{R_n} \boldsymbol{W}\}$ is the signal subspace because $span\{\boldsymbol{R_n} \boldsymbol{V}_r\}$ $= span\{\boldsymbol{R_n} \boldsymbol{V}_r \boldsymbol{Q}\} = span\{\boldsymbol{R_n} \boldsymbol{W}\}$. At the same time, $\boldsymbol{W}$ automatically acquires an orthonormal property, that is, $\boldsymbol{W}^H \boldsymbol{R_n} \boldsymbol{W} = \boldsymbol{I}_r$. Therefore, we can compute the projector $\boldsymbol{E}$ and estimate the signal subspace by solving the minimization problem in (8), without calculating $\boldsymbol{V}_r$ explicitly.

3.2. obPAST Algorithm

To obtain a recursive algorithm for signal subspace tracking, we replace the expectation in (8) with the exponentially weighted sum

$J(\boldsymbol{W}(t)) = \sum_{i=1}^{t} \beta^{t-i} \| \boldsymbol{y}(i) - \boldsymbol{R_n} \boldsymbol{W}(t) \boldsymbol{W}^H(t) \boldsymbol{y}(i) \|^2_{\boldsymbol{R_n^{-1}}}$

Here β is the forgetting factor. As the expectation in (8) and the weighted sum in the above equation are essentially equivalent, the theorems in the above section also hold for $J(\boldsymbol{W}(t))$.

Approximating $\boldsymbol{W}^H(t)\boldsymbol{y}(i)$ with $\boldsymbol{z}(i) = \boldsymbol{W}^H(i-1)\boldsymbol{y}(i)$, we get $J_1(\boldsymbol{W}(t)) = \sum_{i=1}^{t} \beta^{t-i} \| \boldsymbol{y}(i) - \boldsymbol{R_n} \boldsymbol{W}(t) \boldsymbol{z}(i) \|^2_{\boldsymbol{R_n^{-1}}}$

$J_1(\boldsymbol{W}(t))$ is minimized if $\boldsymbol{W}(t) = \boldsymbol{R_n^{-1}} \boldsymbol{C_{yz}}(t) \boldsymbol{C_{zz}^{-1}}(t)$with

$\boldsymbol{C_{yz}}(t) = \sum_{i=1}^{t} \beta^{t-i} \boldsymbol{y}(i) \boldsymbol{z}^H(i) = \beta \boldsymbol{C_{yz}}(t-1) + \boldsymbol{y}(t) \boldsymbol{z}^H(t)$,

$\boldsymbol{C_{zz}}(t) = \sum_{i=1}^{t} \beta^{t-i} \boldsymbol{z}(i) \boldsymbol{z}^H(i) = \beta \boldsymbol{C_{zz}}(t-1) + \boldsymbol{z}(t) \boldsymbol{z}^H(t)$.

Then recursive computation can be applied to $\boldsymbol{W}(t)$ and the obPAST algorithm is finally derived and summarized in Table 1. The signal subspace is given by $span\{\boldsymbol{R_n} \boldsymbol{W}(t)\}$ and the estimator for clean signal is given by $\boldsymbol{R_n} \boldsymbol{W}(t) \boldsymbol{W}^H(t)$. Note from the proof of Theorem 2 that $\boldsymbol{W}^H(t) \boldsymbol{R_n} \boldsymbol{W}(t) = \boldsymbol{I}_r$ when converged and the estimator is unbiased.

If we define $\boldsymbol{S}(t) = \boldsymbol{R_n} \boldsymbol{W}(t)$ and replace the update of $\boldsymbol{W}(t)$ with $\boldsymbol{S}(t)$ in Table 1, we can derive another version of obPAST (obPAST II), which is shown in Table 2. The advantage of obPAST II is that we can directly update the signal subspace given by $span\{\boldsymbol{S}(t)\}$ and the estimator given by $\boldsymbol{S}(t) \boldsymbol{S}^H(t) \boldsymbol{R_n^{-1}}$. Moreover, $\boldsymbol{S}(t)$ satisfies $\boldsymbol{S}^H(t) \boldsymbol{R_n^{-1}} \boldsymbol{S}(t) = \boldsymbol{I}_r$ when converged. Both obPAST I and II need $m^2 + 3mr + \mathsf{O}(r^2)$ operations for each update.

3.3. Discussions

In fact, the signal subspace can also be tracked in colored noise using pre-whitening approach, which requires Cholesky decomposition or SVD for the noise covariance matrix. It can be revealed that obPAST embeds pre-whitening into the update procedure and does not need matrix decomposition.

Obviously, $\boldsymbol{R_n^{-1}}$ is required within each iteration in both obPAST I and II. In real application, it can be calculated in advance using available noise data. Consider a detector for pure noise segments(such as a voice activity detector) is available for the corrupted signal, so that the pure noise data can be identified. We use $\boldsymbol{R_n}(t) = \gamma \boldsymbol{R_n}(t-1) + (1-\gamma) \boldsymbol{n}(t) \boldsymbol{n}^H(t)$ to update $\boldsymbol{R_n}$. Here, γ is the forgetting factor and $\boldsymbol{n}(t)$ is the current noise data. Using matrix inversion lemma, the above equation leads to an efficient update of $\boldsymbol{R_n^{-1}(t)}$ expressed as $\boldsymbol{R_n^{-1}}(t) = \frac{1}{\gamma} \left(\boldsymbol{R_n^{-1}}(t-1) - \frac{\boldsymbol{R_n^{-1}}(t-1) \boldsymbol{n}(t) \boldsymbol{n}^H(t) \boldsymbol{R_n^{-1}}(t-1)}{\gamma/(1-\gamma) + \boldsymbol{n}^H(t) \boldsymbol{R_n^{-1}}(t-1) \boldsymbol{n}(t)} \right)$. Each update of $\boldsymbol{R_n^{-1}}$ needs $2m^2$ operations. In this way, the statistical property of nonstationary noise can be tracked in real time.

Although theoretical analysis gives the result that $\boldsymbol{W}(t)$ acquires an orthonormal property when enough data is processed, real applications do not always guarantee this. Inspired by [2], we can also apply orthonormalization to $\boldsymbol{W}(t)$ in obPAST I as $\boldsymbol{W}(t) := \boldsymbol{W}(t) \left(\boldsymbol{W}^H(t) \boldsymbol{R_n} \boldsymbol{W}(t) \right)^{-1/2}$ after the update for $\boldsymbol{W}(t)$. Recursive implementation of the forced orthonormalization can be derived using the techniques described in [2]. Due to the limitation of space, we omit the detailed procedure here. Similar operation can also be applied to $\boldsymbol{S}(t)$ in obPAST II.

Generalized Estimator for the clean signal in colored noise is given in the speech enhancement scenario [6] by $\hat{\boldsymbol{E}} = \boldsymbol{R_n} \boldsymbol{V}_r \boldsymbol{T} \boldsymbol{V}_r^H$ where $\boldsymbol{T}$ is a diagonal weighting matrix with dimension $r \times r$. We point out that using a deflation version(see [1]) of obPAST, $\boldsymbol{V}_r$(instead of $\boldsymbol{W} = \boldsymbol{V}_r \boldsymbol{Q}$) and $\boldsymbol{\Lambda_y}$ can be updated explicitly, thus $\hat{\boldsymbol{E}}$ can also be calculated recursively.

TABLE 1 obPAST I

Choose $\boldsymbol{P}(0)$ and $\boldsymbol{W}(0)$ suitably
FOR $t = 1, 2, \cdots$ DO
$\boldsymbol{z}_t = \boldsymbol{W}^H(t-1) \boldsymbol{y}_t$
$\boldsymbol{h}_t = \boldsymbol{P}(t-1) \boldsymbol{z}_t$
$\boldsymbol{g}_t = \boldsymbol{h}_t / (\beta + \boldsymbol{z}_t^H \boldsymbol{h}_t)$
$\boldsymbol{P}(t) = (\boldsymbol{P}(t-1) - \boldsymbol{g}_t \boldsymbol{h}_t^H)/\beta$
$\boldsymbol{e}_t = \boldsymbol{R_n^{-1}} \boldsymbol{y}_t - \boldsymbol{W}(t-1) \boldsymbol{z}_t$
$\boldsymbol{W}(t) = \boldsymbol{W}(t-1) + \boldsymbol{e}_t \boldsymbol{g}_t^H$

TABLE 2 obPAST II

Choose $\boldsymbol{P}(0)$ and $\boldsymbol{S}(0)$ suitably
FOR $t = 1, 2, \cdots$ DO
$\boldsymbol{z}_t = \boldsymbol{S}^H(t-1)(\boldsymbol{R}_{\boldsymbol{n}}^{-1}\boldsymbol{y}_t)$
$\boldsymbol{h}_t = \boldsymbol{P}(t-1)\boldsymbol{z}_t$
$\boldsymbol{g}_t = \boldsymbol{h}_t/(\beta + \boldsymbol{z}_t^H\boldsymbol{h}_t)$
$\boldsymbol{P}(t) = (\boldsymbol{P}(t-1) - \boldsymbol{g}_t\boldsymbol{h}_t^H)/\beta$
$\boldsymbol{e}_t = \boldsymbol{y}_t - \boldsymbol{S}(t-1)\boldsymbol{z}_t$
$\boldsymbol{S}(t) = \boldsymbol{S}(t-1) + \boldsymbol{e}_t\boldsymbol{g}_t^H$

4. SIMULATION

The experimental setup includes three different sinusoids impinging an array of nine sensors in different directions of arrival(DOA). Array observation is corrupted by additive colored noise with fixed covariance matrix. We use obPAST II to track the signal subspace. Performance is measured by the subspace distance $D(\boldsymbol{S}(t), \boldsymbol{H}) = \|\boldsymbol{P}_{\boldsymbol{S}(t)} - \boldsymbol{P}_{\boldsymbol{H}}\|_{\boldsymbol{F}}$ [5] with $\boldsymbol{P}_{\boldsymbol{S}(t)} = \boldsymbol{S}(t)\left(\boldsymbol{S}^H(t)\boldsymbol{S}(t)\right)^{-1}\boldsymbol{S}^H(t), \boldsymbol{P}_{\boldsymbol{H}} = \boldsymbol{H}\left(\boldsymbol{H}^H\boldsymbol{H}\right)^{-1}\boldsymbol{H}^H$ and $\|\cdot\|_{\boldsymbol{F}}$ denotes Frobenius norm. The same measure for the direct simultaneous diagonalization approach and PAST is also used for comparison. Furthermore, the asymptotic orthogonality property of $\boldsymbol{S}(t)$ is displayed using measure $\|\boldsymbol{S}^H(t)\boldsymbol{R}_{\boldsymbol{n}}^{-1}\boldsymbol{S}(t) - \boldsymbol{I}_r\|_{\boldsymbol{F}}$.

The noise covariance matrix $\boldsymbol{R}_{\boldsymbol{n}}$ is set as $r_{ij} = 0.01 \cdot \rho \cdot (m-i+1) \cdot (m-j+1)$ with $m = 9$; $\rho = 1$ if $i = j$, else $\rho = 0.9$. Signal scenarios with both smooth and sudden parameter changes are studied. The corresponding spatial frequencies with DOA change from [-0.2 0.3 0.2] to [-0.2 0.2 0.3]in the first 1000 snapshots linearly, then change from [-0.2 0.2 0.3] to [-0.2 0.2 0.4] abruptly in the 1200^{th} snapshot. The SNR level is 6 dB and the forgetting factor is 0.995. The average results of 100 independent runs are shown in Fig. 1, which demonstrate that obPAST leads to smaller deviation from the exact signal subspace than PAST. Furthermore, the performance of obPAST almost coincides with that of the direct simultaneous diagonalization approach after enough snapshots.

5. CONCLUSION

In this paper, a maximum likelihood and minimum variance unbiased estimator for signal corrupted by colored noise is derived using simultaneous diagonalization and oblique projection. Furthermore, obPAST is proposed to update this estimator and track the signal subspace efficiently. The algorithm requires that the noise covariance matrix is full rank and can be estimated. The computational complexity is $\mathsf{O}(m^2)$ where m is the dimension of the noisy signal. This algorithm extends PAST to the colored noise case.

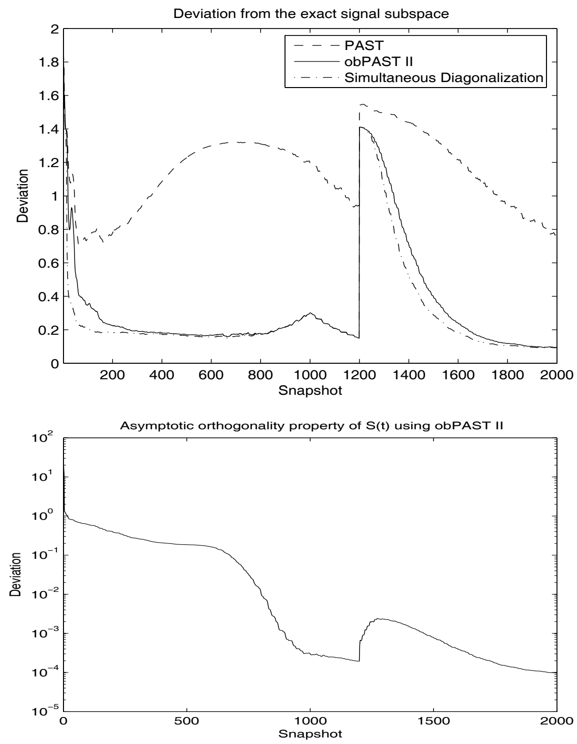

Fig. 1. Subspace tracking performance.

6. ACKNOWLEDGMENT

We would like to thank Ruonan Li at Tsinghua University for his contribution to this paper.

7. REFERENCES

[1] Bin Yang, “Projection approximation subspace tracking,“ *IEEE Trans. Signal Processing*, vol. 43, pp. 95–107, Jan. 1995.

[2] Abed-Meraim, K. and Chkeif, A. and Hua, Y., “Fast orthonormal PAST algorithm,“ *IEEE Signal Processing Letters*, vol. 7, pp. 60–62, Mar. 2000.

[3] Miao, Y. and Hua, Y., “Fast subspace tracking and neural network learning by a novel information criterion,“ *IEEE Trans. Signal Processing*, vol. 46, pp. 1967–1979, Jul. 1998.

[4] Gustafsson, T., “Instrumental variable subspace tracking using projection approximation,“ *IEEE Trans. Signal Processing*, vol. 46, pp. 669–681, Mar. 1998.

[5] Golub GH, Van Loan CF, *Matrix computation*, 2nd ed. Baltimore: The Johns Hopkins University Press, 1989.

[6] Yi Hu and Loizou, P.C., “A generalized subspace approach for enhancing speech corrupted by colored noise,“ *IEEE Trans. Speech and Audio Processing*, vol. 11, pp. 334–341, Jul. 2003.

[7] Behrens, R.T. and Scharf, L.L., “Signal processing applications of oblique projection operators,“ *IEEE Trans. Signal Processing*, vol. 42, pp. 1413–1424, Jun. 1994.