BLIND SUBSPACE-BASED CHANNEL ESTIMATION USING THE EM ALGORITHM

Koji Harada[1] *and Hideaki Sakai*[2]

[1] Agilent Technologies International Japan, Ltd., Kobe, JAPAN; koji_harada@agilent.com
[2] Graduate School of Informatics, Kyoto University, Kyoto, JAPAN; hsakai@i.kyoto-u.ac.jp

ABSTRACT

We propose an application of the Expectation-Maximization (EM) algorithm to the problem of blind estimation of single-input multiple-output (SIMO), finite-impulse-response (FIR) channels. We first assume Gaussian input to formulate an EM-based estimation of the signal subspace of the output covariance matrix. This Gaussian assumption allows us to utilize knowledge from EM-based probabilistic principle component analysis (P-PCA). Next, we show that the equilibrium point of the EM iteration equations is reached without the Gaussian assumption, which suggests usage of non-Gaussian communication input signals. The estimated signal subspace is then utilized to identify the channels. In principle, the proposed method yields the same channel estimates as the widely-known subspace method, but is computationally more efficient. In addition, unlike typical EM applications, the proposed scheme is free from cumbersome parameter initialization issue, which greatly increases flexibility of the proposed scheme.

Index Terms— Blind channel estimation, subspace method, Expectation-Maximization (EM) algorithm, probabilistic principle component analysis (P-PCA)

1. INTRODUCTION

A number of algorithms have been developed to blindly estimate SIMO FIR channels. Among those based on second-order statistics, the popular subspace method [1] has many favorable properties, such as excellent robustness to noise and compatibility with colored input signals. However, it involves evaluation of the output covariance matrix followed by eigen-decomposition of it, hence, is not computationally very efficient when the channel order is large or there are many output observations.

In this paper, a new subspace-based method utilizing the Expectation-Maximization (EM) algorithm is proposed. Here, the *signal subspace* of the output covariance matrix is estimated using the EM algorithm. Then the FIR channels are identified by fitting the *filter subspace* onto the *signal subspace*. This EM-based method retains all the favorable features of the subspace method mentioned above, with additional benefit being computationally more efficient because it avoids having to evaluate the output covariance matrix. Because of the local convergence nature of the EM algorithm, the estimated subspace basis has dependency on initialization of the EM iteration. However, all these bases do span the same subspace, which explains why the final channel estimates of the proposed scheme do not depend on EM initialization.

If we assume isotropic Gaussian input and view it as the latent variable, the probabilistic principle component analysis (P-PCA) [3] technique is directly applicable to estimate to the *signal subspace*. However, our goal is to use possibly non-Gaussian communication signals (white or colored) as the input. It is shown that the equilibrium point of the EM iteration equations is reached without the Gaussian assumption, which suggests usage of non-Gaussian communication input signals.

To demonstrate the above points, numerical simulation is performed with QPSK and 16QAM signals (white or colored) with randomly generated SIMO FIR channels while the E-step initialization for the *signal subspace* is held unchanged. The proposed scheme showed no miss-convergence and good agreement with the standard subspace method.

Most prior work on applying EM to blind channel estimation problems can roughly be classified into *semi-blind* estimation [4], or joint channel and sequence estimation [5] which typically has dependency on symbol constellation. Application to blind SIMO FIR estimation does not seem to be previously proposed, to the best of the authors' knowledge.

2. THE SIMO FIR CHANNEL MODEL

In the $1\times M$ SIMO FIR channel model, the white or colored input $s(n)\in\mathbb{C}$ at time n and the output $\mathbf{x}(n)\in\mathbb{C}^M$ have a convolutional relationship

$$\mathbf{x}(n)=\sum_{l=0}^{L}\mathbf{h}(l)s(n-l)+\mathbf{v}(n),$$

where $\mathbf{h}(l)\in\mathbb{C}^M$, $l\in[0,L]$, is the channel coefficient vector with FIR channel order L, and $\mathbf{v}(n)\in\mathbb{C}^M$ is the observation noise, uncorrelated with $\mathbf{x}(n)$, that follows the M-variate isotropic Gaussian distribution with variance σ_v^2. By stacking successive D outputs as $\mathbf{x}=[\mathbf{x}(n)^T, \mathbf{x}(n\text{-}1)^T, \cdots, \mathbf{x}(n\text{-}D\text{+}1)^T]^T$, with $[\cdot]^T$ denoting transpose, we have

$$\mathbf{x}=\mathcal{H}\mathbf{s}+\mathbf{v}, \tag{1}$$

where $\mathbf{s}=[s(n),s(n-1),\cdots,s(n-D-L+1)]^T$

$\mathbf{v}=[\mathbf{v}(n)^T,\mathbf{v}(n-1)^T,\cdots,\mathbf{v}(n-D+1)^T]^T$

and $\mathcal{H}$ is an $MD\times(L+D)$ filtering matrix constructed as

$$\mathcal{H}=\begin{bmatrix}\mathbf{h}(0) & \cdots & \mathbf{h}(L) & & \\ & \ddots & & \ddots & \\ & & \mathbf{h}(0) & \cdots & \mathbf{h}(L)\end{bmatrix}. \tag{2}$$

The output covariance matrix can then be written as

$$\mathbf{R} = E[\mathbf{x}\mathbf{x}^H] = \mathcal{H}\,E[\mathbf{s}\mathbf{s}^H]\mathcal{H}^H + \sigma_v^2\mathbf{I}_{MD},$$

where $E[\cdot]$ denotes mathematical expectation, $[\cdot]^H$ is complex transpose, and $\mathbf{I}_K$ is the identity matrix of size K. The objective of blind multi-channel estimation is to identify

$$\mathbf{h} = [\mathbf{h}(0)^T, \mathbf{h}(1)^T, \cdots, \mathbf{h}(L)^T]^T,$$

under the assumption that only $\mathbf{x}(n)$ is observable. Full column rank, $L+D$, is assumed on $\mathcal{H}$. This results from the identifiability of $\mathbf{h}$ that the L-th order polynomials constructed from the rows of $[\mathbf{h}(0), \mathbf{h}(1), \cdots, \mathbf{h}(L)]$ are coprime [1].

3. SIGNAL SUBSPACE AND P-PCA

In this section, we show equivalence between the *signal subspace* of the output covariance matrix $\mathbf{R}$ and the *principle subspace* in the P-PCA framework, assuming that $\mathbf{x}$ is an isotropic Gaussian.

3.1. Signal Subspace

The output covariance matrix $\mathbf{R}$ has the eigen-decomposition $\mathbf{R} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$, where $\mathbf{U}$ is a unitary matrix whose columns are the eigenvectors of $\mathbf{R}$, and $\mathbf{\Lambda}$ is a diagonal matrix whose diagonal elements are the eigenvalues of $\mathbf{R}$. We can further decompose it into two terms

$$\mathbf{R} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H = \mathbf{U}_s\mathbf{\Lambda}_s\mathbf{U}_s^H + \mathbf{U}_n\mathbf{\Lambda}_n\mathbf{U}_n^H,$$

using $\mathbf{U} = [\mathbf{U}_s \mid \mathbf{U}_n]$ and $\mathbf{\Lambda} = \left[\begin{array}{c|c} \mathbf{\Lambda}_s & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{\Lambda}_n \end{array}\right]$,

where the diagonal elements of $\mathbf{\Lambda}_s$ are the largest $L+D$ eigenvalues, those of $\mathbf{\Lambda}_n$ are the rest of the eigenvalues, and the columns of $\mathbf{U}_s$ and $\mathbf{U}_n$ are the eigenvectors corresponding to $\mathbf{\Lambda}_s$ and $\mathbf{\Lambda}_n$ respectively. Then the *signal subspace* of $\mathbf{R}$ is defined by the subspace spanned by the columns of $\mathbf{U}_s$ denoted by $\mathrm{Span}\{\mathbf{U}_s\}$. In addition, it is straightforward to show that $\mathcal{H}$ and span the same subspace, that is

$$\mathrm{Span}\{\mathbf{U}_s\} = \mathrm{Span}\{\mathcal{H}\}. \tag{3}$$

3.2. Probabilistic Principle Component Analysis (P-PCA)

In P-PCA [3] an observed data point $\mathbf{y} \in \mathbb{C}^d$ is related to a latent variable $\mathbf{z} \in \mathbb{C}^k$, $k < d$, through the linear mapping

$$\mathbf{y} = \mathbf{W}_0\mathbf{z} + \boldsymbol{\mu} + \boldsymbol{\varepsilon}, \tag{4}$$

where $\mathbf{W}_0 \in \mathbb{C}^{d\times k}$ is a matrix whose columns span the *principle subspace* of $\mathbf{y}$, $\boldsymbol{\mu}$ is the data mean, and $\boldsymbol{\varepsilon}$ is the observation noise. Both the latent variable and the noise are assumed to follow zero mean isotropic Gaussian distributions of the form $p(\mathbf{z}) = \mathcal{N}(\mathbf{z} \mid \mathbf{0}, \mathbf{I}_k)$ and $p(\boldsymbol{\varepsilon}) = \mathcal{N}(\boldsymbol{\varepsilon} \mid \mathbf{0}, \sigma_\varepsilon^2\mathbf{I}_d)$. The marginal likelihood can be obtained by integrating the conditional likelihood over the latent variable, which has the form

$$\begin{aligned} p(\mathbf{y} \mid \mathbf{W}, \boldsymbol{\mu}, \sigma_\varepsilon^2) &= \int p(\mathbf{y} \mid \mathbf{z}, \mathbf{W}, \boldsymbol{\mu}, \sigma_\varepsilon^2)\, p(\mathbf{z})\, d\mathbf{z} \\ &= \mathcal{N}(\mathbf{y} \mid \boldsymbol{\mu}, \mathbf{W}\mathbf{W}^H + \sigma_\varepsilon^2\mathbf{I}_d), \end{aligned} \tag{5}$$

where $\mathbf{W}$ is an estimate of $\mathbf{W}_0$. This marginal distribution is again Gaussian because the linear mapping (4) corresponds to a linear-Gaussian model. For a data set $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_N] \in \mathbb{C}^{d\times N}$ with N observations, the marginal likelihood is simply a product of (5) over the all data points because these are assumed to be independent, yielding log likelihood of the form

$$\ln p(\mathbf{Y} \mid \mathbf{W}, \boldsymbol{\mu}, \sigma_\varepsilon^2) = \sum_{n=1}^{N} \ln \mathcal{N}(\mathbf{y}_n \mid \boldsymbol{\mu}, \mathbf{W}\mathbf{W}^H + \sigma_\varepsilon^2\mathbf{I}_d). \tag{6}$$

Comparing (1) and (4) we see that the columns of $\mathcal{H}$ actually span the principle subspace of $\{\mathbf{x}\}$ if $\mathbf{s}$ is viewed as the latent variable and assumed to be an isotropic Gaussian. Therefore, the maximum-likelihood (ML) solution of $\mathcal{H}$ can be obtained through P-PCA by maximizing (6) with respect to $\mathbf{W} = \mathcal{H}$, and with $\boldsymbol{\mu}$ set to zero.

4. EM ALGORITHM FOR P-PCA

Although direct maximization of (6) with regard to $\mathbf{W}$ has a closed-form solution, an EM algorithm [2] can still be derived for iterative maximization [3]. This avoids explicit calculation of the data covariance matrix and its eigen-decomposition, hence, can lead to substantial computation savings especially if $k \ll d$.

The EM algorithm in general alternates the following two steps from some initial (old) parameter estimates $\boldsymbol{\theta}_{old}$ until convergence is achieved.

E-step: $\quad Q(\boldsymbol{\theta}, \boldsymbol{\theta}_{old}) = E_z[\ln p(\mathbf{Y}, \mathbf{Z} \mid \boldsymbol{\theta}) \mid \boldsymbol{\theta}_{old}]$

M-step: $\quad \boldsymbol{\theta}_{old} \leftarrow \boldsymbol{\theta}_{new} = \arg\max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}, \boldsymbol{\theta}_{old})$

To derive the EM algorithm for estimation of the signal subspace using P-PCA, the parameter vector $\boldsymbol{\theta}$ is replaced by ($\boldsymbol{\mu}=\mathbf{0}$, $\mathbf{W}$, σ_v^2), also, $\mathbf{y}$, $\mathbf{z}$ and $\boldsymbol{\varepsilon}$ are replaced by $\mathbf{x}$, $\mathbf{s}$ and $\mathbf{v}$. Here, we can directly apply results from [3]. Maximization with respect to $\mathbf{W}$ and σ_v^2 yields the M-step equations

$$\mathbf{W}_{new} = \left[\sum_{n=1}^{N} \mathbf{x}_n E[\mathbf{s}_n \mid \mathbf{x}_n]^H\right]\left[\sum_{n=1}^{N} E[\mathbf{s}_n\mathbf{s}_n^H \mid \mathbf{x}_n]\right]^{-1}, \tag{7}$$

$$\sigma_{v,new}^2 = \frac{1}{ND}\sum_{n=1}^{N}\left\{\|\mathbf{x}_n\|^2 - 2E[\mathbf{s}_n \mid \mathbf{x}_n]^H\mathbf{W}^H\mathbf{x}_n + \mathrm{Tr}\left(E[\mathbf{s}_n\mathbf{s}_n^H \mid \mathbf{x}_n]\mathbf{W}^H\mathbf{W}\right)\right\} \tag{8}$$

where

$$E[\mathbf{s}_n \mid \mathbf{x}_n] = (\mathbf{W}^H\mathbf{W} + \sigma_v^2\mathbf{I}_k)^{-1}\mathbf{W}^H\mathbf{x}_n$$

$$E[\mathbf{s}_n\mathbf{s}_n^T \mid \mathbf{x}_n] = \sigma_v^2(\mathbf{W}^H\mathbf{W} + \sigma_v^2\mathbf{I}_k)^{-1} + E[\mathbf{s}_n \mid \mathbf{x}_n]E[\mathbf{s}_n \mid \mathbf{x}_n]^H,$$

are calculated in the E-step with parameter values fixed to their old values. It should also be noted that, for zero noise limit ($\sigma_v^2 \to 0$) we can still obtain a valid EM-like algorithm [6] given by

E-step: $\quad \mathbf{\Omega} = E[\mathbf{S}] = (\mathbf{W}^H\mathbf{W})^{-1}\mathbf{W}^H\mathbf{X}$ (9)

M-step: $\quad \mathbf{W}_{new} = \mathbf{X}\mathbf{\Omega}^H(\mathbf{\Omega}\mathbf{\Omega}^H)^{-1}$ (10)

where $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N] \in \mathbb{C}^{MD\times N}$ and $E[\mathbf{S}] = [\,E[\mathbf{s}_1|\mathbf{x}_1], E[\mathbf{s}_2|\mathbf{x}_2], \cdots, E[\mathbf{s}_N|\mathbf{x}_N]\,] \in \mathbb{C}^{(L+D)\times N}$.

4.1. EM Equilibrium Point

By substituting (9) into (10), we have the recursion

$$\mathbf{W}_{new} = \mathbf{X}\mathbf{X}^H\mathbf{W}\left(\mathbf{W}^H\mathbf{X}\mathbf{X}^H\mathbf{W}\right)^{-1}\mathbf{W}^H\mathbf{W}.$$

For sufficiently large N, $(1/N)\mathbf{X}\mathbf{X}^H = (1/N)\sum_{n=1}^{N}\mathbf{x}_n\mathbf{x}_n^H$ becomes $E[\mathbf{x}_n\mathbf{x}_n^H] = \mathbf{W}_0\mathbf{R}_s\mathbf{W}_0^H$ by the law of large number, with $\mathbf{R}_s = E[\mathbf{s}_n\mathbf{s}_n^H]$ and $\sigma_v^2 \to 0$. Thus the recursion can be rewritten as

$$\mathbf{W}_{new} = \mathbf{W}_0\mathbf{R}_s\mathbf{W}_0^H\mathbf{W}\left(\mathbf{W}^H\mathbf{W}_0\mathbf{R}_s\mathbf{W}_0^H\mathbf{W}\right)^{-1}\mathbf{W}^H\mathbf{W}. \quad (11)$$

To find the equilibrium point of (11), we express $\mathbf{W}$ in the following form

$$\mathbf{W} = \mathbf{W}_0\mathbf{U} + \Delta\mathbf{W}, \quad (12)$$

where $\mathbf{U}$ is an arbitrary unitary matrix representing the rotational ambiguity inherent in P-PCA. Substituting (12) into (11) and, to focus on local stability, ignoring the second-order terms of $\Delta\mathbf{W}$ yields

$$\mathbf{W}_{new} = \mathbf{W}_0\mathbf{U} + \mathbf{W}_0(\mathbf{W}_0^H\mathbf{W}_0)^{-1}\mathbf{W}_0^H\Delta\mathbf{W}. \quad (13)$$

Here the second term is the orthogonal projection of columns of $\Delta\mathbf{W}$ onto the space spanned by columns of $\mathbf{W}_0$, Span$\{\mathbf{W}_0\}$. Therefore, the EM equilibrium point (or rather, space) can be written using an arbitrary $(L+D)\times(L+D)$ matrix $\mathbf{A}$ as

$$\mathbf{W}_{EQ} = \mathbf{W}_0\mathbf{A}, \ \ \mathbf{A} \in \mathbb{C}^{(L+D)\times(L+D)}. \quad (14)$$

This has an interpretation that the EM iteration works so as to bring columns of $\mathbf{W}$ into Span$\{\mathbf{W}_0\}$.

Also, we note that the EM iteration equations (9) into (10) are derived based on the Gaussian assumption on the input. Nevertheless, the equilibrium point (14) is reached by non-Gaussian, colored input signals too. This actually agrees with simulation results we will soon examine.

5. SIGNAL SUBSPACE FITTING

As can be seen from the EM equilibrium point (14) (and (13)), the estimate of $\mathbf{W}$ is subject to latent space rotation $\mathbf{U}$ and uncertainty within Span$\{\mathbf{W}_0\}$. In other words, the likelihood function (6) is multi-modal, causing estimate of $\mathbf{W}$ be dependent on its EM initialization. In order to avoid this initialization issue, we perform subspace fitting so that the P-PCA estimate of $\mathbf{W}$, denoted by $\mathbf{W}_{PCA}$, and the filtering matrix $\mathcal{H}$ in (2) span the same signal subspace. This can be achieved by maximizing the following quantity

$$\text{Maximize} \ \ J(\mathbf{h}) = \sum_i \left\| \mathcal{H}^H \tilde{\mathbf{w}}_{PCA,i} \right\|^2$$

as suggested in [1], where $\tilde{\mathbf{w}}_{PCA,i}$ denotes the i-th orthonormal basis of $\mathbf{W}_{PCA}$, which is obtained using, for example, the Gram-Schmidt orthogonalization procedure. We now apply the structural relation described in *Lemma* 1 in [1] to the above equation to get

$$\sum_i \left\| \mathcal{H}^H \tilde{\mathbf{w}}_{PCA,i} \right\|^2 = \sum_i \left\| \mathcal{W}_{PCA,i}^H \mathbf{h} \right\|^2$$

where $\mathcal{W}_{PCA,i} = \begin{bmatrix} \boldsymbol{\omega}_{i,0}^T & \cdots & \boldsymbol{\omega}_{i,D-1}^T & & \\ & \ddots & & \ddots & \\ & & \boldsymbol{\omega}_{i,0}^T & \cdots & \boldsymbol{\omega}_{i,D-1}^T \end{bmatrix}$

$\tilde{\mathbf{w}}_{PCA,i} = [\boldsymbol{\omega}_{i,0}^T, \boldsymbol{\omega}_{i,1}^T, \cdots, \boldsymbol{\omega}_{i,D-1}^T]^T$.

Therefore, the maximization problem is equivalent to

$$\hat{\mathbf{h}} = \arg\max_{\mathbf{h}} J(\mathbf{h}) = \arg\max_{\mathbf{h}} \mathbf{h}^H \left(\sum_i \mathcal{W}_{PCA,i}\mathcal{W}_{PCA,i}^H \right) \mathbf{h},$$

under the constraint $\|\mathbf{h}\|^2 = 1$, which is easily implemented by looking for the eigenvector corresponding to the largest eigenvalue of $\sum_i \mathcal{W}_{PCA,i}\mathcal{W}_{PCA,i}^H$.

6. COMPUTATIONAL COMPLEXITY

We focus on computational complexity in computation of the signal subspace because it is the computationally most demanding step. Direct eigen-decomposition involves explicit evaluation of the output covariance matrix $O(Nd^2) = O(N(MD)^2)$. This typically exceeds eigen-decomposition $O(d^3) = O((MD)^3)$ because, in most practical cases, there are more observations than its dimensionality, that is, $d < N$, or $MD < N$. The EM-based P-PCA does not construct the covariance matrix explicitly. Instead the most demanding steps are those involving sums over the data set that are $O(Ndk) = O(NMD(L+D))$. For large d, and $k \ll d$, this can be a significant saving, and can offset the iterative nature of the EM algorithm. In the SIMO FIR cases, the EM-based estimation is very efficient for cases where M is large, and $D>L$. Computation of subspace fitting involves eigen-decomposition $O(M^3(L+1)^3)$, which typically is much smaller than computation of the signal subspace involving $O(Ndk) = O(NMD(L+D))$. Matrix inverses that appear in (9) and (10) are taken on Toeplitz matrices, therefore algorithms involving only $O((L+D)^2)$ are available.

7. SIMULATIONS

Numerical simulations were performed with SIMO FIR channels of L=7 (8 taps), M=4 and D=15. To stress that the local convergence nature of the EM-based P-PCA is avoided, the channel coefficients $\mathbf{h} \in \mathbb{C}^{M(L+1)}$ are generated randomly by drawing a sample from complex normal Gaussian distribution every time, while the E-step initialization $\mathbf{W}_{init} \in \mathbb{C}^{MD\times(L+M)}$ is held constant at

$$\mathbf{W}_{init} = \begin{bmatrix} \mathbf{w}_0 & \cdots & \mathbf{w}_L & & \\ & \ddots & & \ddots & \\ & & \mathbf{w}_0 & \cdots & \mathbf{w}_L \end{bmatrix} \quad (15)$$

where

$$[\mathbf{w}_0, \cdots, \mathbf{w}_L] = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \end{bmatrix}$$

The number of input symbols used to generate one estimate is N=1000. The statistical metrics are either mean, sum of mean and the standard deviation, or the maximum of the normalized squared estimation error (NSE) over 50 estimates, where the normalized squared error for the i-th estimate given by

$$NSE_i = \min_{\beta_i} \left\| \beta_i\hat{\mathbf{h}}_i - \mathbf{h}_{true} \right\|^2 \Big/ \left\| \mathbf{h}_{true} \right\|^2$$

where β_i is a complex constant that compensates for the

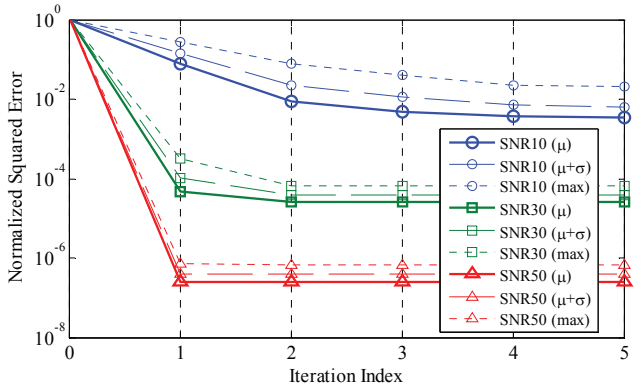

Fig.1 Channel estimation error versus EM iterations for various SNR settings. Channel coefficients are randomly generated, and the E-step initialization is fixed. White QPSK input.

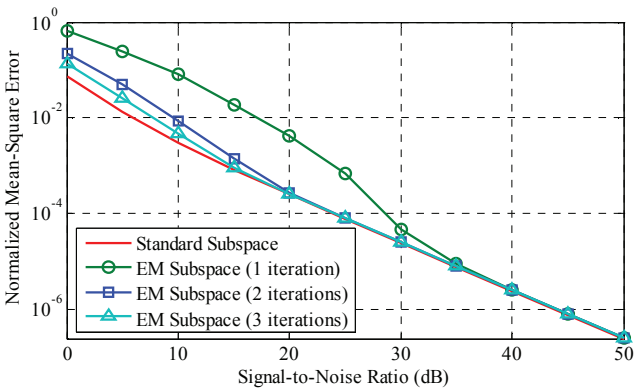

Fig.2 Convergence toward the standard subspace method. White QPSK input into randomly generated channel.

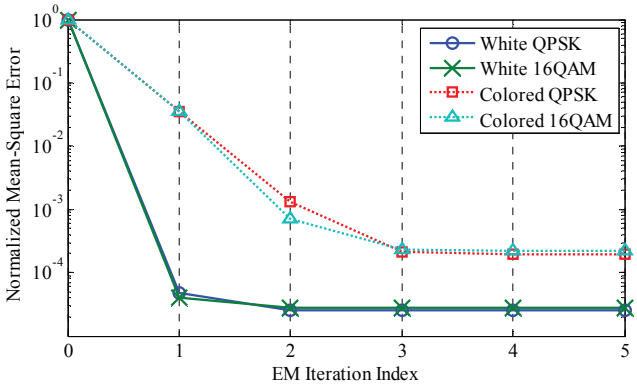

Fig.3 Channel estimation error versus EM iterations for various types of inputs at SNR=30dB.

inherent ambiguity of the *i*-th estimate, associated with blind channel estimation in general. The E-step and M-step equations used in the simulations are zero-noise versions, (9) and (10), because no noticeable performance degradation was observed compared with the noise-included versions, (7) and (8).

Fig.1 shows channel estimation error in function of EM iterations for white QPSK input. We see that EM converges within a few iterations, especially in high SNR settings. Also we see that the maximum NSE lies within a reasonable range from the mean, which verifies successful estimation for every randomly generated channel with a single E-step initialization $\mathbf{W}_{init}$ in (15).

Fig.2 illustrates the mean of NSE versus SNR for white QPSK input. From this figure we see that the estimates by EM-based P-PCA approaches to those by the standard subspace method.

Fig.3 demonstrates channel estimation error in function of the number of EM iterations for various types of inputs, white or colored, and QPSK or 16QAM modulated. We see that the difference in modulation makes only slight difference, as we predicted in Section 4.1. The FIR coefficients of (0.7071, 0.7071) were used as a coloring filter. Although the proposed algorithm did converge, this colorness brought slower convergence and increased error.

7. CONCLUSIONS

A new implementation of blind estimation of SIMO FIR channels has been proposed. This EM-based implementation yields the same channel estimates as the widely-known subspace method, but is computationally more efficient. Unlike typical EM applications, the proposed scheme is free from the cumbersome parameter initialization issue. It is shown that the proposed scheme is applicable to non-Gaussian, colored input signals too.

8. REFERENCES

[1] E. Moulines, P. Duhamel, J.-F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Processing*, vol. 43, pp. 516–525, Feb. 1995.

[2] A. P. Dempster, N. M. Laird and D. B. Rubin, "Maximum Likelihood from Incomplete Data Via the EM Algorithm", *Journal of the Royal Statistical Society*, vol. 39, no. 1, pp 1-38, 1977.

[3] M. E. Tipping and C. M. Bishop. "Probabilistic principal component analysis," *Journal of the Royal Statistical Society*, Series B, 21(3):611–622, 1999.

[4] A. Dogandzic, W. Mo, and Z. Wang, "Semi-blind SIMO flat-fading channel estimation in unknown spatially correlated noise using the EM algorithm," *IEEE Trans. Signal Processing*, vol. 52, pp. 1791–1797, June 2004.

[5] H. Nguyen, and B. C. Levy, "The Expectation-Maximization Viterbi Algorithm for Blind Adaptive Channel Equalization," *IEEE Trans. Communications*, vol. 53, pp. 1671-1678, Oct. 2005.

[6] S. Roweis, "EM Algorithms for PCA and SPCA", *Advances in Neural Information Processing Systems*, vol. 10, The MIT Press, 1998.

[7] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, New York, 2006.

[8] Neal, R. M. & Hinton, G. E. "A view of the EM algorithm that justifies incremental and other variants," *Learning in Graphical Models,* Kluwer Academic Publishers, pp. 355–368, 1998.