14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4-8, 2006, copyright by EURASIP

ICA-Based Semi-Blind Spatial Equalization for MIMO Systems

Zhiguo Ding, T. Ratnarajah, Colin Cowan and Yu Gong
ECIT, Queen's University of Belfast, Queen's Road, Queen's Island
Belfast, BT3 9DT, UK
e-mail: Z.Ding@ecit.qub.ac.uk

***Abstract*— Blind equalization is one challenging problem for multiple-input multiple-output systems, to which independent component analysis (ICA) is applicable. However direct application of ICA could yield low convergence speed and poor performance. In this paper we propose two semi-blind ICA-based algorithms, which incorporate information both from the training and unknown sequences. During each iterative/adaptive step, the training information is utilized to supervise the unconstrained blind ICA-based measure. Simulation results show that the two proposed semi-blind approaches can outperform both the training-based and conventional ICA method. Furthermore we report a special case of MIMO systems which does not require the algorithm of source separation, whose proof is also provided.**

I. Introduction

Multiple-input multiple-output (MIMO) has recently become an area of intense development in the wireless industry. Since multiple transmitted sequences are sharing the same radio resources, MIMO systems are operated in the presence of co-channel interference as well as intersymbol interference. To fully explore the performance gain promised by MIMO, robust techniques for spatial and temporal equalization are required at the receiver.

In this paper, we concentrate on how to achieve spatial equalization and assume that the frequency selective MIMO channels have been reduced to multiple flat fading channels. As proved in [1] second order statistics (SOS) are sufficient to de-convolute frequency selective MIMO channels to an instantaneous mixture. Recall that methods using high order statistics (HOS) have slow convergence speed and high computational complexity. So a reasonable strategy of MIMO equalization is to first achieve the temporal equalization by using SOS. And in the second step, the spatial equalization is completed by utilizing HOS. While a lot of work has been done for deconvolution [2], [3], research in the area of spatial equalization for MIMO systems is still ongoing.

When channel information is not available at the receiver, training based methods [4] are traditionally proposed to achieve reliable detection. As the training will consume either time or frequency resources, these training based methods are bandwidth inefficient, and hence blind approaches which only rely on the statistics of observations become preferable. Interestingly the data model of blind spatial equalization for MIMO systems is the same as the one in a classical area, known as blind source separation [5]–[8]. Specially we are interested in the approach of independent component analysis (ICA), which utilizes the measure of non-Gaussianity to achieve blind source separation. However there is some residual order and phase ambiguity in the result of ICA which requires the training-based methods to be concatenated.

Instead of considering the training and blind information separately, the idea of so-called semi-blind has received a lot of attention recently [9]. By incorporating information both from the training and unknown sequence, semi-blind approaches are able to achieve better performance than both training and blind methods. Furthermore it can also overcome the shortcoming of blind methods which typically require quite a large amount of data to converge. In [10] a semi-blind method using the prior information of constant modulus was proposed for single-input single-output systems. Semi-blind approaches for MIMO systems with space-time block coding can be found in [11], [12] which utilize the redundant information of space time coding and hence are not applicable to general MIMO systems.

In this paper we consider semi-blind spatial equalization for general MIMO systems. By using ICA, a cost function to yield the desired equalizer is constructed to measure the non-Gaussianity of the output of the spatial equalizer. Then a training based cost function is incorporated to supervise the ICA cost function and to find the desired solution more quickly and accurately. Two algorithms are proposed to solve the combined measure, the semi-blind iterative and adaptive algorithm. The iterative method has a fast convergence speed and better performance. The adaptive algorithm can avoid the computation of the matrix inverse and also be able to track time varying systems. Simulation shows that both proposed semi-blind algorithms can achieve better performance than the pure training and conventional ICA method. Also we will report a special case of MIMO systems which does not require the algorithm of source separation with a formal proof provided.

II. Data Model

Consider a discrete-time system model with $M > 1$ transmit antennas and $K \geq M$ receive antennas. We assume that the channel is time invariant during one data block and frequency flat fading. Let N denote the length of one block. During the time interval $[t_i + nT, t_i + (n+1)T]$, the kth receive antenna receive a symbol which is denoted as $r_k(n)$. Hence we obtain

$$r_k(n) = \sum_{m=1}^{M} h_{k,m} s_m(n) + w_k(n), \tag{1}$$

This work was supported by the UK Engineering and Physical Sciences Research Council under grant number EP/C004132/1.

where $s_m(n), n = 1, \ldots, N$, is the nth symbol transmitted from the mth antenna, $h_{k,m}$ denotes the coefficient of the flat fading path from the mth transmit antenna to the kth receive antenna.

To make ICA applicable, we assume that the multiple transmitted sequences are independent to each other and $s_m(n)$ is a random variable with equal probability chosen from the constellation which is known at the receiver. Also assume that the noise is a white process, having $E\{w_k(n)w_i^*(j)\} = \sigma^2\delta_{k,i}\delta_{n,j}$, where $\delta_{i,j}$ is the Kronecker delta function.

Let $\mathbf{s}(n) = \begin{bmatrix} s_1(n) & \cdots & s_M(n) \end{bmatrix}^T$ denote the vector of M transmitted symbols at the nth time interval. Stacking over K receive antennas, we obtain

$$\mathbf{r}(n) = \mathbf{H}\mathbf{s}(n) + \mathbf{w}(n), \tag{2}$$

where $\mathbf{r}(n) = \begin{bmatrix} r_1(n) & \cdots & r_K(n) \end{bmatrix}^T$, $\mathbf{H}$ is the $K \times M$ block-stationary channel matrix with its i,jth element equal to $h_{i,j}$ and $\mathbf{w}(n)$ is constructed from $w_k(n)$ similarly to $\mathbf{r}(n)$.

As a common strategy for most ICA algorithms, the received symbols are first prewhitened

$$\mathbf{y}(n) = \mathbf{P}\mathbf{r}(n), \tag{3}$$

where $\mathbf{P}$ is a $K \times K$ prewhitening matrix. One choice of the prewhitening matrix is $\mathbf{P} = \mathbf{U}\mathbf{\Lambda}^{-1/2}\mathbf{U}^H$, where $\mathbf{U}$ and $\mathbf{\Lambda}$ are obtained from the eigenvalue decomposition of the covariance matrix of the observations $E\{\mathbf{r}(n)\mathbf{r}(n)^H\} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$. As pointed out in [7], the strategy of prewhitening is to reduce the number of parameters to be estimated and the degree of freedom of the mixing matrix. In fact the power of the prewhitening process can be demonstrated from the following claim:

Claim 1 : *Consider a MIMO system with two antennas at each end $M = K = 2$ and the transmitted symbols belong to the BPSK constellation. The original transmitted symbols can be separated by only using a prewhitening filter whose output contain only order and phase ambiguity.*

This means that in some communication scenarios, the prewhitening filter itself can accomplish blind source separation and hence there is no need for another source separation algorithm concatenated to the prewhitening filter. The proof of the claim 1 is given in the Appendix.

III. SEMI-BLIND ICA CRITERION

The aim of the spatial equalizer is to suppress both co-channel interference and additive noise, and recover the transmitted symbols as accurately as possible. Let $\mathbf{g}_m \triangleq \begin{bmatrix} g_{1,m} & \cdots & g_{K,m} \end{bmatrix}^T$ denote the coefficients of the equalizer for the symbols from the mth transmit antenna. Hence we obtain

$$\hat{s}_m(n) = \mathbf{g}_m^H\mathbf{y}(n) = \mathbf{g}_m^H\mathbf{P}\mathbf{H}\mathbf{s}(n), \tag{4}$$

which should be equal to the $s_m(n)$ ideally. Although we initially neglect additive noise for clarity in developing the algorithm, we later assess its performance in environments with additive noise.

A lot of criteria have been proposed to estimate the coefficients of the equalizer. In this paper, we will concentrate on two types, the ICA-based and training-based criteria.

A. ICA based criterion

Recall that $\hat{s}_m(n)$ in (4) can be treated as a linear combination of the original symbols. The basic idea of ICA is that a linear combination of multiple variables should be more gaussian than the original variables, unless the outputs of the filter are the same as the original ones. So the aim of ICA algorithms is to find a measure of so-called non-Gaussianity. In [7], [13] one such criterion is proposed as

$$J(\mathbf{g}_m) = E\{f(|\hat{s}_m(n)|^2)\}, \tag{5}$$

where $f(\cdot)$ is the nonquadratic function defined in [7]. Many classical measures of non-Gaussianity can fit into this general framework. For example, standard fourth order cumulant is one classical measure of non-Gaussianity, which is defined as

$$K(\hat{s}_m(n)) = E\{|\hat{s}_m(n)|^4\} - 2E\{|\hat{s}_m(n)|^2\} - |E\{(\hat{s}_m(n))^2\}|^2.$$

Following from the fact that $\mathbf{y}(n)$ has been prewhitened, we can obtain

$$K(\hat{s}_m(n)) = E\{|\hat{s}_m(n)|^4\} - 2. \tag{6}$$

So for kurtosis, the nonquadratic function can be written as

$$f(x) = x^2 - 2. \tag{7}$$

So the desired $\mathbf{g}_m$ can be found by minimizing the following cost function,

$$J_{ICA}(\mathbf{g}_m) = E\{f(|\hat{s}_m(n)|^2)\} = E\{f(|\mathbf{g}_m^H\mathbf{y}(n)|^2)\}, \tag{8}$$

with the constraint $\mathbf{g}_m^H\mathbf{g}_m = 1$.

Since $f(\cdot)$ is not a quadratic function, there is no closed form solution of (8). Instead, many numerical methods have been proposed to approximate the desired solution, such as the Newton method in [13]. However most of these algorithms require a long data block to acquire the high order statistics as accurately as possible, which may cause the delay of detection or may be not possible due to a fast fading channel. Furthermore, as pointed out in many papers, such as [5], [7], there will be order ambiguity as well as phase ambiguity which can not be solved by only using blind ICA algorithms. So training symbols are inevitably required in order to achieve correct detection.

B. Training based criterion

Consider the case where there are $N_t, N \geq N_t \geq M$, known symbols at the head of each data block. Hence from (4) we obtain

$$\mathbf{Z}_m = \mathbf{g}_m^H\mathbf{T}, \tag{9}$$

where $\mathbf{Z} = \begin{bmatrix} \hat{s}_m(N - N_t + 1) & \cdots & \hat{s}_m(N) \end{bmatrix}$ and $\mathbf{T} = \begin{bmatrix} \mathbf{P}\mathbf{H}\mathbf{s}(N - N_t + 1) & \cdots & \mathbf{P}\mathbf{H}\mathbf{s}(N) \end{bmatrix}$. Since both $\mathbf{Z}$ and $\mathbf{T}$ are known, the desired coefficients of the filter can be estimated by minimizing the following cost function

$$J_{training}(\mathbf{g}_m) = (\mathbf{g}_m^H\mathbf{T} - \mathbf{Z}_m)(\mathbf{g}_m^H\mathbf{T} - \mathbf{Z}_m)^H. \tag{10}$$

Training based estimation algorithms typically have the advantages that there is no ambiguity and only small time delay. However, training methods only incorporate the information in the training sequence and completely ignore the statistics in the unknown part which might be helpful to improve the system performance.

C. Semi-blind ICA cost fuction

Since a ideal solution should be able to minimize both ICA and training based criteria, it is desirable to combine these two criteria as

$$J_{semi}(\mathbf{g}_m) = E\{f(|\mathbf{g}_m^H\mathbf{y}(n)|^2)\} + \alpha(\mathbf{g}_m^H\mathbf{T} - \mathbf{Z}_m)(\mathbf{g}_m^H\mathbf{T} - \mathbf{Z}_m)^H, \quad (11)$$

where α is a scalar weight coefficient that determines the contribution of the blind and training components of the cost function. One difficulty remaining is that how to solve the semi-blind criterion in (11) which still does not have a close-form solution. In the following section, we will propose two algorithms to solve this nonquadratic criterion.

IV. SEMI-BLIND ICA ALGORITHMS

In this section, we will propose two algorithms. One is using the Newton method to solve the minimization iteratively, which is named as I-ICA. The other is using the steepest decent method and find the optimal adaptively, which is named as A-ICA.

A. Semi-blind iterative ICA algorithm

Given the existence of the first order and second order derivatives of the nonquadratic function, $f(\cdot)$, the desired solution of $\mathbf{g}_m$ should satisfy the following equation

$$\frac{\partial J_{semi}(\mathbf{g}_m)}{\partial \mathbf{g}_m} = 0, \quad (12)$$

where

$$\frac{\partial J_{semi}(\mathbf{g}_m)}{\partial \mathbf{g}_m} = \{\frac{1}{N}\sum_{n=1}^{N}\{f'(|\mathbf{g}_m^H\mathbf{y}(n)|^2)\mathbf{y}(n)\mathbf{y}(n)^H\mathbf{g}_m\}^* \quad (13)$$
$$+\alpha(\mathbf{T}\mathbf{T}^H\mathbf{g}_m - \mathbf{T}\mathbf{Z}_m)^H\}^*$$

and $f'(\cdot)$ is the first order derivative of the function $f(\cdot)$. So the optimal solution to minimize (11) should be the root of (12), which can be found by using the Newton method.

Let i denote the index of the iteration step. So during each iteration, the coefficients of the spatial equalizer are updated as

$$\mathbf{g}_m(i+1) = \mathbf{g}_m(i) - (J'_{semi}(\mathbf{g}_m))\left(\frac{\partial J'_{semi}(\mathbf{g}_m)}{\partial \mathbf{g}_m}\right)^{-1} \quad (14)$$

where

$$\frac{\partial J'_{semi}(\mathbf{g}_m)}{\partial \mathbf{g}_m} = \frac{1}{N}\sum_{n=1}^{N}\{f'(|\mathbf{g}_m^H\mathbf{y}(n)|^2)\mathbf{y}(n)\mathbf{y}(n)^H\} \quad (15)$$
$$+\frac{1}{N}\sum_{n=1}^{N}\{f'(|\mathbf{g}_m^H\mathbf{y}(n)|^2)|\mathbf{g}_m^H\mathbf{y}(n)|^2\mathbf{y}(n)\mathbf{y}(n)^H\} + \alpha\mathbf{T}\mathbf{T}^H,$$

and $J'_{semi}(\mathbf{g}_m) = (\frac{\partial J_{semi}(\mathbf{g}_m)}{\partial \mathbf{g}_m})^*$. Note that (14) requires the computation of the inverse of a $K \times K$ matrix, which has been assumed to be invertable.

The initialization of the iterative semi-blind ICA algorithm is achieved by only using the training information

$$\mathbf{g}_m(0) = (\mathbf{T}\mathbf{T}^H)^{-1}\mathbf{T}\mathbf{Z}_\mathbf{m}{}^H. \quad (16)$$

Note that our current semi-blind scheme requires at least as many training symbols as the number of receive antennas $N_t \geq K$. The advantage of the I-ICA is its fast convergence speed due to the property of the Newton method. However, during each iteration, the inverse of a matrix with dimension $K \times K$ has to be computed, which will cause extra computational complexity. Furthermore, I-ICA is not suitable for the time-varying scenario as it assumes the block-stationary fading-channel.

B. Semi-blind adaptive ICA algorithm

The basic idea of the adaptive semi-blind ICA algorithm is to continually adapt the coefficients of the spatial equalizer as new observations become available. So for the semi-blind adaptive ICA algorithm, the semi-blind cost function can be written as

$$J_{semi}(\mathbf{g}_m(n)) = \tfrac{1}{n}\textstyle\sum_{j=1}^{n} f(|\mathbf{g}_m(n)^H\mathbf{y}(j)|^2) + \quad (17)$$
$$\alpha(\mathbf{g}_m(n)^H\mathbf{T} - \mathbf{Z}_m)(\mathbf{g}_m(n)^H\mathbf{T} - \mathbf{Z}_m)^H,$$

where $n = N_t + 1, \ldots, N$. Note that here we assume the channel is constant during one block. If the channel is time varying, a forgetting factor should be introduced to adjust the cost function. To find the desired solution of (17), many classical adaptive methods could be applied. We choose the steepest decent method to avoid the inverse of the matrix in (14).

So during each time interval, the coefficients of the equalizer is updated as

$$\mathbf{g}_m(n+1) = \mathbf{g}_m(n) - \mu\frac{\partial J_{semi}(\mathbf{g}_m(n))}{\partial \mathbf{g}_m(n)}, \quad (18)$$

where μ is the step size parameter which controls the convergence speed of the adaptive algorithm and

$$\frac{\partial J_{semi}(\mathbf{g}_m(n))}{\partial \mathbf{g}_m(n)} = \{\frac{1}{n}\sum_{j=1}^{n}\{f'(|\mathbf{g}_m(n)^H\mathbf{y}(j)|^2)\mathbf{y}(j)\mathbf{y}(j)^H \quad (19)$$
$$\mathbf{g}_m(n)\}^* + \alpha(\mathbf{T}\mathbf{T}^H\mathbf{g}_m(n) - \mathbf{T}\mathbf{Z}_m)^H\}^*.$$

As can be seen from (18) the matrix inverse is no longer required compared to the iteration step in (14). Hence A-ICA is more computationally efficient than I-ICA. For fast time-varying systems, the choice of the step size parameter could be critical to ensure the fast and steady convergence speed, a low steady-state error and good tracking performance. One technique to find the optimal choice of the step size was proposed in [10], where the original cost-function of $\mathbf{g}_m$ is treated as a function of μ and then a method of steepest descent is used adaptively to find the optimal choice of μ. More parameter estimation techniques, such as the Kalman filter and the recursive least-squares method, can be found in [14], [15]. Due to the space limitation, we will report the result for the optimization of the step size in the future. In fact, since the channel is assumed to be constant over a block, the proposed semi-blind algorithm with a constant choice of μ is ready to outperform both the training and conventional ICA algorithms.

V. Numerical Study

Consider a MIMO system with $M = 4$ transmit antennas and $K = 4$ receive antennas. Both training and information symbols are chosen from a QPSK constellation. The elements of the channels and noise matrices were zero-mean, circular complex Gaussian random variables, with variances chosen to achieve the desired SNR.

The two proposed iterative and adaptive algorithms are referred as "I-ICA" and "A-ICA". The proposed algorithms are compared with the least-square training-based method [4], termed "LS training", where the coefficients of the equalizer are estimated during the training period and remain the same during the data period. For the another compared scheme, termed as "LS-ICA", the least-square training-based method is first used during the training period to estimate the coefficients of the equalizer. Then during the data period, a pure blind ICA method [7] is used to update the equalizer which is initialized by the results from the training period. Also we show the MMSE bound with perfect knowledge of the channel, where $\mathbf{g}_m = [\mathbf{H}\mathbf{H}^H + \sigma^2\mathbf{I}]^{-1}\mathbf{H}\mathbf{e}_m$ and $\mathbf{e}_m$ is an all zero vector except its mth element equal to 1.

We first show the performance of the methods as a function of SNR. The block length is chosen as $N = 200$ and the number of training symbols is $N_t = 4$. As can be seen from Fig.1, the proposed semi-blind algorithm can achieve better performance than the both lease-square based training methods. The reason for this performance gain is that, the semi-blind methods jointly incorporate the information in the known training sequence and the statistics property of the unknown sequence whereas the LS-ICA scheme separately utilizes two kinds of information during each of the two periods and the LS training scheme completely ignores the existence of the blind information. Hence for the semi-blind algorithm, there is more information available during each iteration/adaption step to improve its performance than both of the two compared schemes. Since the performance of LS-ICA is very close to the LS training scheme, we will only show the performance of LS training scheme in the sequel. As can be seen from Fig. 1, the adaptive semi-blind algorithm can outperform the iterative method, which is due to the instability of the Newton method. Recall that the Newton method can be treated as an adaptive method with constant step size $\left(\frac{\partial J'_{semi}(\mathbf{g}_m)}{\partial \mathbf{g}_m}\right)$ which might not be optimal in some circumstances and therefore cause violation.

Recall that the advantage of the proposed iterative method is its fast convergence speed, which could be vital for time varying systems. Hence in Fig. 2 we show the performance of the proposed algorithms as a function of the number of symbols used, while keeping SNR fixed. Since we assume the first four symbols used as the training, we only provide the performance when $N > 4$. As expected, the iterative method has very fast convergence speed and 20 symbols are enough to ensure its convergence although A-ICA can eventually outperform the iterative algorithm.

Since the performance of semi-blind algorithms is dependent on how much training information is available, we show the performance of the proposed algorithms in Fig. 3 as a function of the length of the training sequence, with fixed SNR. As can be seen from the figure, the performance of both the compared scheme and semi-blind algorithms can be improved by increasing training information. One interesting observation from Fig. 3 is that the performance gain of the training method with increased training information is much larger than that of the proposed algorithms, which means that semi-blind methods are less sensitive to the training information. This is due to the fact that their performance with very few training symbols is already very close to the MMSE bound, which is consistent to Fig. 1.

VI. Conclusion

In this paper, we have studied how to achieve spatial equalization for MIMO systems. We proposed two semi-blind ICA-based algorithms, which incorporate information both from the training and unknown sequences. During each iterative/adaptive step, the training information was utilized to supervise the unconstrained blind ICA-based measure. The performance of the algorithm was also demonstrated through simulation results.

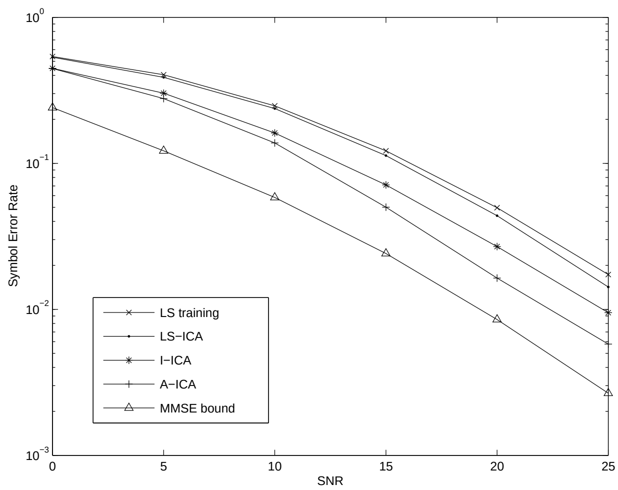

Fig. 1. Symbol error rate performance for the semi-blind algorithms and the compared schemes versus SNR.

Appendix

Proof for Claim 1: Recall that the data model can be written as[1]

$$\mathbf{y}(n) = \mathbf{P}\mathbf{r}(n) = \mathbf{P}\mathbf{H}\mathbf{s}(n), \tag{20}$$

The proof for Claim 1 is separated into two steps. First we prove that for 2×2 systems, the matrix $\mathbf{PH}$ always has the following structure

$$\mathbf{PH} = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}. \tag{21}$$

[1]To simplify the development of the proof, we only study the noiseless real-value data model here. For BPSK signals, the assumption of the real-value system will not damage the generality of the proof as the complex fading channel can be separated into two independent components, I and Q channels.

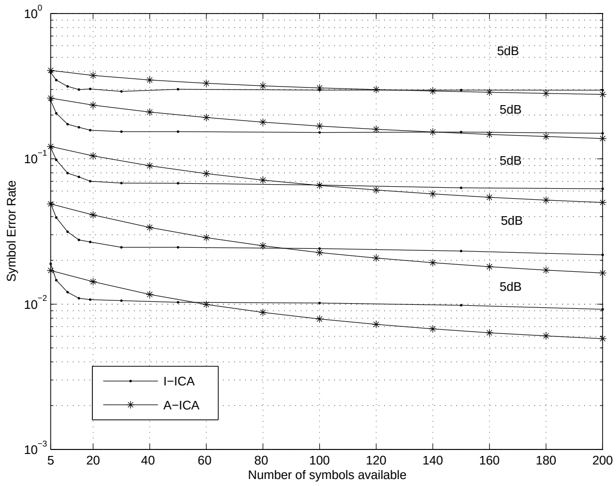

Fig. 2. Symbol error rate performance for the adaptive and iterative algorithms versus the number of symbols used.

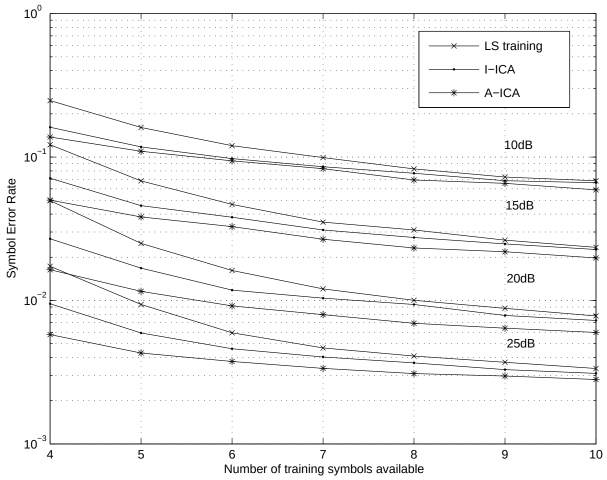

Fig. 3. Symbol error rate performance for the training and semi-blind algorithms versus the number of training symbols used.

Recall that $\mathbf{P} = \mathbf{U}\mathbf{\Lambda}^{-1/2}\mathbf{U}^T$ and the eigenvalue decomposition of the observation covariance matrix is

$$\mathbf{H}\mathbf{H}^T = E\{\mathbf{r}(n)\mathbf{r}(n)^T\} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T. \quad (22)$$

Hence the singular value decomposition of the channel matrix can be written as

$$\mathbf{H} = \mathbf{U}\mathbf{\Lambda}^{1/2}\mathbf{V}^T, \quad (23)$$

where $\mathbf{V}$ is the right singular vector of the channel matrix. Hence we obtain

$$\mathbf{P}\mathbf{H} = \mathbf{U}\mathbf{\Lambda}^{-1/2}\mathbf{U}^T\mathbf{U}\mathbf{\Lambda}^{1/2}\mathbf{V}^T = \sum_{m=1}^{M} \mathbf{u}_m \mathbf{v}_m^T, \quad (24)$$

where $\mathbf{U} = \begin{bmatrix}\mathbf{u}_1 & \cdots & \mathbf{u}_M\end{bmatrix}$ and $\mathbf{V} = \begin{bmatrix}\mathbf{v}_1 & \cdots & \mathbf{v}_M\end{bmatrix}$.

Recall that both $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices. Hence for a 2×2 system, both two matrices should have the structure as

$$\mathbf{U} = \begin{bmatrix} a_1 & -a_2 \\ a_2 & a_1 \end{bmatrix}, \quad \mathbf{V} = \begin{bmatrix} b_1 & -b_2 \\ b_2 & b_1 \end{bmatrix}. \quad (25)$$

So we can find that the structure of $\mathbf{PH}$ will be

$$\mathbf{PH} = \begin{bmatrix} a_1b_1 + a_2b_2 & a_1b_2 - a_2b_1 \\ a_2b_1 - a_1b_2 & a_1b_1 + a_2b_2 \end{bmatrix} = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}. \quad (26)$$

In the second step, we will prove that if $\mathbf{PH}$ has the structure given as in (21), the BPSK symbols $\mathbf{s}(n)$ are ready to be detected from $\mathbf{y}(n)$ with only order ambiguity. As the event of $a = b$ is a small probability event, we only consider the case where $a \neq b$.

If $a < b$, from (20) and (26) we know that the sign of the first element of $\mathbf{y}(n)$ is determined by the second element of $\mathbf{s}(n)$. And for the second element of $\mathbf{y}(n)$, the first element of $\mathbf{s}(n)$ becomes dominating component. Hence the hard decision of $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \mathbf{y}(n)$ should be equal to $\mathbf{s}(n)$ with a scale ambiguity.

If $a < b$, the result is similar and the hard decision of $\mathbf{y}(n)$ should be the same as $\mathbf{s}(n)$ with a scale ambiguity.

Summarizing these two steps, we conclude that the original transmitted symbols can be separated by only using a prewhitening filter whose output contain only order and phase ambiguity.

REFERENCES

[1] Y. Inouye and R. W. Liu, "A system-theoretic foundation for blind equalization for an FIR MIMO channel system," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 49, pp. 425–436, Apr. 2002.

[2] A. Chevreuil and P. Loubaton, "MIMO blind second-order equalization method and conjugate cyclostationarity," *IEEE Trans. Signal Processing*, vol. 42, pp. 572–578, Feb. 1999.

[3] H. Chen, X. Cao, H. Fang, and J. Zhu, "Nonlinear adaptive blind whitening for MIMO channels," *IEEE Trans. Signal Processing*, vol. 53, pp. 2635–2648, Aug. 2005.

[4] S. Kay, *Fundamentals of statistical signal processing : estimation theory*, Prentice-Hall International, 1993.

[5] P. Comon, "Independent component analysis, a new concept ?," *Signal Processing*, vol. 36, pp. 287–314, Apr. 1994.

[6] C. B. Papadias, "Globally convergent blind source separation based on a multiuser kurtosis maximization criterion," *IEEE Trans. Signal Processing*, vol. 48, pp. 3508–3519, Dec. 2000.

[7] A. Hyvarinen and E. Oja, "Independent component analysis: A tutorial," *Neural Networks*, vol. 13, pp. 411–430, 2000.

[8] Z. Ding, T. Ratnarajah and C. Cowan, "Adaptive semi-blind ICA-based spatial equalization for MIMO Rayleigh fading channels with optimal step size," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing*, May 2006, Toulouse, France.

[9] G. B. Giannakis, Y. Hua, P. Stoica, and L. Tong, *Signal processing advances in wireless and mobile communicationas*, vol. 1, Prentice-Hall International, 2001.

[10] V. Zarzoso and P. Comon, "Semi-blind constant modulus equalization with optimal step size," in *Proc. of International Conference on Acoustics, Speech, and Signal Processing*, Philadelphia,USA, 2005, pp. 18–23.

[11] A. L. Swindlehurst and G. Leus, "Blind and semi-blind equalization for generalized space-time block codes," *IEEE Trans. Signal Processing*, vol. 50, pp. 2489–2498, Oct 2002.

[12] Z. Ding, D. B. Ward, and W. H. Chin, "A general scheme for equalization of space-time block-coded systems with unknown CSI," *IEEE Trans. Signal Processing*, (to appear).

[13] E. Bingham and A. Hyvarinen, "A fast fixed-point algorithm for independent component analysis of complex valued signals," *International Journal of Neural Systems*, vol. 10, pp. 1–8, Feb. 2000.

[14] S. Haykin, *Adaptive filter theory*, Prentice-Hall International, 4th edition, 2001.

[15] S. C. Douglas and A. Cichocki, "Adaptive step size techniques for decorrelation and blind source separation," in *Proc. of the Asilomar Conference on Signals, Systems and Computers*, Pacific, Grove,CA, 1998, pp. 1191–1195.