Data-aided Maximum Likelihood Symbol Timing Estimation in MIMO Correlated Fading Channels

Yik-Chung Wu and Erchin Serpedin
Department of Electrical Engineering, Texas A&M University, College Station, TX 77843-3128, USA.
Email: {ycwu, serpedin}@ee.tamu.edu

***Abstract*— In this paper, the maximum likelihood (ML) symbol timing estimator in MIMO correlated channel based on training data is derived. It is shown that the approximated ML algorithm in [4] and [9] is just a special case of the proposed algorithm. Furthermore, the modified Cramer-Rao bound (MCRB) is also derived for comparison. Simulation results under different operating conditions (e.g., number of antennas and correlation between antennas) are given to assess the performances of the ML estimator and it is found that the mean square errors (MSE)s of the ML estimator i) are close to the MCRBs; ii) are approximately independent of the number of transmit antennas; iii) are inversely proportional to the number of receive antennas and iv) correlation between antennas has no effect on the MSE performance.**

I. Introduction

Communication over Multiple-input multiple-output (MIMO) channel has attracted much attention recently [1]-[8] due to the huge capacity gain over single antenna system. While many different techniques and algorithms have been proposed to explore the potential capacity, synchronization in MIMO channel received relatively less attention.

Symbol timing synchronization in MIMO uncorrelated flat fading channel was first studied by Naguib *et al.* [4], where orthogonal training sequences are transmitted at different transmit antennas to simplify the maximization of the over-sampled approximated log-likelihood function. This algorithm was extended by the authors in [9] to increase its estimation accuracy. Unfortunately, the algorithms in [4] and [9] are derived in an ad hoc fashion and there is no objective criteria for comparison.

In this paper, the true ML symbol timing estimator in MIMO channel based on training data is derived. Particularly, we consider correlated fading between antennas. It is shown that the approximated ML algorithm in [4] and [9] is just a special case of the proposed algorithm. Furthermore, the modified Cramer-Rao bound (MCRB) is also derived for comparison. Simulation results under different operating conditions (e.g., number of antennas and correlation between antennas) are given to assess the performances of the ML estimator and it is found that i) the MSEs of the ML estimator are close to the corresponding MCRBs; ii) the MSEs are approximately independent of the number of transmit antennas; iii) the MSEs are inversely proportional to the number of receive antennas and iv) correlation between antennas has no effect on the MSE performance.

II. Signal Model

Consider a MIMO communication system with N transmit and M receive antennas. At each receiving antenna, a superposition of independently faded signals from all the transmit antennas plus noise is received. The complex envelope of the received signal at the j^{th} receive antenna can be written as

$$r_j(t) = \sqrt{\frac{E_s}{NT}} \sum_{i=1}^{N} h_{ij} \sum_{n} d_i(n) g(t - nT - \varepsilon_o T) + \eta_j(t), \quad j = 1, 2, ..., M \tag{1}$$

where E_s/N is the symbol energy; h_{ij}'s are the complex channel coefficients between the i^{th} transmit antenna and the j^{th} receive antenna; $d_i(n)$ is the zero-mean complex valued symbol transmitted from the i^{th} transmit antenna; $g(t)$ is the transmit filter with unit energy; T is the symbol duration; $\varepsilon_o \in [0, 1)$ is the unknown timing offset and $\eta_j(t)$ is the complex-valued circularly distributed Gaussian white noise at the j^{th} receive antenna, with power density N_o. Throughout this paper, it is assumed that the channel is frequency flat and quasi-static.

After passing through the anti-aliasing filter, the received signal is then sampled at rate $f_s = 1/T_s$, where $T_s \triangleq T/Q$. Note that the oversampling factor Q is determined by the frequency span of $g(t)$; if $g(t)$ is bandlimited to $f = \pm 1/T$ (an example of which is the root raised cosine pulse), $Q = 2$ is sufficient. The received vector $\mathbf{r}_j$, which consists of $L_o Q$ consecutive received samples (L_o is the observation length), can be expressed as (without loss of generality, we consider the received sequence start at $t = 0$)

$$\mathbf{r}_j = \xi \mathbf{A}_{\varepsilon_o} \mathbf{Z} \mathbf{H}_{j,:}^T + \boldsymbol{\eta}_j, \tag{2}$$

where[1] $\xi \triangleq \sqrt{E_s/NT}$,

$$\mathbf{r}_j \triangleq [r_j(0)\ r_j(T_s)\ ...\ r_j((L_o Q - 1)T_s)]^T, \tag{3}$$

$$\mathbf{A}_\varepsilon \triangleq [\mathbf{a}_{-L_g}(\varepsilon)\ \mathbf{a}_{-L_g+1}(\varepsilon)\ ...\ \mathbf{a}_{L_o+L_g-1}(\varepsilon)], \tag{4}$$

$$\mathbf{a}_i(\varepsilon) \triangleq [g(-iT - \varepsilon T)\ g(T_s - iT - \varepsilon T)\ ...\ g((L_o Q - 1)T_s - iT - \varepsilon T)]^T, \tag{5}$$

$$\mathbf{Z} \triangleq [\mathbf{d}_1\ \mathbf{d}_2\ \cdots\ \mathbf{d}_N], \tag{6}$$

$$\mathbf{d}_i \triangleq [d_i(-L_g)\ d_i(-L_g + 1)\ \cdots\ d_i(L_o + L_g - 1)]^T \tag{7}$$

$$\boldsymbol{\eta}_j \triangleq [\eta_j(0)\ \eta_j(1)\ ...\ \eta_j(L_o Q - 1)]^T, \tag{8}$$

[1]Notation $\mathbf{x}^T$ denotes the transpose of $\mathbf{x}$, and $\mathbf{x}^H$ denotes the transpose conjugate of $\mathbf{x}$.

$$\mathbf{H} \triangleq \begin{bmatrix} h_{11} & h_{21} & \cdots & h_{N1} \\ h_{12} & h_{22} & \cdots & h_{N2} \\ \vdots & & & \vdots \\ h_{1M} & h_{2M} & \cdots & h_{NM} \end{bmatrix}, \tag{9}$$

with $\mathbf{H}_{j,:}$ denotes the j^{th} row of matrix $\mathbf{H}$, $\eta(i) \triangleq \eta(iT/Q)$, and L_g denotes the number of symbols affected by the inter-symbol interference (ISI) introduced by one side of $g(t)$.

Stacking the received vectors from all the M received antennas gives[2]

$$\mathbf{r} = \xi(\mathbf{I}_M \otimes \mathbf{A}_{\varepsilon_o})\text{vec}(\mathbf{Z}\mathbf{H}^T) + \boldsymbol{\eta} \tag{10}$$

where $\mathbf{r} \triangleq [\mathbf{r}_1^T \ \mathbf{r}_2^T \ ... \ \mathbf{r}_M^T]^T$, $\boldsymbol{\eta} \triangleq [\boldsymbol{\eta}_1^T \ \boldsymbol{\eta}_2^T \ ... \ \boldsymbol{\eta}_M^T]^T$ and $\mathbf{I}_M$ is the $M \times M$ identity matrix.

In order to include the correlation between channel coefficients, we write

$$\mathbf{H} = \sqrt{\boldsymbol{\Phi}_R}\mathbf{H}_{\text{i.i.d.}}\sqrt{\boldsymbol{\Phi}_T}^T \tag{11}$$

where $\boldsymbol{\Phi}_T$ and $\boldsymbol{\Phi}_R$ are the power correlation matrices [15] of transmit antennas and receive antennas arrays (which is assumed known), respectively; $\mathbf{H}_{\text{i.i.d.}} \in \mathbb{C}^{M\times N}$ containing independently and identically distributed (i.i.d.) zero-mean, unit-variance, circularly symmetric complex Gaussian entries and the square roots are the matrix square root (i.e., Cholesky factorization) such that $\sqrt{\boldsymbol{\Phi}}\sqrt{\boldsymbol{\Phi}}^H = \boldsymbol{\Phi}$. Note that (11) models the correlation among transmit and receive antennas array independently. This model is based on the assumption that only immediate surroundings of the antenna array impose the correlation between antennas array elements and have no impact on the correlation at the other end of the communication link. The validity of this model is verified by recent measurements [5]-[7]. Putting (11) into (10), we have

$$\mathbf{r} = \xi(\mathbf{I}_M \otimes \mathbf{A}_{\varepsilon_o})\text{vec}(\mathbf{Z}\sqrt{\boldsymbol{\Phi}_T}\mathbf{H}_{\text{i.i.d.}}^T\sqrt{\boldsymbol{\Phi}_R}^T) + \boldsymbol{\eta}. \tag{12}$$

III. ML ESTIMATOR WITH KNOWN TRAINING DATA

In this case, the matrix $\mathbf{Z}$ contains the known training sequences and the only unknown is $\mathbf{H}_{\text{i.i.d.}}$. Noting the fact that $\text{vec}(\mathbf{AYB}) = (\mathbf{B}^T \otimes \mathbf{A})\text{vec}\mathbf{Y}$, then (12) becomes

$$\begin{aligned} \mathbf{r} &= \xi(\mathbf{I}_M \otimes \mathbf{A}_{\varepsilon_o})(\sqrt{\boldsymbol{\Phi}_R} \otimes \mathbf{Z}\sqrt{\boldsymbol{\Phi}_T})\text{vec}(\mathbf{H}_{\text{i.i.d.}}^T) + \boldsymbol{\eta} \\ &= \xi(\sqrt{\boldsymbol{\Phi}_R} \otimes \mathbf{A}_{\varepsilon_o}\mathbf{Z}\sqrt{\boldsymbol{\Phi}_T})\text{vec}(\mathbf{H}_{\text{i.i.d.}}^T) + \boldsymbol{\eta} \end{aligned} \tag{13}$$

where the last line come from the fact that $(\mathbf{A}\otimes\mathbf{B})(\mathbf{C}\otimes\mathbf{D}) = (\mathbf{AC}) \otimes (\mathbf{BD})$.

From (13), the joint maximum likelihood estimate of ε_o and $\text{vec}(\mathbf{H}_{\text{i.i.d.}}^T)$ is given by maximizing

$$p(\mathbf{r}|\varepsilon, \mathbf{h}) = \frac{1}{(\pi N_o)^{L_o Q}} \exp\left[-\frac{(\mathbf{r} - \bar{\mathbf{A}}_\varepsilon \mathbf{h})^H(\mathbf{r} - \bar{\mathbf{A}}_\varepsilon \mathbf{h})}{N_o}\right], \tag{14}$$

or equivalently minimizing

$$J(\mathbf{r}|\varepsilon, \mathbf{h}) = (\mathbf{r} - \bar{\mathbf{A}}_\varepsilon \mathbf{h})^H(\mathbf{r} - \bar{\mathbf{A}}_\varepsilon \mathbf{h}), \tag{15}$$

[2]Notation $\otimes$ denotes Kronecker products and $\text{vec}(\mathbf{H})$ denotes a $MN \times 1$ vector formed by stacking the columns of $\mathbf{H}$ under each other.

where $\bar{\mathbf{A}}_\varepsilon \triangleq \xi(\sqrt{\boldsymbol{\Phi}_R}\otimes\mathbf{A}_\varepsilon\mathbf{Z}\sqrt{\boldsymbol{\Phi}_T})$, ε and $\mathbf{h}$ are the trial values for ε_o and $\text{vec}(\mathbf{H}_{\text{i.i.d.}}^T)$, respectively.

From the linear signal model given in (13), the ML estimate for $\text{vec}(\mathbf{H}_{\text{i.i.d.}}^T)$ (when ε is fixed) is [10]

$$\hat{\mathbf{h}} = (\bar{\mathbf{A}}_\varepsilon^H \bar{\mathbf{A}}_\varepsilon)^{-1}\bar{\mathbf{A}}_\varepsilon^H \mathbf{r}. \tag{16}$$

Plugging (16) into (15), after some straightforward manipulations and dropping the irrelevant terms, the timing delay is estimated by maximizing the following likelihood function

$$\Lambda(\varepsilon) = \mathbf{r}^H \bar{\mathbf{A}}_\varepsilon(\bar{\mathbf{A}}_\varepsilon^H \bar{\mathbf{A}}_\varepsilon)^{-1}\bar{\mathbf{A}}_\varepsilon^H \mathbf{r}. \tag{17}$$

Using the fact $(\mathbf{A}\otimes\mathbf{B})^{-1} = \mathbf{A}^{-1}\otimes\mathbf{B}^{-1}$ and $(\mathbf{A}\otimes\mathbf{B})^H = \mathbf{A}^H \otimes \mathbf{B}^H$ to expand $\bar{\mathbf{A}}_\varepsilon(\bar{\mathbf{A}}_\varepsilon^H\bar{\mathbf{A}}_\varepsilon)^{-1}\bar{\mathbf{A}}_\varepsilon^H$, we have

$$\begin{aligned} &\bar{\mathbf{A}}_\varepsilon(\bar{\mathbf{A}}_\varepsilon^H\bar{\mathbf{A}}_\varepsilon)^{-1}\bar{\mathbf{A}}_\varepsilon^H \\ &= [\sqrt{\boldsymbol{\Phi}_R}(\sqrt{\boldsymbol{\Phi}_R}^H\sqrt{\boldsymbol{\Phi}_R})^{-1}\sqrt{\boldsymbol{\Phi}_R}^H] \\ &\otimes [\mathbf{A}_\varepsilon\mathbf{Z}\sqrt{\boldsymbol{\Phi}_T}(\sqrt{\boldsymbol{\Phi}_T}^H\mathbf{Z}^H\mathbf{A}_\varepsilon^H\mathbf{A}_\varepsilon\mathbf{Z}\sqrt{\boldsymbol{\Phi}_T})^{-1}\sqrt{\boldsymbol{\Phi}_T}^H\mathbf{Z}^H\mathbf{A}_\varepsilon^H] \\ &= \mathbf{I}_M \otimes \mathbf{A}_\varepsilon\mathbf{Z}(\mathbf{Z}^H\mathbf{A}_\varepsilon^H\mathbf{A}_\varepsilon\mathbf{Z})^{-1}\mathbf{Z}^H\mathbf{A}_\varepsilon^H \end{aligned} \tag{18}$$

where in the second equality, we note that $\sqrt{\boldsymbol{\Phi}_R}$ and $\sqrt{\boldsymbol{\Phi}_T}$ are both square matrices. Putting this result into (17), the likelihood function is given by

$$\begin{aligned} \Lambda(\varepsilon) &= \mathbf{r}^H(\mathbf{I}_M \otimes \mathbf{A}_\varepsilon\mathbf{Z}(\mathbf{Z}^H\mathbf{A}_\varepsilon^H\mathbf{A}_\varepsilon\mathbf{Z})^{-1}\mathbf{Z}^H\mathbf{A}_\varepsilon^H)\mathbf{r} \\ &= \sum_{j=1}^{M} \mathbf{r}_j^H\mathbf{A}_\varepsilon\mathbf{Z}(\mathbf{Z}^H\mathbf{A}_\varepsilon^H\mathbf{A}_\varepsilon\mathbf{Z})^{-1}\mathbf{Z}^H\mathbf{A}_\varepsilon^H\mathbf{r}_j \end{aligned} \tag{19}$$

and the problem of symbol timing estimation can be written as

$$\hat{\varepsilon} = \arg\max_{\varepsilon} \Lambda(\varepsilon). \tag{20}$$

We make the following remarks:

1) The maximization of the likelihood function usually involves a two-step approach. The first step (coarse search) computes $\Lambda(\varepsilon)$ over a grid of timing delay $\varepsilon_k \triangleq k/K$ for $k = 0, 1, ..., K-1$, and then the ε_k that maximizes $\Lambda(\varepsilon)$ is selected. The second step (fine search) finds the global maximum by using either gradient method [14], dichotomous search [12], or interpolation [12]. In this paper, we employ the parabolic interpolation in the second step due to its implementation simplicity. More specifically, assume that $\Lambda(\varepsilon_{\hat{k}})$ is identified as the maximum among all $\Lambda(\varepsilon_k)$ in the first step. Define $I_1 \triangleq \Lambda(\varepsilon_{\hat{k}-1})$, $I_2 \triangleq \Lambda(\varepsilon_{\hat{k}})$ and $I_3 \triangleq \Lambda(\varepsilon_{\hat{k}+1})$, then [12]
$$\hat{\varepsilon} = \varepsilon_{\hat{k}} + \frac{I_1 - I_3}{2K(I_1 + I_3 - 2I_2)}. \tag{21}$$
2) The likelihood function at each receive antenna can be calculated independently and then added together to obtain the overall likelihood function.
3) The correlations in the transmit and receive antenna arrays do not appear in the estimator. That is, the ML symbol timing estimator is independent of the antennas correlation.

4) In order for the estimate of $\text{vec}(\mathbf{H}^T_{\text{i.i.d.}})$ to hold in (16), it is required that $\bar{\mathbf{A}}_\varepsilon$ is full rank [10], or equivalently $\sqrt{\mathbf{\Phi}_R}$, $\mathbf{A}_\varepsilon$, $\mathbf{Z}$ and $\sqrt{\mathbf{\Phi_T}}$ are all full rank. Note that $\sqrt{\mathbf{\Phi}_R}$ and $\sqrt{\mathbf{\Phi_T}}$ are lower triangular matrices with positive diagonal elements [11], so they are full rank. Furthermore, for $g(t)$ being a root raised cosine pulse (which is the most frequently used pulse shape), numerical calculations show that $\mathbf{A}_\varepsilon$ is full rank. Finally, $\mathbf{Z}$ can be made full rank by properly designing the training data. A sufficient condition is that, for $i \neq j$,
$$[d_i(a)\ \cdots\ d_i(b)] \cdot [d_j(a)\ \cdots\ d_j(b)]^H = 0 \qquad (22)$$
for some $a, b \in \{-L_g, -L_g+1, \ldots, L_o+L_g-1\}$ with $a < b$.

5) If $g(t)$ is a root raised cosine pulse, and $L_g = 0$, then the $(i,j)^{th}$ element of $\mathbf{A}_\varepsilon^H \mathbf{A}_\varepsilon$ $(i,j = 0,1,...,L_o-1)$ is
$$\begin{aligned}&[\mathbf{A}_\varepsilon^H \mathbf{A}_\varepsilon]_{ij} \\ &= \sum_{n=0}^{L_oQ-1} g(nT_s - iT - \varepsilon T) g(nT_s - jT - \varepsilon T) \approx \delta_{ij}.\end{aligned} \qquad (23)$$
Furthermore, if the training sequences from different transmit antennas are orthogonal and with the same norm (i.e., $\|\mathbf{d}_i\|^2 = c$ is a constant), then
$$\Lambda(\varepsilon) \approx \frac{1}{c}\sum_{j=1}^{M} \mathbf{r}_j^H \mathbf{A}_\varepsilon \mathbf{Z}\mathbf{Z}^H \mathbf{A}_\varepsilon^H \mathbf{r}_j = \frac{1}{c}\sum_{j=1}^{M}\sum_{i=1}^{N} |\mathbf{d}_i^H \mathbf{A}_\varepsilon^H \mathbf{r}_j|^2. \qquad (24)$$
For sufficiently large observation interval, $\mathbf{A}_\varepsilon^H \mathbf{r}_j$ is the matched filtering with one output sample per symbol with delay ε [13]. This reduces to the approximated ML function proposed in [4].

IV. The Modified CRB

For the model in (13), it is known that for a specific timing delay ε_o, the MCRB is given by [13]
$$\text{MCRB}(\varepsilon_o) = \frac{\sigma^2}{2\text{tr}(\bar{\mathbf{D}}_{\varepsilon_o}^H \bar{\mathbf{D}}_{\varepsilon_o} \mathbf{\Gamma_h})} \qquad (25)$$
where $\sigma^2 = N_o f_s = N_o Q/T$ is the noise variance, tr(.) denotes the trace of a matrix,
$$\bar{\mathbf{D}}_\varepsilon \triangleq \frac{d\bar{\mathbf{A}}_\varepsilon}{d\varepsilon} = \xi \sqrt{\mathbf{\Phi}_R} \otimes \mathbf{D}_\varepsilon \mathbf{Z} \sqrt{\mathbf{\Phi_T}} \qquad (26)$$
with $\mathbf{D}_\varepsilon \triangleq d\mathbf{A}_\varepsilon / d\varepsilon$ and
$$\mathbf{\Gamma_h} \triangleq \mathbb{E}[\text{vec}(\mathbf{H}^T_{\text{i.i.d.}})\text{vec}(\mathbf{H}^T_{\text{i.i.d.}})^H] = \mathbf{I}_{MN} = \mathbf{I}_M \otimes \mathbf{I}_N. \quad (27)$$
Plugging (26) and (27) into (25), after some calculations, we get
$$\begin{aligned}&\text{MCRB}(\varepsilon_o) \\ &= \frac{QN}{2\text{tr}(\sqrt{\mathbf{\Phi}_R}^H \sqrt{\mathbf{\Phi}_R})\text{tr}(\sqrt{\mathbf{\Phi_T}}^H \mathbf{Z}^H \mathbf{D}_{\varepsilon_o}^H \mathbf{D}_{\varepsilon_o} \mathbf{Z} \sqrt{\mathbf{\Phi_T}})} \left(\frac{E_s}{N_o}\right)^{-1} \\ &= \frac{1}{2M\text{tr}(\tilde{\mathbf{Z}}^H \tilde{\mathbf{D}}_{\varepsilon_o}^H \tilde{\mathbf{D}}_{\varepsilon_o} \tilde{\mathbf{Z}} \mathbf{\Phi_T})} \left(\frac{E_s}{N_o}\right)^{-1}\end{aligned} \qquad (28)$$
where $\tilde{\mathbf{Z}} \triangleq \mathbf{Z}/\sqrt{N}$ and $\tilde{\mathbf{D}}_\varepsilon \triangleq \mathbf{D}_\varepsilon/\sqrt{Q}$. From second line to the third line, we used the fact that $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$ and the diagonal elements of $\mathbf{\Phi_R}$ must be all one, regardless what the correlation matrix is.

We make the following remarks concerning the MCRB:

1) Since the timing delay ε_o is assumed uniformly distributed, the average MCRB can be calculated by numerical integration of (28).
2) The MCRB does not depends on the receive antenna array correlation. Furthermore, the MCRB are inversely proportional to the number of receive antenna. That is, the MCRB will be reduced by half when the number of receive antennas is double, regardless the correlation between the receive antennas.

V. Simulation results and discussions

In this section, the MSE performance of the proposed symbol timing estimator is assessed by Monte Carlo simulations and then compared with the MCRB derived in Section IV. In all the simulations, $L_o = 32$, $L_g = 4$ (i.e., the total length of training data is 40), $Q = 2$, $K = 16$, ε_o is uniformly distributed in the range $[0,1)$ and $g(t)$ being a root raised cosine filter with roll-off factor $\alpha = 0.3$. The training sequences used are the sequences proposed in [9]. Each point is obtained by averaging 10^4 simulation runs. The MCRBs are also shown for comparison.

First, let assume $\mathbf{\Phi}_T = \mathbf{I}_N$ and $\mathbf{\Phi}_R = \mathbf{I}_M$ for the moment. The effect of correlation among antennas will be examined in the last figure. The effect of the number of transmit antennas N is shown Fig. 1 with $M = 4$. It can be seen that the increase of transmit antenna does not improve the estimation accuracy. Closer examination reveals that the MCRBs for different number of transmit antennas basically coincide. Therefore, increasing N does not improve the performance. Next, the effect of number of receive antennas M is shown in Fig. 2 with $N = 4$. It is clear that increasing M leads to considerable MSE improvements. Since from (28), the MCRB is inversely proportional to M and from Fig. 2, the performances of the proposed estimator are very close to their corresponding MCRBs, it can be concluded that the MSE of ML estimator is approximately inversely proportional to M.

Fig. 3 shows the effect of fading correlation among antennas. Measured correlation matrices from Nokia [15] are used in the simulation:
$$\mathbf{\Phi}_T = \begin{bmatrix} 1 & 0.4154 & 0.2057 & 0.1997 \\ 0.4154 & 1 & 0.3336 & 0.3453 \\ 0.2057 & 0.3336 & 1 & 0.5226 \\ 0.1997 & 0.3453 & 0.5226 & 1 \end{bmatrix} \qquad (29)$$
$$\mathbf{\Phi}_R = \begin{bmatrix} 1 & 0.3644 & 0.0685 & 0.3566 \\ 0.3644 & 1 & 0.3245 & 0.1848 \\ 0.0685 & 0.3245 & 1 & 0.3093 \\ 0.3566 & 0.1848 & 0.3093 & 1 \end{bmatrix}. \qquad (30)$$
It can be seen that the fading correlation among antennas does not change the MSE performance of the ML estimator or the

MCRB. Finally, note that from all 3 figures, the MSEs of the ML estimator are very close to the corresponding MCRBs.

VI. Conclusions

The ML symbol timing estimator and the MCRB in MIMO correlated channel based on training data has been derived. It is shown that the approximated ML algorithm in [4] and [9] is just a special case of the proposed algorithm. Simulation results were given to assess the performances of the ML estimator and it is found that i) the MSEs of the ML estimator are close to the MCRBs; ii) the MSEs are approximately independent of the number of transmit antennas; iii) the MSEs are inversely proportional to the number of receive antennas and iv) correlation between antennas has little effect on the MSEs and the MCRBs.

References

[1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, pp. 311-335, Mar. 1998.

[2] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas in Commun.*, vol. 16, pp. 1451-1458, Oct. 1998.

[3] A. F. Naguib, N. Seshadri and A. R. Calderbank, "Increasing data rate over wireless channels," *IEEE Signal Processing Magazine*, vol. 17, pp. 76-92, May 2000.

[4] A. F. Naguib, V. Tarokh, N. Seshadri and A. R. Calderbank, "A space-time coding modem for high-data-rate wireless communications," *IEEE J. Select. Areas in Commun.*, vol. 16, pp. 1459-1478, Oct. 1998.

[5] K. Yu, M. Bengtsson, B. Ottersten, D. McNamara, P. Karlsson and M. Beach,"Second Order Statistics of NLOS Indoor MIMO Channels based on 5.2 GHz Measurements," *Proceedings of IEEE GLOBECOM 01*, vol. 1, pp. 156-160, Nov. 2001.

[6] J. P. ermoal, L. Schumacher, K. I. Pedersen, P. E. Mogensen and F. Frederisken, "A stochastic MIMO radio channel model with experimental validation," *IEEE J. Select. Areas in Commun.*, vol. 20, pp. 1211-1226, Aug. 2002.

[7] D. Chizhik, J. Ling, P. W. Wolniansky, R. A. Valenzuela, N. Costa and K. Huber, "Multiple-input-multiple-output measurements and modeling in Manhattan," *IEEE J. Select. Areas in Commun.*, vol. 21, pp. 321-331, Apr. 2003.

[8] D. Gesbert, M. Shafi, D.-S. Shiu, P. J. Smith and A. Naguib, "From theory to practice: an overview of MIMO space-time coded wireless systems," *IEEE J. Select. Areas in Commun.*, vol. 21, pp. 281-301, Apr. 2003.

[9] Y.-C. Wu, S. C. Chan and E. Serpedin, "Symbol-timing estimation in space-time coding systems based on orthogonal training sequences," submitted to *IEEE Trans. on Wireless Communications*, Mar. 2003. Available online: http://ee.tamu.edu/∼serpedin

[10] S. M. Kay, *Fundamentals of Statistical Signal Processing - Estimation Theory*. Prentice Hall, 1993.

[11] R. A. Horn and C. R. Johnson, "Matrix analysis," Cambridge University Press, 1990.

[12] Y. V. Zakharov, V. M. Baronkin and T. C. Tozer, "DFT-based frequency estimators with narrow acquisition rnage," *IEE Proc.-Commun.*, Vol. 148, No. 1, pp. 1-7, Feb. 2001.

[13] G. Vazquez and J. Riba, "Non-data-aided digital synchronization," in Signal Processing Advanced in Wireless and Mobile Communications: Volume 2 (edited by G. B. Giannakis, Y. Hua, P. Stoica and L. Tong), Prentice Hall 2001.

[14] B. Ottersten, M. Viberg, P. Stoica and A. Nehorai, "Exact and large sample maximum likelihood techniques for parameter estimation and detectionin array processing," in *Radar array processing*, Springer-Verlag, 1993.

[15] L. Schumacher, J. P. Kermoal, F. Frederiksen, K. I. Pedersen, A. Algrans and P. E. Mogensen, "MIMO channel characterisation," *Deliverable D2 V1.1 of IST-1999-11729 METRA project*, pp. 1-57, Feb. 2001. Available online: http://www.ist-metra.org

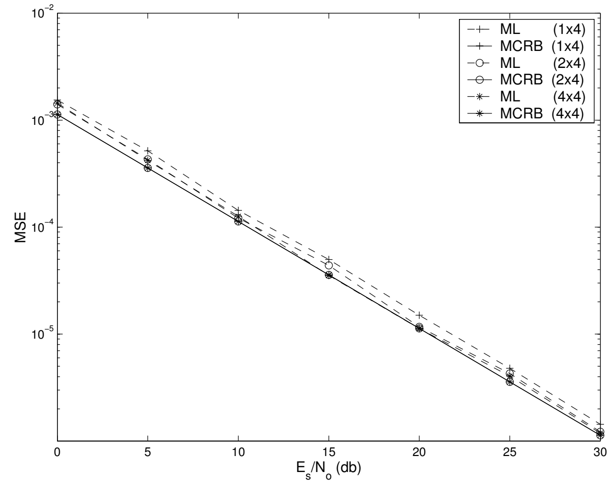

Fig. 1. MSEs of the proposed estimator with different number of transmit antennas (assuming $\mathbf{\Phi}_T = \mathbf{I}_N$ and $\mathbf{\Phi}_R = \mathbf{I}_M$).

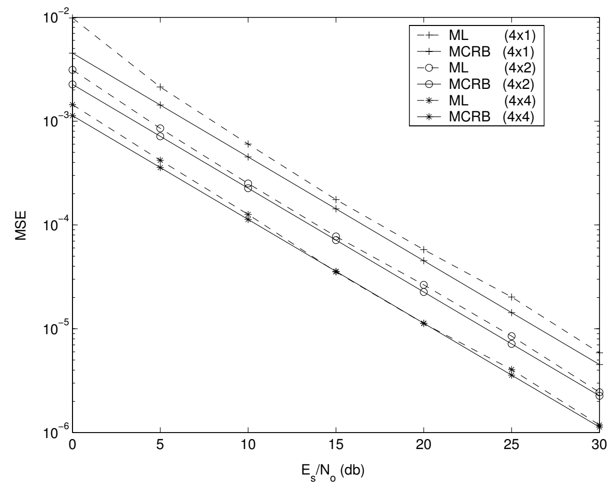

Fig. 2. MSEs of the proposed estimator with different number of receive antennas (assuming $\mathbf{\Phi}_T = \mathbf{I}_N$ and $\mathbf{\Phi}_R = \mathbf{I}_M$).

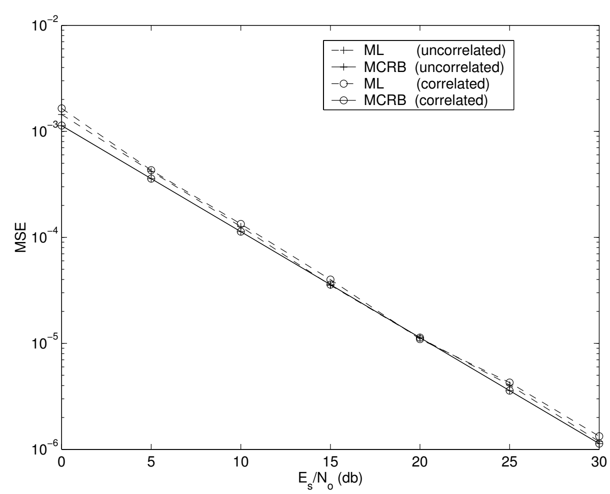

Fig. 3. MSEs of the proposed estimator with and without fading correlation between antennas for a 4×4 system.