Joint Channel and Frequency Offset Estimation and Training Sequence Design for MIMO Systems over Frequency Selective Channels

Tao Cui and Chintha Tellambura
Department of Electrical and Computer Engineering
University of Alberta
Edmonton, AB, Canada T6G 2V4
Email: {taocui, chintha}@ece.ualberta.ca

***Abstract*— We consider the joint estimation of channel impulse response (CIR) and frequency offset (FO) for a multiple-input multiple-output (MIMO) frequency selective fading channel. We average the joint likelihood function of the CIR and the FO over the distribution of CIR and maximize the resulting marginal likelihood function to estimate the FO. We also derive the design criteria for training sequence (TS) which will maximize the performance of our proposed estimator. Simulation results show that the optimal TS can improve the performance of the joint ML estimator.**

I. Introduction

To cater for the increased demand for wireless high-speed data transmission, the MIMO paradigm has been embraced by both the academic and industrial research community, which uses multiple antennas at both the transmitter and receiver. It has been proven effective in combating fading and enhancing data rates [1]. However, MIMO systems suffer from two problems: (1) The performance is highly sensitive to FO; (2) channel estimation becomes more complicated as the number of antennas increases. Thus FO correction and channel estimation (CE) are crucial to MIMO systems.

Many existing CE methods use Pilot-Symbol-Assisted Modulation (PSAM), in which known pilot symbols are embedded in each frame to aid CE and/or FO estimation. Although the selection of pilot symbols or TSs is very important to the accuracy of the estimator in PSAM systems, only few studies address this issue. Tellambura *et.al.* [2] give the optimal sequences design for channel estimation. Morelli and Mengali [3] consider TS design for FO estimation. The optimal TS design for joint channel and FO estimation has been proposed in [4] using a Cramér-Rao Bound (CRB) approach.

In this paper, for a MIMO frequency selective channel, we derive an ML FO estimator using a Baysian approach [5]. We average the joint likelihood function over the channel statistics and get the marginal likelihood for the FO. A ML channel estimator is also derived. We analyze the performance of the proposed joint FO and channel estimator and derive the mean square error (MSE) of the estimated system parameters. Design criteria for optimal sequences which will minimize the MSE are derived. The CRB [6] for the estimation of FO is also derived. Simulation results show that the optimal TS can reduce the MSE of both the channel and FO estimation.

The rest of the paper is organized as follows. Section II briefly reviews the basic MIMO system model. Section III introduces the joint ML FO and channel estimation. The analysis of the ML estimator and TS design criteria are derived in Section IV. Computer simulation results are given in Section V and final conclusions are made in Section VI.

Notation: Boldface letters will be used for matrices and column vectors; $(\cdot)^H$ will denote Hermitian (conjugate transpose); I_N denotes the $N \times N$ identity matrix; diag$\{\mathbf{x}\}$ stands for the diagonal matrix with the column vector $\mathbf{x}$ on its diagonal. $\mathrm{tr}(\mathbf{A}) = \sum_{i=1}^{N} a_{ii}$ is the trace of matrix A. $\jmath = \sqrt{-1}$.

II. System Model

The baseband MIMO system considered in this paper has N_t transmit antennas and N_r receive antennas. Each channel is frequency selective but the CIR and FO remain constant for a data frame with K data symbols. N pilot symbols are inserted in the front of each data frame. The transmit symbols are $a_t(k) \in \mathcal{Q}$ where finite alphabet $\mathcal{Q}$ is a linear modulation such as PSK or QAM. We assume the frequency offsets between different transmit-receive antenna pairs are the same, and denoted by Δf_q. The case of different FO's will be considered in the full journal version of this paper. The received signal samples at receive antenna q, sampled at symbol rate $f_s = 1/T_s$, can be expressed by

$$r_q(k) = \sum_{t=1}^{N_t} \sum_{l=0}^{L-1} a_t(k-l) e^{\jmath 2\pi v_q k} h_{t,q}(l) + w(k) \qquad (1)$$

where $k = 0, 1, \ldots, N-1$, $q = 1, 2, \ldots, N_r$, $v_q = \Delta f_q T_s$ is the normalized FO; $h_{t,q}(l)$ is the response of the equivalent channel, baseband pulse shape and the matched filter; $w(n)$ is a zero mean Gaussian additive noise with variance σ_n^2.

In each data frame, a preamble of TS is inserted to estimate the channel and FO. The TS is sent every K symbols.

The estimates are used to demodulate the remaining data symbols in the frame. We assume $N+L-1$ training symbols $a_t(n), -L+1 \leq n \leq N-1$ are transmitted at each transmit antenna. We assume that $a_t(-L+j) = a_t(N-L+j)$ for $1 \leq j \leq L-1$. Dropping the first $L-1$ samples (precursors), we define the received vector at q receive antenna $\mathbf{r}_q = [r_q(0), r_q(1), \cdots, r_q(N-1)]^T$ that corresponding to the TS and $\mathbf{h}_{t,q} = [h_{t,q}(0), h_{t,q}(1), \cdots, h_{t,q}(L-1)]$ the channel vector between the t-th transmit antenna and the q-th receive antenna, which are i.i.d. Gaussian random variables with zero mean and variance $\sigma_{t,q}^2$. Eq. (1) can be written in compact matrix form as

$$\mathbf{r}_q = \mathbf{\Gamma}(v_q) \sum_{t=1}^{N_t} \mathbf{A}_t \mathbf{h}_{t,q} + \mathbf{w}_q \tag{2}$$

where $\mathbf{A}_t$ is a $N \times L$ circulant matrix with entries

$$[\mathbf{A}_t]_{i,j} = a_t(i-j), 0 \leq i \leq N-1, 0 \leq j \leq L-1 \tag{3}$$

and $\mathbf{\Gamma}(v_q)$ is a diagonal matrix

$$\mathbf{\Gamma}(v_q) = \text{diag}\left\{1, e^{j2\pi v_q}, e^{j4\pi v_q}, \cdots, e^{j2\pi(N-1)v_q}\right\} \tag{4}$$

v_q is the FO for the q receive antenna. The matrix form of (2) becomes

$$\mathbf{r}_q = \mathbf{\Gamma}(v_q)\mathbf{A}\mathbf{h}_q + \mathbf{w}_q \tag{5}$$

where $\mathbf{A} = [\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_{N_t}]$ and $\mathbf{h}_q = [\mathbf{h}_{1,q}{}^T, \mathbf{h}_{2,q}{}^T, \cdots, \mathbf{h}_{N_t,q}{}^T]^T$.

III. Frequency offset and channel estimation for MIMO system

The received symbol vector $\mathbf{r}_q$ is Gaussian with mean $\mathbf{D}_q\mathbf{h}_q$ and covariance matrix $\sigma_n^2\mathbf{I}_N$. The likelihood function for the unknown parameters $\mathbf{h}_q$ and v_q is given by

$$\Lambda(\mathbf{r}_q|\mathbf{h}_q, v_q) = \exp\left\{-\frac{1}{2\sigma_n^2}\|\mathbf{r}_q - \mathbf{\Gamma}(v_q)\mathbf{A}\mathbf{h}_q\|^2\right\}. \tag{6}$$

A. Frequency offset estimation

We use the ML approach to estimate the FO. The authors in [3] develop a method based on ML. They first fix v and find the $\mathbf{h}(v)$ to maximize the likelihood function. They next fix $\mathbf{h}$ and find $v(\mathbf{h})$ to maximize the likelihood function. Finally, they substitute the $\mathbf{h}(v)$ to $v(\mathbf{h})$ and get the optimal v. Here we take a different approach to this problem.

Since channel response can be modelled as complex Gaussian process, we can compute the average of $\Lambda(\mathbf{r}_q|\mathbf{h}_q, v_q)$ with respect to $\mathbf{h}_q$, which gives the marginal likelihood of v_q. It removes the likelihood function's dependence on $\mathbf{h}$. In this way, we can get the FO estimator independent of CIR. The marginal likelihood function is given by

$$\begin{aligned}\Lambda(\mathbf{r}_q|v_q) &= \int \Lambda(\mathbf{r}_q|\mathbf{h}_q, v_q) f(\mathbf{h}_q) d\mathbf{h}_q \\ &= c\exp\left\{\frac{1}{2\sigma_n^2}\left[\mathbf{r}_q{}^H\mathbf{\Gamma}(v_q)\mathbf{G}\mathbf{\Gamma}^H(v_q)\mathbf{r}_q\right]\right\}\end{aligned} \tag{7}$$

where $\mathbf{G}$ is

$$\mathbf{G} = \mathbf{A}(\mathbf{A}^H\mathbf{A} + 2\sigma_n^2\mathbf{R}_{\mathbf{h}_q}^{-1})^{-1}\mathbf{A}^H. \tag{8}$$

c is a constant, $\mathbf{R}_{\mathbf{h}_q} = E\{\mathbf{h}_q\mathbf{h}_q^H\}$ is the autocorrelation matrix of $\mathbf{h}_q$. Maximizing (7) is equivalent to maximizing

$$g(v_q) = \mathbf{r}_q{}^H\mathbf{\Gamma}(v_q)\mathbf{G}\mathbf{\Gamma}^H(v_q)\mathbf{r}_q \tag{9}$$

(9) can be further reduced to

$$\begin{aligned}g(v_q) =& \gamma^T(v_q)\mathbf{R}_q{}^H\mathbf{G}\mathbf{R}_q\gamma^*(v_q) \\ =& \gamma^T(v_q)\mathbf{B}\gamma^*(v_q)\end{aligned} \tag{10}$$

where $\mathbf{B} = \mathbf{R}_q^H\mathbf{G}\mathbf{R}_q$ is Hermitian. In high SNR region, $\mathbf{G} \approx \mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H$ and

$$\hat{v}_q = \arg\max_{v_q} \mathbf{r}_q{}^H\mathbf{\Gamma}(v_q)\mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\mathbf{\Gamma}^H(v_q)\mathbf{r}_q \tag{11}$$

which is the MLE#1 in [3]. This shows that our proposed estimator has the same asymptotical property as MLE#1 in [3].

When the FO is constant and and the pilot symbols are the same during J data frames, the J received signals can be combined for FO estimation. Let $\mathbf{r}_q(k)$ denote the received signal in the k-th frame. Since $\mathbf{r}_q(k)$ for $k = 1, 2, \ldots, J$ are independent, the optimal FO estimator can be readily obtained as

$$\hat{\mathbf{v}}_q = \arg\max_{v_q} \sum_{k=1}^{J} \mathbf{r}_q^H(k)\mathbf{\Gamma}(v_q)\mathbf{G}\mathbf{\Gamma}^H(v_q)\mathbf{r}_q(k) \tag{12}$$

Remarks:

- From (9), the estimation of v_q and $\mathbf{h}$ is decoupled, which means that v_q can be calculated and subsequently used for estimating $\mathbf{h}$. The $\hat{v}_q$ given by this estimator is used as the initial estimation in the channel estimator in Section III. B.
- In the high SNR region, (9) reduces to

$$\begin{aligned}\hat{v}_q &= \arg\max_{v_q} g(v_q) \\ &= \arg\max_{v_q} \|\mathbf{\Pi}_A\mathbf{\Gamma}^H(v_q)\mathbf{r}_q\|^2\end{aligned} \tag{13}$$

  where $\mathbf{\Pi_A} = \mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H$ is the orthogonal projection matrix on the subspace of $\mathbf{A}$. Hence (13) can be interpreted as choosing v_q that can make $\mathbf{r}_q$ orthogonal to $\mathbf{\Pi_A}$ or equivalently parallel to the orthogonal projection matrix on the zero space of $\mathbf{A}$, $\mathbf{\Pi}_{\mathbf{A}}^{\perp} = \mathbf{I} - \mathbf{\Pi_A}$. Therefore, the rank of the zero space $N - LN_t$ determines the accuracy of the estimate.
- The matrix $\mathbf{G}$ in (9) can be precomputed. The main computational cost is to find v in (9). Eq. (10) can be written as

$$g(v_q) = \sum_{i=0}^{N-1}\sum_{k=0}^{N-1} b_{i,k}e^{j2\pi(i-k)v_q} \tag{14}$$

which can viewed as a polynomial, where $b_{i,k}$ is the (i,k)-th entry of $\mathbf{B}$. As discussed in [7], the minima of the cost function can be performed in a two-step procedure: coarse search and fine search. Fast algorithm can applied to find the minimum of (14) using FFT and parabolic interpolation.

- Eq. (9) reveals that $g(v_q)$ is periodic. If v_q out of the interval $-0.5 < v_q \leq 0.5$, the proposed FO estimator will give ambiguous estimates. Therefore $-0.5 < v_q \leq 0.5$ is the estimation range of the ML estimator.

B. Channel estimation

In this subsection, an ML channel estimator based on the likelihood function and the above estimated FO $\hat{v}_q$ is derived. We assume that the FO estimate is perfect, $\hat{v}_q = v_q$.

If FO v_q is known, the least squares (LS) estimate of $\mathbf{h}_q$ is given by

$$\hat{\mathbf{h}}_q = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\mathbf{\Gamma}^H(v_q)\mathbf{r}_q. \tag{15}$$

Note that $\mathbf{A}^H\mathbf{A}$ and $\mathbf{A}^H\mathbf{A}$ are assumed to be invertible. The LS estimate relates to the true CIR by

$$\hat{\mathbf{h}}_q = \mathbf{h}_q + (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\mathbf{\Gamma}^H(v_q)\mathbf{w}_q. \tag{16}$$

That is, the LS estimate is equal to the true CIR perturbed by additive noise. The variance of the noise is determined by the TS and FO.

IV. Performance analysis and optimal training sequence design

We now derive the MSEs of the FO estimation and LS channel estimation. Optimal TS design criteria based on these MSEs are also derived. The optimal TS design has been considered in [4] using the CRB. The TS is designed by minimizing the trace of the CRB matrix which is equivalent to minimizing the weighted sum of FO CRB and CE CRB. However, it is impossible to minimize the CRB's for FO and CE individually. Instead of CRB, we use the MSE criterion to design TS.

A. Training sequence for channel estimation

From (16), the MSE of the LS channel estimator can be readily obtained as

$$\text{MSE}(\hat{\mathbf{h}}_q) = \frac{\sigma_n^2}{LN_t}\text{tr}\{(\mathbf{A}^H\mathbf{A})^{-1}\}. \tag{17}$$

Note that the MSE in (17) is independent of FO v_q. Hence, the optimal TS design is decoupled. Our goal is to design optimal TS to minimize the MSE of $\hat{\mathbf{h}}_q$. Following the same deduction as in [8], the MSE of the ML channel estimator is minimized if matrix $\mathbf{A}$ satisfies $\mathbf{A}^H\mathbf{A} = \mathcal{P}\mathbf{I}_{LN_t}$, where $\mathcal{P}$ is the power dedicated for training. The minimum MSE is given by

$$\text{MSE}(\hat{\mathbf{h}}_q) = \frac{\sigma_n^2}{\mathcal{P}}. \tag{18}$$

Therefore we have the following design criterion:

Criterion 1. The optimal training sequences for channel estimation must satisfy $\mathbf{A}^H\mathbf{A} = \mathcal{P}\mathbf{I}_{LN_t}$.

Remarks

- Criterion 1 implies that the TS's of different transmitted antennas must be orthogonal to each other regardless of the FO. Furthermore, each TS should be orthogonal to its circular shifted sequence, (i.e. ideal autocorrelation). If zero signal is allowed in the MIMO system, $\mathbf{A}$ can be chosen that in each column only one entry is non-zero and the zero position is different in different columns. Hence, at any time, only one transmit antenna is active.
- In practice, if zero signal is not allowed, TS must satisfy $\mathbf{A}^H\mathbf{A} = \mathcal{P}\mathbf{I}_{LN_t}$, which is very difficult to find with long length. Note that similar condition arises in the selection of spreading sequence for CDMA systems. The theory of spreading code design can be applied here. This suggests the use of m-sequences and gold sequences.

B. Training sequence for frequency offset estimation

The performance analysis of the ML FO estimator is hard. However, using [3], we can find the expectation and variance of the ML estimator (9) in the high SNR region as

$$E[\hat{v}_q] \approx v_q - \frac{E[\dot{g}(v_q)]}{E[\ddot{g}(v_q)]},\ \text{var}[\hat{v}_q] \approx \frac{E\{[\dot{g}(v_q)]^2\}}{E^2[\ddot{g}(v_q)]} \tag{19}$$

where $\dot{g}(v_q)$ and $\ddot{g}(v_q)$ denote the first and second derivatives of $g(\hat{v}_q)$ at $\hat{v}_q = v_q$. The estimate is asymptotically unbiased or $E[\hat{v}_q] = v_q$ and

$$\text{var}[\hat{v}_q] = \frac{\sigma_n^2}{8\pi^2\text{tr}\left\{\mathbf{A}^H\mathbf{M}^H\left(\mathbf{I}_N - \mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\right)\mathbf{M}\mathbf{A}\mathbf{R}_{\mathbf{h}_q}\right\}} \tag{20}$$

where $\mathbf{M}$ is a diagonal matrix given by

$$\mathbf{M} = \text{diag}\{0, 1, 2, \cdots, N-1\}. \tag{21}$$

Furthermore, the Cramér-Rao Bound (CRB) [6] for the estimation of v_q is

$$\text{CRB} = -\frac{\sigma_n^2}{4\pi^2\text{tr}\left(\mathbf{A}^H\mathbf{D}^{(2)}\mathbf{A}\mathbf{R}_{\mathbf{h}_q}\right)} \tag{22}$$

where

$$\mathbf{D}^{(2)} = 2\mathbf{M}\mathbf{G}\mathbf{M} - \mathbf{G}\mathbf{M}^2 - \mathbf{M}^2\mathbf{G}. \tag{23}$$

Eqs. (20) and (22) show that the variance and the CRB are equal, which suggests that our proposed FO estimator satisfies the asymptotic efficiency property of the ML estimator. From (20) the MSE depends only on the TS and $\mathbf{R}_{\mathbf{h}_q}$, which suggests a second design criterion. In a MIMO system, the optimal TSs must be designed by minimizing the average $\text{var}[\hat{v}_q]$ over all the receive antennas, which means the optimal TS must achieve the minimum average MSE. This leads to the following design criterion

Criterion 2. The optimal training sequences for frequency offset estimation must satisfy

$$\mathbf{A} = \arg\max_{\mathbf{A}} \frac{1}{N_r} \sum_{q=1}^{N_r} \mathrm{tr}\left\{ \mathbf{A}^H \mathbf{M}^H \left(\mathbf{I}_N - \frac{1}{N} \mathbf{A}\mathbf{A}^H \right) \mathbf{M}\mathbf{A}\mathbf{R}_{\mathbf{h}_q} \right\} \tag{24}$$

Remarks

- Since it is impossible to maximizing the MSEs of FO estimation and CE at the same time, the optimal TS's for joint FO and CIR estimation can be found by searching the TS's that meet Criterion 1 first, and then choosing the TS's that meet (24).
- For SISO systems with one transmit and one receive antenna and frequency selective fading channels, the second design criterion becomes

$$\mathbf{A} = \arg\max_{\mathbf{A}} \mathrm{tr}\left\{ \mathbf{A}^H \mathbf{M}^H \left(\mathbf{I}_N - \frac{1}{N} \mathbf{A}\mathbf{A}^H \right) \mathbf{M}\mathbf{A}\mathbf{R}_{\mathbf{h}_q} \right\} \tag{25}$$

C. Training sequence length design

The length of the TS is a critical parameter. The TS needs to be long enough for the channel to be identified. Eq. (16) is valid if and only if $\mathbf{A}^H\mathbf{A}$ is invertible. The TS matrix $\mathbf{A}$ must have full column rank. Therefore we have

Identifiability: The necessary condition for $\mathbf{A}$ to have full column rank and (16) to be identifiable is $N \geq LN_t$.

In practical systems, the transmit power within each data frame is fixed. We assume fixed transmit energy $\mathcal{E}$ in each frame. The energy allocated to information data becomes $\mathcal{E} - \mathcal{P}$. An increase in $\mathcal{P}$ increase the channel estimation accuracy and, thus, decrease the output error probability. On the other hand, an increase of the pilot energy will decrease the effective data SNR and thus, the output error probability will increase. Therefore, an optimal $\mathcal{P}$ exists given a fixed SNR and channel length L. The analysis of the impact of CIR and FO estimation error on performance is beyond the scope of this paper. Assuming all the pilots have equal power, the optimal N can be found via simulation.

V. Simulation

In our simulations, we assume that the channel between each transmit and receive antenna has the same power delay profile and i.i.d. Gaussian. A 3-ray (L=3), exponential power delay profile (EPDP) channel model is used, which can be expressed as

$$E\left[|h_{t,q}(l)|^2\right] = \frac{1-\exp\left(-\frac{1}{5}\right)}{1-\exp\left(-\frac{L}{5}\right)} \exp\left(-\frac{l}{5}\right) \tag{26}$$

where $l = 0, 1, \ldots, L-1, t = 1, 2, \ldots, N_t, q = 1, 2, \ldots, N_r$. The TS and data symbols are both modulated by quadrature phase-shift keying (QPSK). The performance of FO estimator (9), denoted by MLE, is compared with the FO estimator MLE#1 in [3], dentoed by MME. The optimal data detection for the frequency selective SISO channel is accomplished via the sphere decoding (SD) algorithm of [9].

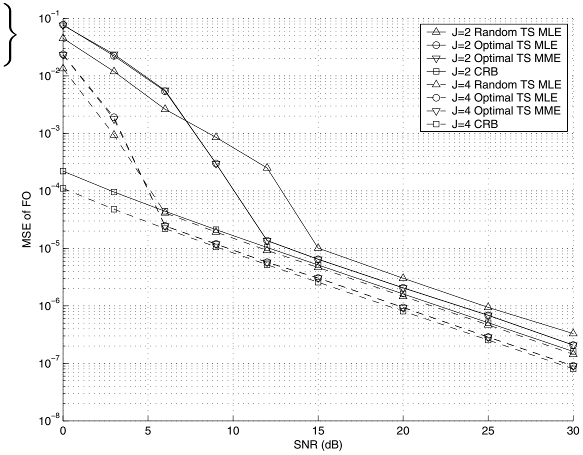

Fig. 1. The MSE of FO versus SNR with different J in a SISO channel

A. SISO Test

We test the performance of our derived sequences in SISO system ($N_t = N_r = 1$). We select $N = 8$. By using the two design criteria, we find there are 7 optimal TSs. We select $[1+j, -1-j, 1+j, 1+j, -1-j, 1+j, 1+j, 1+j]$. We compare the performance of optimal TS with that of randomly selected TS. Figs. 1 and 2 show the average MSE of CIR and FO, respectively. We find the optimal TS can have 2dB gain over random TS at MSE=10^{-6} for FO estimation (Fig. 1) and 2.5dB gain at MSE=10^{-3} for CIR estimation (Fig. 2) with $J = 2$. Increasing the number of frames used for estimation, all the estimators perform better and the thresholds of MME [3] and MLE can be reduced. The proposed MLE can asymptotically achieve the FO CRB. In low SNR region, our proposed MLE is slightly better than the MME. They have the same performance in high SNR region. From Fig. 1 we also find the random TS has better performance than that of optimal TS. This is because we optimize the TS by using the asymptotical variance.

Fig. 3 compares the BER of different estimators. The system with perfect CIR and FO is used as a benchmark. Since the optimal TS is designed for the high SNR region, the performance of random TS is better than that of optimal TS in the low SNR region. The optimal TS has 0.5dB gain over the random TS at BER=10^{-2} while it has 1.5dB loss compared with the benchmark. This paper designs TS by minimizing MSE of estimator. Such 1.5dB loss suggests optimal TS design for the BER performance.

B. MIMO test

We select $N_t = N_r = 2$, $L = 2$ and $N = 12$. Similarly, the optimal TS performs better than random TS in high SNR region. MLE and MME have almost the same performance in all SNR. The gap between optimal TS and random TS is

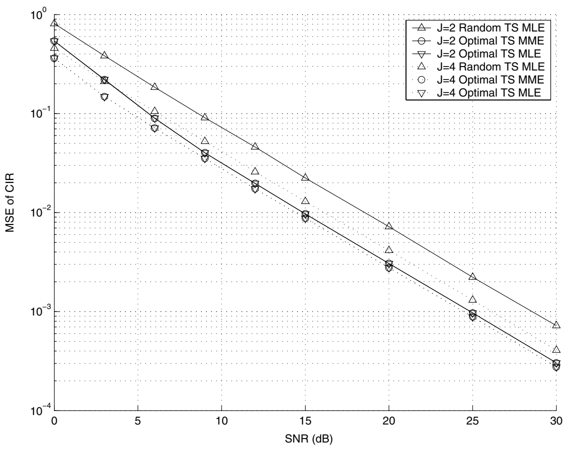

Fig. 2. The MSE of CIR versus SNR with different J in a SISO channel

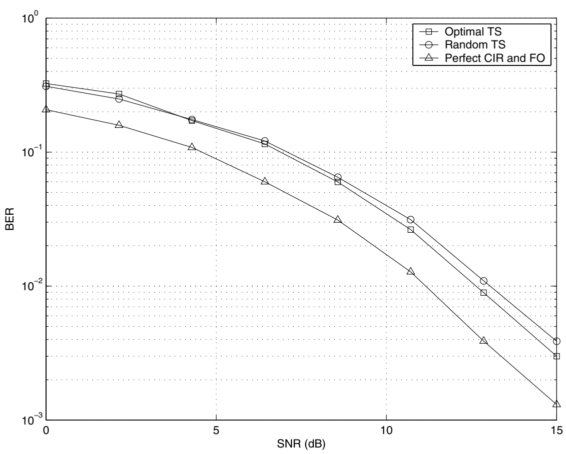

Fig. 3. BER versus SNR with randomly selected TS and optimal TS in a SISO channel

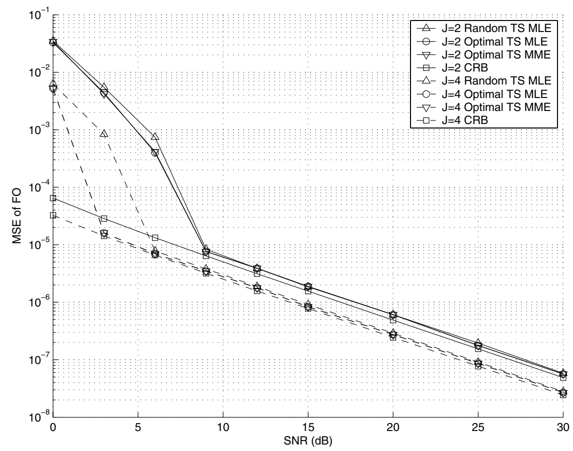

Fig. 4. The MSE of FO versus SNR with different J in a MIMO channel

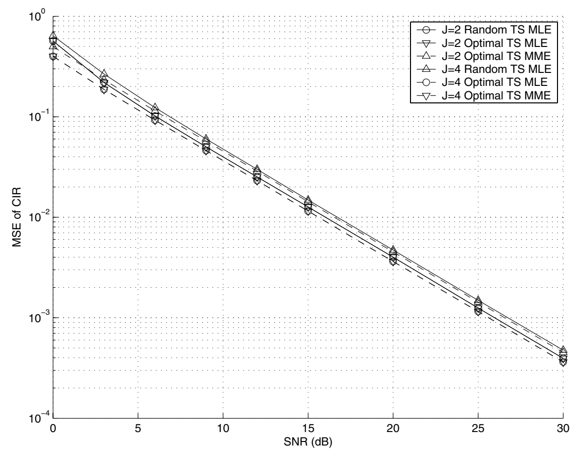

Fig. 5. The MSE of CIR versus SNR with different J in a MIMO channel

0.3dB at MSE=10^{-7} for FO estimation (Fig. 4) and 0.8dB at MSE=10^{-3} for CIR estimation (Fig. 5). The estimator's accuracy is improved in a MIMO system compared with a SISO system.

VI. Conclusion

In this paper, we have addressed the FO and CIR estimation for a MIMO frequency selective fading channel. We propose a novel FO estimator independent of the channel response and analyze the performance of our ML estimator. Moveover, two design criteria are derived for TS which will minimize the MSE of the estimated parameters. Performance results show that the proposed TSs perform better than randomly selected TSs in MIMO system.

References

[1] A. F. Naguib, N. Seshadri, and A. R. Calderbank, "Increasing data rate over wireless channels," *IEEE Signal Processing Magazine*, vol. 17, no. 3, pp. 76–92, May. 2000.

[2] C. Tellambura, M. Parker, Y. Guo, S. Shepherd, and S. Barton, "Optimal sequences for channel estimation using discrete fourier transform techniques," *IEEE Trans. on Commun.*, vol. 47, pp. 230–238, Feb. 1999.

[3] M. Morelli and U. Mengali, "Carrier-frequency estimation for transmissions over selective channels," *IEEE Trans. on Commun.*, vol. 48, no. 9, pp. 1580–1589, Sep. 2000.

[4] P. Stoica and O. Besson, "Training sequence design for frequency offset and frequency-selective channel estimation," *IEEE Trans. on Commun.*, vol. 51, no. 11, pp. 1910 – 1917, Nov. 2003.

[5] E. Chiavaccini and G. M. Vitetta, "Maximum-likelihood frequency recovery for OFDM signals transmitted over multipath fading channels," *IEEE Trans. on Commun.*, vol. 52, no. 2, pp. 244 – 251, Feb. 2004.

[6] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation and Time Series Analysis*, 1st ed., R. Roberts, Ed. Addison Wesley, 1991.

[7] D. Rife and R. Boorstyn, "Single tone parameter estimation from discrete-time observations," *IEEE Trans. on Inform. Theory*, vol. 20, no. 5, pp. 591 – 598, Sept. 1974.

[8] T.-L. Tung, K. Yao, and R. Hudson, "Channel estimation and adaptive power allocation for performance and capacity improvement of multiple-antenna OFDM systems," in *Proc. of SPAWC*, Mar. 2001, pp. 82–85.

[9] U. Fincke and M. Pohst, "Improved methods for calculating vectors of short length in a lattice, inlcuding a complexity analysis," *Math. Computation*, vol. 44, pp. 463–471, Apr. 1985.