Channel Estimation for MIMO OFDM beamforming Systems

***Najoua Achoura*† and *Ridha Bouallegue*††**

National Engineering School of Tunis, National Engineering School of Sousse

Summary
Transmit beamforming with receive combining is a simple method for exploiting the significant diversity provided by multiple-input multiple-output (MIMO) systems, and the use of orthogonal frequency division multiplexing (OFDM) enables low complexity implementation of this scheme over frequency selective MIMO channels. A major challenge to MIMO-OFDM systems is how to obtain the channel state information accurately and promptly for coherent detection of information symbols and channel synchronization. This paper describes channel estimation and tracking scheme for MIMO-OFDM systems based on pilot tones.

Key words:
MIMO, OFDM, beamforming, CSI, estimation

1.Introduction

An increasing interest has been put on smart antennas technology for exploiting the capacity of the scattered wireless channel. Classically, this kind of techniques have been defined for the receiver, however in last years, it has been considered the possibility of using arrays of antennas also at the transmitter, configuring a Multi-Input-Multi-Output (MIMO) channel. Simultaneously, the Orthogonal Frequency Division Multiplexing (OFDM) has been successfully proposed for many communications and broadcasting systems because of its easy modulation/demodulation and equalization processes. Both MIMO and OFDM are suitable for the new user requirements and multimedia applications, which demand higher bit rates, capacity and Quality of Service.
In this paper we analyze the combination of MIMO systems with OFDM. For this kind of structures different approaches are possible. In this paper, joint beamforming, consisting in the extension to the transmitter side of the classical receive beamforming, is used.
Typical ways of identifying the channel utilize several blocks that consist completely of training symbols for SISO and MIMO systems. In such systems the CSI is first estimated prior to any transmission of data. A retraining block is transmitted when the CSI changes significantly [5]. In rapidly time varying environment, such systems must continuously re-train to re-estimate the CSI. Between re-training these systems experience increased BER due to their outdated estimates of the channel. In this paper channel estimation and tracking scheme based on pilot tones is proposed.
Channel estimation for MIMO-OFDM systems is widely studied. Two estimators are usually considered: LS and MMSE .We present a comparison between these two estimators for the MIMO-OFDM system using beamforming.
This paper is organized as follows. In Section 2 we briefly overview the basic system model .Section 3 we introduce channel estimation; the analysis of the channel estimator is derived in section 4 and 5 through this analysis optimal training sequences is derived. Finally in section 6 some simulation results and conclusions are presented.

2.System Models

Adaptive Beamforming is a technique in which an array of antennas is exploited to achieve maximum reception in a specified direction by estimating the signal arrival from a desired direction (in the presence of noise) while signals of the same frequency from other directions are rejected [2]. This is achieved by varying the weights of each of the antennas used in the array. This spatial separation is exploited to separate the desired signal from the interfering signals. In adaptive beamforming the optimum weights are iteratively computed using complex algorithms based upon different criteria.
The communication over a frequency selective MIMO channel with N_T transmits and N_R receive antenna can be represented in multi-carrier fashion (see Fig. 1) as:

$$y_k = H_k s_k + n_k \qquad 1 \le k \le N \tag{1}$$

Where k denotes the carrier index , N is the number of carriers, $s_k \in \mathbb{C}^{n_T \times 1}$ is the transmitted vector , $H_k \in \mathbb{C}^{n_R \times n_T}$ is the channel matrix, $y_k \in \mathbb{C}^{n_R \times 1}$ is the received signal vector, and $n_k \in \mathbb{C}^{n_R \times 1}$ is a zero-mean circularly symmetric complex Gaussian noise vector with arbitrary covariance matrix R_k .Since transmit beamforming is used at each carrier, the transmitted signal is:

$$s_k = b_k x_k \qquad 1 \le k \le N \tag{2}$$

Where b_k is the transmit beamvector and x_k is the transmitted symbol at the k^{th} carrier [2].The receiver also uses beamforming:

$$\hat{x}_k = a_k^H y_k \qquad 1 \le k \le N \tag{3}$$

Where $a_k \in \mathbb{C}^{n_R \times 1}$ is the receive beamvector and x_k is estimated symbol at the k^{th} carrier. The transmitted is constrained in its average transmit power as:

$$\sum_{k=1}^{N} E\left\{\left\|s_k\right\|^2\right\} = \sum_{k=1}^{N}\left\|b_k\right\|^2 \le P_t \tag{4}$$

where P_T is the power in units of energy per block-transmission.

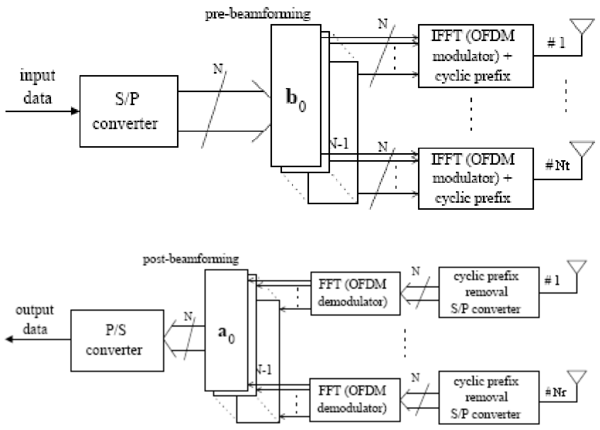

Fig 1: Emission and reception scheme

3.Channel Estimation

The two basic channel estimations in OFDM systems are illustrated in Figure 2 [3]. The first one, block-type pilot channel estimation, is developed under the assumption of slow fading channel, and it is performed by inserting pilot tones into all subcarriers of OFDM symbols within a specific period. The second one, comb-type pilot channel estimation, is introduced to satisfy the need for equalizing when the channel changes even from one OFDM block to the subsequent one. It is thus performed by inserting pilot tones into certain subcarriers of each OFDM symbol, where the interpolation is needed to estimate the conditions of data subcarriers.

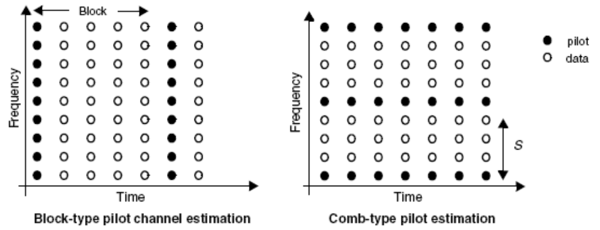

Fig 2: Comparison between two types

MIMO channel requires estimation of Nt × Nr sub-channels, moreover this great number, difficulties also is in inter-symbols interferences generating by Nt emission antennas. A simple solution to this problem is to estimate, in succession different emission antenna channels. Nr channels for each emission antenna are estimated using pilot sequence so other antenna don't send anything. Despite its simplicity, this technique has many problems. First, decreasing efficiency when antenna number increases. The result is important intolerable delay for wireless systems especially for joint beamforming systems where channel estimation in reception must be sent to emission part .moreover such schema needs synchronisation .Another proposition consists in sending OFDM pilot sequences simultaneously and to modulate some sub-carriers (see figure 3).For instance, in a system with Nt=3, each emission antenna modulate one carrier per three .So orthogonality between subcarriers is maintained .for LS estimation case, we must send Nt blocs at least either interpolate modulated carriers .is to send simultaneously blocs modulating all carriers using pilot symbols interfering as little as possible

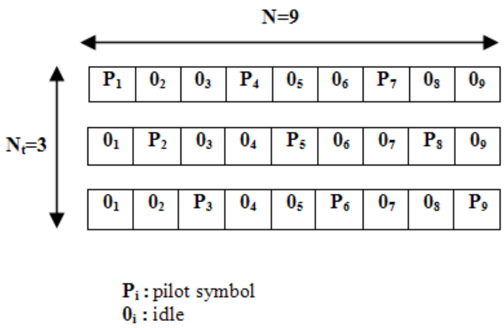

Fig 3: pilot symbols

4.Estimation LS

The received block at the q^{th} antenna after removing the cyclic prefix is given by:

$$y^q(n) = \sum_{p=1}^{N_t} D^{q,p}(n) F^H s^p(n) + \eta^q(n) \tag{6}$$

Where $D^{q,p}$ is channel matrix with first column given by $\left[h^{q,p^T} \ 0_{1\times N-L} \right]^T$ where $h^{q,p}(n) = \left[h_0^{q,p}(n) \cdots h_{L-1}^{q,p}(n) \right]^T$ is the channel impulse response between the p[th] transmit antenna and the q[th] receive antenna. $s^p(n) = \left[s_0^p(n) \cdots s_{N-1}^p(n) \right]^T$ is the transmitted vector from the p[th] transmit antenna at time index n. F is the N× N FFT matrix (normalized columns)[4]. Taking the FFT of $y^q(n)$ we obtain:

$$Y^q(n) = \sum_{p=1}^{N_t} H^{q,p}(n) s^p(n) + \Xi^q(n) \qquad (7)$$

Where $H^{q,p}(n)$ is a diagonal matrix with the frequency response of channel as its diagonal, and $\Xi^q(n)$ is the frequency domain noise.

Let $s^p(n) = e^p(n) + b^p(n)$ where $e^p(n)$ is some arbitrary $N\times 1$ data vector and $b^p(n)$ is some arbitrary $N\times 1$ training sequence vector. Substituting $s^p(n)$ in (7) we obtain:

$$Y^p(n) = \sum_{p=1}^{N_t} (diag\left\{e^p(n)\right\} + diag\left\{b^p(n)\right\}) \times F_{0:L-1} h^{q,p}(n) + \Xi^q(n) \qquad (8)$$

where $F_{0:L-1}$ is the first L columns of F .Defining $E_d^p(n) = diag\left\{e^p(n)\right\}$ equation (8) can be written as:

$$Y^q(n) = \sum_{p=1}^{N_t} E_d^p(n) F_{0:L-1} h^{q,p}(n) + \sum_{p=1}^{N_t} B_d^p(n) F_{0:L-1} h^{q,p}(n) + \Xi^q(n) \qquad (9)$$

Define $A = \left[B_d^1 \ F_{0:L-1} \cdots B_d^{N_t} \ F_{0:L-1} \right]$ and

$$\hat{h}^q(n) = A^\dagger Y^q(n) \qquad (10)$$

where the pseudo inverse matrix $A^? = (A^H A)^- \ A$

Using (9) we can obtain the following:

$$h^q(n) = h^q(n) + A^? \sum_{p=1}^{N_t} S_d^p(n) F_{0:L-1} h^{q\ p}(n) + A\ \Xi^q(n) \qquad (11)$$

where $h^q(n) = \left[h^{q,1^T}(n) \cdots h^{q,N_t^T}(n) \right]^T$ .By imposing the following condition :

$$A^\dagger S_q^p(n) = 0 \quad for \quad p = 1, \cdots, N_t \qquad (12)$$

we obtain:

$$\hat{h}^q(n) = h^q(n) + A^\dagger \Xi^q(n) \qquad (13)$$

Equation (13) indicates that $\hat{h}^q(n)$ is a combination of the true $h^q(n)$ plus a term affected only by the noise in the system.

Satisfying (12) implies $B_d^{r^H}(n) E_d^p(n) = 0$ for any $r, p = 1 \cdots N_t$ .for such a scheme where both data and training are embedded in the same OFDM block, the only way of satisfying this is by choosing disjoint sets of tones for training and data, i.e. zeros in $b^p(n)$ where $e^p(n)$ contains data and vice versa. This allows us to write (5) in simplified form: $\hat{h}^p(n) = \tilde{A}^\dagger \tilde{Y}^q(n)$ where $\tilde{A} = \left[\tilde{B}_d^1 \tilde{F} \cdots \tilde{B}_d^{N_t} \tilde{F} \right]$, $\tilde{B}_d^p = diag\left\{ nonzero\ entries\ of\ B^p \right\}$

$\tilde{F}$, $\tilde{Y}^q(n)$ and $\tilde{\Xi}^q(n)$ are the corresponding rows of $F_{0:L-1}$, $Y^q(n)$ and $\Xi^q(n)$ respectively. Then the channel estimate becomes:

$$\hat{h}^q(n) = h^q(n) + \tilde{A}^\dagger \tilde{\Xi}^q(n) \qquad (14)$$

- *Optimal Training Sequences*

Let P be the number of pilot tones required for training, with the necessary condition of identifiability $P \geq LN_t$ .Let $\left\{k_0, k_1, \ldots, k_{P-1}\right\}$ be the set of P tones used for training [1]. Then $\tilde{F}$ can be written as $\tilde{F} = \left[f_0, \ldots, f_{L-1} \right]$ where $f_l = \left[W^{lk_0} \quad W^{lk_1} \cdots W^{lk_{P-1}} \right]^T$ and $W = e^{-j\frac{2\pi}{N}}$

$\tilde{A}^H \tilde{A}$ can be written as:

$$\tilde{A}^H \tilde{A} = \begin{bmatrix} \tilde{F}^H \tilde{B}_d^{1H} \tilde{B}_d^1 \tilde{F} \cdots \tilde{F}^H \tilde{B}_d^{1H} \tilde{B}_d^{N_t} \tilde{F} \\ \vdots \qquad \ddots \qquad \vdots \\ \tilde{F}^H \tilde{B}_d^{N_t H} \tilde{B}_d^1 \tilde{F} \cdots \tilde{F}^H \tilde{B}_d^{N_t H} \tilde{B}_d^{N_t} \tilde{F} \end{bmatrix} \qquad (15)$$

Define

$$C_{ij} = \tilde{F}^H \tilde{B}_d^{iH} \tilde{B}_d^{jH} \tilde{F}^H \qquad (16)$$

As the $(i,j)^{th}$ $L\times L$ sub- matrix of $\tilde{A}^H \tilde{A}$. For optimal channel estimation we require $\tilde{A}^H \tilde{A}$ to be diagonal i.e.,

$$C_{ij} = \begin{cases} cI_P & for\ i = j \\ 0_{L\times L} & for\ i \neq j \end{cases} \qquad (17)$$

Case i=j Assuming constant modulus training sequences ($\tilde{B}_d^{iH} \tilde{B}_d^j = I_P$) we obtain from (16):

$$C_{ii} = \tilde{F}^H \tilde{F} \qquad (18)$$

The $(r,s)^{th}$ entry of the sub-matrix C_{ii} (denoted by $\left[C_{ii}\right]_{r,s}$) can be written by : $\left[C_{ii}\right]_{r,s} = f_r^H f_s$ which is equal to:

$$\left[C_{ii}\right]_{r,s} = \begin{cases} P & for\ r = s \\ \sum_{p=0}^{P} W^{k_p(s-r)} & for\ r \neq s \end{cases} \qquad (19)$$

From (19), the only set of tones that achieves the diagonality of C_{ii} is $\left\{k_p\right\}_{p=0}^{P-1}$ with $k_p = p_0 + pV$ where p_0 is some reference $\in \left\{0, \cdots, V-1\right\}$ and $\frac{PV(s-r)}{N} \in \mathbb{Z}, \frac{V(s-r)}{N} \notin \mathbb{Z}$ $\forall s, r \in \left\{0, \cdots, L-1\right\}$

For a minimum number of pilot tones and a maximum spacing $PV = N$ where $V = \frac{N}{P}$ .Note that the set of equispaced tones is the optimal tone set that achieves the first part of equation (16) under the assumption that constant modulus training sequences are used

Case i≠j We now investigate the conditions imposed by the second part of equation (16). We assume maximum spaced pilot tones that satisfy the first part of equation (16). The $(r,s)^{th}$ entry of C_{ii} can be written as

$$\left[C_{ij}\right]_{r,s} = W^{p_0(s-r)} \sum_{p=1}^{P} \left[\tilde{B}_d^{i^*} \tilde{B}_d^j\right]_{p,p} W^{(s-r)pV}$$

$$= W^{p_0(s-r)} \sum_{p=1}^{P} \left[\tilde{B}_d^{i^*} \tilde{B}_d^j\right]_{p,p} e^{-j\frac{2\pi}{N}pV(s-r)} \quad (20)$$

Substituting for $V = \frac{N}{P}$ in equation (16), we obtain:

$$\left[C_{ij}\right]_{r,s} = W^{p_0(s-r)} \sum_{p=1}^{P} \left[\tilde{B}_d^{i^*} \tilde{B}_d^j\right]_{p,p} e^{-j\frac{2\pi}{P}p(s-r)} \quad (21)$$

Considering the case when s=r; we see that equation (21) imposes the constraint of orthogonality on the training sequences. However, this is not sufficient, since we must also count for the case when $s \neq r$ .Since $s-r \in \{-L+1,\cdots,L-1\}$ we require $B_d^{i^H} D_\phi B_d^i = 0$ $\forall \phi \in \{-L+1,\cdots,L-1\}$ where D_ϕ represents a phase shift matrix given by

$$D_\phi = diag\left\{\left[1 \quad e^{-j\frac{2\pi}{P}\phi} \cdots e^{-j\frac{2\pi(P-1)}{P}\phi}\right]\right\} \quad (22)$$

In other words, in MIMO systems the training sequences on different transmit antennas must not only be orthogonal but also phase shift orthogonal for phase shifts in the range $\{-L+1,\cdots,L-1\}$ This phase shift orthogonality in frequency corresponds to shift orthogonality in time. In a system with multipath propagation, the training sequence of one antenna must not be only orthogonal to the training sequences of other antennas but also to delayed copies of these sequences.

An optimal training sequence can be designed as follows:

$$\left[\tilde{B}_d^{i^*} \tilde{B}_d^i\right]_{p,p} = e^{-j\frac{2\pi}{P}n_{ij}p} \quad (23)$$

where $\left[\tilde{B}_d^{i^*}\right]_{p,p} = e^{-j\frac{2\pi}{P}n_i p}$ and $n_{ij} = n_j - n_i$ with $n_{i,j} \in \mathbb{Z}$

Hence the second part of equation (16) is satisfied if $\frac{n_{ij}+s-r}{P} \notin \mathbb{Z} \ \forall s,r \in \{0,\cdots,L-1\}, \forall i,j \in \{1,\cdots,N_t\}$ with $i \neq j$ One possible choice is $n_i = (i-1)L \ \forall i,j \in \{1,\cdots,N_t\}$

5.Estimation MMSE

The MMSE estimator employs the second-order statistics of the channel conditions to minimize the mean-square error.

Denote by R_{gg} , R_{HH} , and R_{YY} the autocovariance matrix of g , H , and Y, respectively, and by R_{gY} the cross covariance matrix between g and Y. with g , Y are the sampled channel impulse and the output data of DFT block at the receiver respectively, and H is defined as the DFR of g. Also denote by σ_N^2 the noise variance $E\{|N|^2\}$. Assume the channel vector g and the noise N are uncorrelated, it is derived that:

$$R_{HH} = E\{HH^H\} = E\{(Fg)(Fg)^H\} = FR_{gg}F^H$$

$$R_{gY} = E\{gY^H\} = E\{g(XFg+N)^H\} = R_{gg}F^H X^H$$

$$R_{YY} = E\{YY^H\} = XFR_{gg}F^H X^H + \sigma_N^2 I_N$$

Assume R_{gg} (thus R_{HH}) and σ_N^2 are known at the receiver in advance, the MMSE estimator of g is given by [6]. At last, it is calculated

$$\hat{H}_{MMSE} = F\hat{g}_{MMSE} = F\left[\left(F^H X^H\right)^{-1} R_{gg}^{-1}\sigma_N^2 + XF\right]^{-1} Y$$

$$= FR_{gg}\left[\left(F^H X^H XF\right)^{-1}\sigma_N^2 + R_{gg}\right] F^{-1}\hat{H}_{LS}$$

$$= R_{HH}\left[R_{HH} + \sigma_N^2\left(XX^H\right)^{-1}\right]^{-1}\hat{H}_{LS}$$

The MMSE estimator yields much better performance than LS estimators, especially under the low SNR scenarios. A major drawback of the MMSE estimator is its high computational complexity, especially if matrix inversions are needed each time the data in changes.

6.Simulations and Discussions:

In this section, we evaluate the performance of estimation method presented in the previous section. For this, we define the mean square error (MSE) as:

$$MSE = \frac{1}{N_t \times N_r \times N} \sum_{i=1}^{N_r} \sum_{j=1}^{N_t} \sum_{k=1}^{N} \left\| H_{i,j}(k) - \hat{H}_{i,j}(k) \right\|^2$$

We consider only the case when all training sequences are dedicated for estimation i.e. there is no data blocs.

We compare estimation's result for three types of sequence: random sequences (RAND), orthogonal phase shift sequences (OPS) as defined in previous section, and sequences without interferences (WI) consisting at transmitting one symbol per carrier to avoid interferences. (See Figure 4)

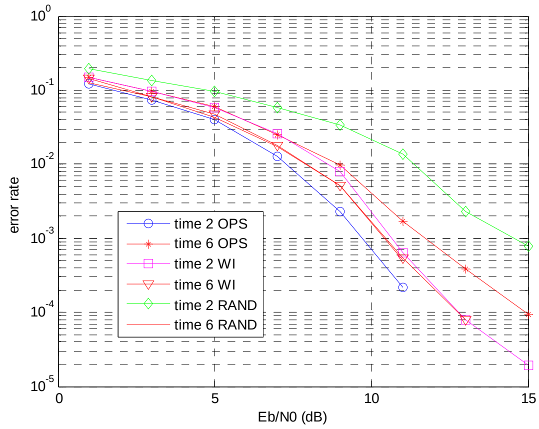

Fig 4: comparison of three type pilot for time estimation

We remark that WI and OPS sequences have almost the same performance. Same results are obtained if we have sequence training number N_t times less. Two reasons can explain this: first, the total absence of interference between different antenna .Second, the power constraint so power for WI sequences is N_t more important then OPS case .As we can expect random sequences are the worst. However, this difference is reduced as well as bloc's number increase. Thus, we see in the case of 6 training blocs difference between curves is negligible.

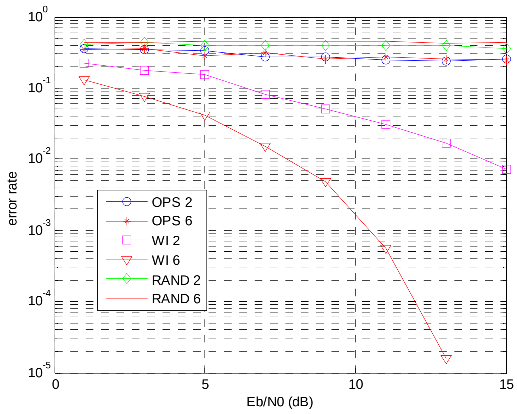

Fig 5: comparison of three type pilot for frequency estimation

Figure 5 show performance of frequency estimation, first we remark non efficiency of random sequences .in other hand, orthogonality of OPS sequences provide a better gain but not acceptable performances. Logically, SI sequences are the most performant thus they can't reach quality of temporal estimation with RAND and OPS sequences.

- *Sub-carrier number*

We have just seen that temporal estimation is better than frequency one .However, this affirmation is not always true .In fact, figure 6 shows performance of a system for N=16 using WI sequences .We can see that frequency estimation is better then temporal one (L=21) this is true while L>N .For OPS sequences, temporal estimation still the best .Generally, in OFDM systems we have L << N that why we prefer using temporal estimation.

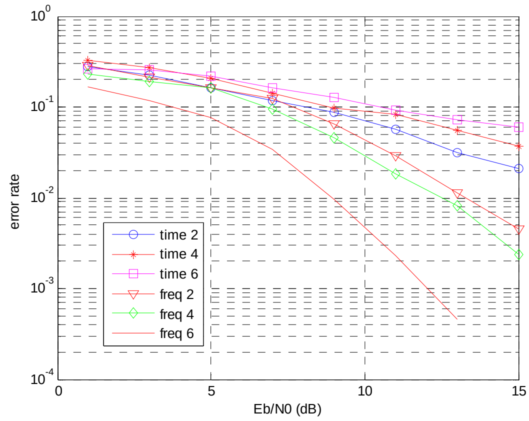

Fig 6: comparison for type WI, N=16

- *Estimator comparison*

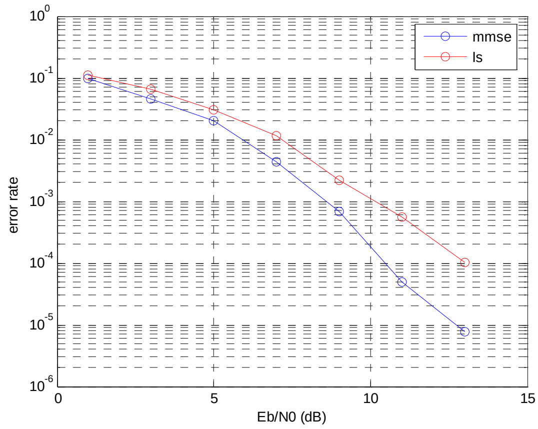

Fig 7: LS and MMSE comparison, Nt=Nr=2, frequency estimation, OPS pilot, N=64

Figure 7 shows that Linear Minimum Mean Squared Error estimator based on statical channel proprieties performs better then Least Square one .Nevertheless, this second method is less complicated and more robust against imperfect knowledge of channel parameters.

References

[1] I. Barhumi, G. Leus, and M. Moonen, "Optimal training sequences for channel estimation in OFDM systems with transmitter diversity in mobile wireless channels," Proceedings of International Zurich Seminar, February 2002.

[2] A.P. Iserte, A.I. Perez-Neira, D.P. Palomar, and M.A. Lagunas, "Power allocation techniques for joint beamforming in OFDM-MIMO channels," Proceedings EUSIPCO 2002, (Toulouse, France), September 2002.

[3] Y.Shen , Ed. Martinez « Channel Estimation in OFDM Systems » Freescale Semiconductor, Application Note Rev. 0,1.2006

[4] Auer, G., "Channel Estimation in Two Dimensions for OFDM Systems with Multiple Transmit Antennas," *GLOBECOM*, pp. 322–326, 2003.

[5] Hlaing Minn ,Naofal Al-Dhahir ,Yinghui Li, « Optimal Training Signals for MIMO OFDM Channel Estimation in the Presence of Frequency Offset and Phase Noise », IEEE Transactions on communications,vol. 54, no. 10, october 2006

[6] Coleri, S., Ergen, M., Puri, A., and Bahai, A., "Channel Estimation Techniques Based on Pilot Arrangement in OFDM Systems," IEEE Transactions on Broadcasting, vol. 48, pp. 223–229, Sept. 2002.