TURBO ESTIMATION OF CHANNEL AND SYMBOLS IN PRECODED MIMO SYSTEMS

Anna Scaglione and Azadeh Vosoughi

School of Electrical and Computer Engineering Cornell University, Ithaca, NY 14853 anna@ece.cornell.edu;av78@cornell.edu

ABSTRACT

We consider a block fading frequency selective multi-input multioutput (MIMO) channel in additive white Gaussian noise (AWGN). The channel input is a training vector *superimposed* on a linearly precoded vector of Gaussian symbols. To achieve a better performance over the conventional least-squares (LS), we utilize the linear mean square error (LMMSE) symbol estimate to improve the initial LS estimate and update the symbol estimation accordingly. We provide the guidelines to design training which minimizes the MSE of the initial LS estimate.

1. INTRODUCTION

Precoding is an effective tool to facilitate estimation and equalization for time and/or frequency selective channels. In fact, as shown in [1] and the references therein, the redundancy introduced by the precoder can be exploited to guarantee the symbol recovery in the absence of noise and to blindly estimate deterministic frequency selective single-input single-output (SISO) channels up to a scale. To acquire the channel state information (CSI) without ambiguity training is required.

Superimposed training technique which consists in adding a known training sequence to the unknown data sequence at the transmitter, is utilized in [2] for the purpose of channel estimation. By construction, there is no increase in the bandwidth. Affine precoding schemes [3] enjoy the benefits of both linear precoding and training.

For frequency selective SISO channels [5] designed affine precoders which *decouple* channel estimation from symbol detection and optimized the LS channel estimator. The class of affine precoders (F, t) which admit the decoupling, satisfies a special form of orthogonality [5].

For block fading frequency selective MIMO channels, [4] we derived the CRB and we found that satisfying an orthogonality constraint analogous to [5] is indeed a *sufficient* condition to reduce the CRB. After imposing the orthogonality constraint to the CRB, the CRB is reduced to two terms where the first term depends on the training t, whereas the second one depends on the precoder F, but not on the training. As we will show in this paper, there is a substantial gap between the CRB and the LS estimate of the channel. The interesting results shown in the following are that by using an "orthogonal design" for (F, t) w can fill the gap in the estimation performance at SNR $\gg 1$ by iteratively performing symbol estimations and channel estimation in a *turbo* fashion. However, such gains are not attained if the design of (F, t) does

not satisfy the orthogonality condition. The importance of the orthogonality condition motivates us to investigate designs that satisfy it. These designs have not been studied for the MIMO channel.

Notation: Boldface upper and lower cases denote matrices and column vectors respectively. The tr(A) is trace of A. The column vector formed by stacking vertically the columns of A is a = vec(A). I_M is the identity matrix of size M. Δ_{ii} is a $K \times K$ matrix whose all elements are zeros, except the *ii*-th element which is one. \otimes is the Kronecker product. The (i, j) entry of A is indicated with $[A]_{ij}$. diag(a) is a diagonal matrix whose diagonal elements are the components of a. Complex conjugate, Hermitian, transpose, pseudoinverse and expectation operations are represented by $(.)^*$, $(.)^H$, $(.)^T$, $(.)^{\dagger}$, $E\{.\}$ respectively.

2. SYSTEM MODEL

The system considered has K transmit and R receive antennas. We assume a block fading model where P is the coherence time of the channel, i.e., the equivalent discrete-time impulse response of the channel does not change during the transmission of P snapshots. In our setup The channel has finite memory L. The information sequence s[n] is parsed into blocks of size N, namely s_i . Each block is precoded by a tall $KP \times N$ precoding matrix **F**. An $KP \times 1$ training vector t, which is known to the receiver, is added to the precoding block Fs_i to obtain the transmitted data block $x_i =$ $Fs_i + t$. The $PK \times 1$ vector x_i is obtained by stacking P transmit snapshots $\boldsymbol{x}_i := vec([\boldsymbol{x}[iP], \dots, \boldsymbol{x}[iP+P-1]]))$, where $\boldsymbol{x}[iP+p]$ is the $K \times 1$ (coded) symbol vector, emitted by the K transmit antennas. For the mapping from s_i to x_i to be invertible, we require F to be full column rank. Stacking M = P - L received snapshots in an $MR \times 1$ vector $\boldsymbol{y}_i := vec([\boldsymbol{y}[iP+L], \dots, \boldsymbol{y}[iP+P-1]]))$, in which we eliminated the first L vectors to cancel the inter-block interference (IBI), we obtain $\boldsymbol{y}_i = \boldsymbol{H}\boldsymbol{x}_i$, where \boldsymbol{H} is an $RM \times KP$ block Toeplitz matrix with block components $\mathbf{H}[l]$ $l = 0, \dots, L$ where $\{\mathbf{H}[l]\}_{r,k}$ is the *l*-th sample of the impulse response characterizing the channel between the k-th transmitter and the r-th receiver. We require $N \leq \operatorname{rank}(\boldsymbol{H}) \leq \min(KP, RM)$. The received signal is $\boldsymbol{z}_i = \boldsymbol{y}_i + \boldsymbol{n}_i$ where $\boldsymbol{n}_i \sim \mathcal{N}(\boldsymbol{0}, \sigma_{nn}^2 \boldsymbol{I}_{RM})$. Furthermore, we assume that $s_i \sim \mathcal{N}(\mathbf{0}, \sigma_{ss}^2 \boldsymbol{I}_N)$. and \boldsymbol{n}_i and \boldsymbol{s}_i are uncorrelated. Combining all, we obtain $z_i = HFs_i + Ht + n_i$. Without IBI, we assume that the resulting channel estimator operates on a block-by-block basis and we omit the block index *i*. We let $h \in C^{KR(L+1)}$ be the vector containing the channel coefficients we wish to estimate $\mathbf{h} := vec ([\mathbf{H}[0], \dots, \mathbf{H}[L]]^T)$. To simplify the CRB derivation we rewrite Hx explicitly as a function of h. To this end we introduce a mapping $\Phi: x \to X$ such

This work is supported in part by the National Science Foundation under grant CCR-0133635.

that Hx = Xh where $X := \Phi(x)$ is defined as:

$$\mathbf{X} := \begin{bmatrix} \mathbf{I}_R \otimes \boldsymbol{\mathcal{X}}_{(1,:)} \\ \mathbf{I}_R \otimes \boldsymbol{\mathcal{X}}_{(2,:)} \\ \vdots \\ \mathbf{I}_R \otimes \boldsymbol{\mathcal{X}}_{(M,:)} \end{bmatrix}$$
(1)

in which $\mathcal{X}_{(i,:)}$ is the *i*-th row of the $M \times K(L+1)$ block Toeplitz structure matrix \mathcal{X} :

$$\boldsymbol{\mathcal{X}} := \begin{bmatrix} \boldsymbol{x}^{T}[L] & \cdots & \boldsymbol{x}^{T}[0] \\ \boldsymbol{x}^{T}[L+1] & \cdots & \boldsymbol{x}^{T}[1] \\ \vdots & \ddots & \vdots \\ \boldsymbol{x}^{T}[P-1] & \cdots & \boldsymbol{x}^{T}[M-1] \end{bmatrix}$$
(2)

3. DETERMINISTIC CHANNEL CRB EXPRESSION

The channel CRB is the inverse of the complex Fisher information matrix \mathcal{J}_{c} .

Lemma 1 [4] \mathcal{J}_c is given by:

$$\boldsymbol{\mathcal{J}}_{c} = \mathbf{T}^{H} \boldsymbol{R}_{z}^{-1} \mathbf{T} + \sigma_{ss}^{4} \boldsymbol{\mathcal{E}}^{H} (\boldsymbol{D}^{T} \otimes \boldsymbol{R}_{z}^{-1}) \boldsymbol{\mathcal{E}}$$
(3)

$$\mathbf{T} = \boldsymbol{\Phi}(t)$$

$$\mathbf{R}_{z} = \sigma_{ss}^{2} \boldsymbol{H} \boldsymbol{F} \boldsymbol{F}^{H} \boldsymbol{H}^{H} + \sigma_{nn}^{2} \boldsymbol{I}_{RM}$$

$$\boldsymbol{D} = \boldsymbol{F} \boldsymbol{F}^{H} \boldsymbol{H}^{H} \boldsymbol{R}_{z}^{-1} \boldsymbol{H} \boldsymbol{F} \boldsymbol{F}^{H}$$

$$\boldsymbol{\mathcal{E}} = \left[vec(\boldsymbol{E}_{1}), \dots, vec(\boldsymbol{E}_{(L+1)RK}) \right]$$

$$\boldsymbol{E}_{i} = \frac{\partial \boldsymbol{H}}{\partial h_{i}} \quad i = 1, \dots, (L+1)RK$$

Regarding the structure of \mathcal{J}_c , we can make two remarks:

Remark $1 : \mathbf{T}^{H} \mathbf{R}_{z}^{-1} \mathbf{T}$ can be read as the inverse of the error covariance of the LS channel estimation when the precoded symbols' contribution is *purely* increasing the noise level.

Remark 2: $\mathcal{E}^{H}(\mathbf{D}^{T} \otimes \mathbf{R}_{z}^{-1})\mathcal{E}$, which is independent of the training, is the contribution coming from the blind estimation component. Indeed, this term quantifies the gain one can achieve by extracting the channel information embedded in the term HFs. The presence of this term testifies to the *sub-optimality* of the LS channel estimation which is based on the training symbols only [5], specially when this term is dominating, which occurs whenever most of the transmit signal energy is spent on the data.

4. DESIGN GUIDELINES: CRB CRITERION

We look for a family of (F, t) which provide us a lower $tr(\mathcal{J}_c^{-1})$ than the others, *independent* of the underlying FIR MIMO channel.

Lemma 2 [4] The pair (F, t) which satisfy the following orthogonality constraint provides a lower $tr(\mathcal{J}_c^{-1})$:

$$\boldsymbol{\mathcal{T}}^{H}\boldsymbol{\mathcal{F}}_{i} = \boldsymbol{0} \quad i = 1, 2, \dots, N \tag{4}$$

where \mathcal{T} and \mathcal{F}_i are defined according to (1)-(2) using \mathbf{t} and \mathbf{f}_i (the *i*-th column of \mathbf{F}). Under (4) \mathcal{J}_c reduces to:

$$\boldsymbol{\mathcal{J}}_{c} = \sigma_{nn}^{-2} \, \mathbf{T}^{H} \mathbf{T} + \sigma_{ss}^{4} \, \boldsymbol{\mathcal{E}}^{H} (\boldsymbol{D}^{T} \otimes \boldsymbol{R}_{z}^{-1}) \boldsymbol{\mathcal{E}}$$
(5)

We wish to characterize the pairs (F, t) which satisfy (4).We select our general design such that the matrix F and the vector t incorporate cyclic prefix (CP) [4]. We define $F := \Psi \overline{F}$, $t := \Psi \overline{t}$ where Ψ is the CP-inducing matrix. Under the CP assumption, \mathcal{T} and \mathcal{F}_i s will be block circulant matrices and can be *diagonalized* using the fast Fourier transform (FFT) matrix.

4.1. Design of Affine Precoding With CP

Let $W := \exp(j2\pi/M)$, U be the $M \times M$ FFT matrix with $[U]_{mn} = M^{-1/2}W^{-(m-1)(n-1)}$ and $U_{0:L}$ denote the first L+1 columns of U. We can express \mathcal{T} as $\mathcal{T} = U^H \Delta_{\overline{t}}(U_{0:L} \otimes I_K)$ where $\Delta_{\overline{t}}$ is an $M \times MK$ block diagonal matrix defined as $\Delta_{\overline{t}} := diag[\tilde{t}^T(1), \tilde{t}^T(W), \cdots, \tilde{t}^T(W^{M-1})]$ and $\tilde{t}(z) := \sum_{m=0}^{M-1} \overline{t}[m]z^{-m}$ is the \mathcal{Z} transform of \overline{t} . Similarly, we can establish the FFT-based diagonalization of \mathcal{F}_i and define $\tilde{f}_i(z) := \sum_{m=0}^{M-1} \overline{f}_i[m]z^{-m}$.

Lemma 3 A sufficient condition for (4) to be satisfied is that $\tilde{t}(W^m)\tilde{f}_i^H(W^m) = 0$ for $m = 0, 1, \dots, M-1$ and $i = 1, \dots, N$.

Suppose $\Delta_{\overline{t}}$ has $m_0^* \leq M$ nonzero block entries. Let \mathcal{I}_0 be the set of *ordered* indices containing these m_0^* indices [5]:

$$\mathcal{I}_0 := \left\{ q_n | \; ilde{m{t}}(W^q)
eq m{0}, \; q_n < q_{n+1}, \; n = 0, \dots, m_0^* - 1
ight\}$$

The design in Lemma 3 offers degrees of freedom to optimize m_0^* , $\tilde{t}(W^m)$, the fraction of the power allocated to the training ζ (where $\zeta := ||t||^2 / \mathcal{P}$ and $\mathcal{P} = ||x||^2 = \sigma_{ss}^2 tr(FF^H) + ||t||^2$ is the total transmit power and is constant), and the position of the pilot tones via \mathcal{I}_0 . In Section 6 we exploit these degrees of freedom to design *t* which minimizes the MSE of an initial LS channel estimator (Section 5).

Remark: For the frequency selective SISO channels, the constraint in Lemma 3 is necessary and sufficient to satisfy (4) [5]. Note that for the MIMO it is only a sufficient condition. In Section 7 we investigate whether the transmit diversity can be exploited to design $(\overline{F}, \overline{t})$ which have common tones and still they maintain the orthogonality constraint in (4).

5. TURBO CHANNEL AND SYMBOL ESTIMATION

To build the estimator, we confine ourselves to the affine precoders with CP. Under the orthogonality constraint in (4) we have $\mathbf{T}^{\dagger}\mathbf{F}_{i} = \mathbf{0}$ for i = 1, ..., N. Hence an *initial* LS channel estimator is given by $\hat{\mathbf{h}}_{0} = \mathbf{T}^{\dagger}\mathbf{z}$ where the MSE expression for the LS estimate is:

$$E\{||\hat{\boldsymbol{h}}_0 - \boldsymbol{h}||^2\} = tr(\mathbf{T}^{\dagger}\mathbf{T}^{\dagger H}) = tr\left((\mathbf{T}^H\mathbf{T})^{-1}\right)$$
(6)

(6) is minimized if and only if $\mathbf{T}^{H}\mathbf{T} = c_{0}\boldsymbol{I}_{RK(L+1)}$ or equivalently $\boldsymbol{\mathcal{T}}^{H}\boldsymbol{\mathcal{T}} = c_{0}\boldsymbol{I}_{K(L+1)}$ for some nonzero constant $c_{0} < \zeta$ (in particular $\zeta = c_{0} + \sum_{l=0}^{L-1} ||\boldsymbol{t}[l]||$). The LLMSE symbol estimator is given by $\hat{\boldsymbol{s}} = \boldsymbol{R}_{sz}\boldsymbol{R}_{zz}^{-1}\boldsymbol{z}$ where $\boldsymbol{R}_{sz} = \sigma_{ss}^{2}\boldsymbol{F}^{H}\hat{\boldsymbol{H}}^{H}$. Assuming that $\mathbf{T}^{H}\mathbf{T} = c_{0}\boldsymbol{I}_{RK(L+1)}$, \boldsymbol{s} and $\hat{\boldsymbol{h}}$ are uncorrelated, and $E\{\tilde{\boldsymbol{h}}\boldsymbol{n}^{H}\} = \sigma_{nn}^{2}\mathbf{T}^{\dagger}, \boldsymbol{R}_{zz}$ is:

$$\begin{aligned} \boldsymbol{R}_{zz} &= \hat{\boldsymbol{H}}(\sigma_{ss}^{2}\boldsymbol{F}\boldsymbol{F}^{H} + \boldsymbol{t}\boldsymbol{t}^{H})\hat{\boldsymbol{H}}^{H} \\ &+ c_{0}^{-1}(\sigma_{ss}^{2}\sum_{i=1}^{N}\boldsymbol{F}_{i}\boldsymbol{F}_{i}^{H} + (1 + 2\sigma_{nn}^{2})\boldsymbol{T}\boldsymbol{T}^{H})) + \sigma_{nn}^{2}\boldsymbol{I} \end{aligned}$$

The LLMSE symbol estimate can be incorporated in to the LS channel estimate. Suppose the symbol estimation error is sufficiently small. We may approximate the received vector as $\boldsymbol{z} \approx \mathcal{A}\boldsymbol{h} + \boldsymbol{n}$ where $\mathcal{A} := \left(\sum_{i=1}^{N} \mathbf{F}_i \hat{\boldsymbol{s}}(i) + \mathbf{T}\right)$ from which we can obtain an improved LS channel estimate $\hat{\boldsymbol{h}}_1 = \mathcal{A}^{\dagger}\boldsymbol{z}$. The channel estimation followed by symbol estimation can be performed *iteratively* to obtain enhanced symbol recovery. Note that $\boldsymbol{s}(i)$ s are Gaussian RVs and $\hat{\boldsymbol{s}}(i)$ s are soft estimates, hence this method is not a decision directed method.

6. OPTIMAL DESIGN FOR PILOT TONES

In this Section, we provide the guidelines to design t such that T or equivalently T is semi-unitary.

Lemma 4 Let $\Gamma_{\overline{t}}$ be the $m_0^* \times K$ matrix whose rows are the nonzero block entries of $\Delta_{\overline{t}}$. \mathcal{T} is full column rank if and only if (i) $M > m_0^* \geq K(L+1)$, where the minimum redundancy corresponds to $m_0^* = K(L+1)$, and (ii) $\Gamma_{\overline{t}}$ is full column rank.

Let $\Theta(m) := \tilde{\boldsymbol{t}}^*(W^m) \tilde{\boldsymbol{t}}^T(W^m) \ m \in \mathcal{I}_0$ be an $K \times K$ matrix. Substituting the FFT based diagonalized expression for $\boldsymbol{\mathcal{T}}$, we can rewrite $\boldsymbol{\mathcal{T}}^H \boldsymbol{\mathcal{T}}$ as the following:

$$\begin{bmatrix} \sum_{m \in \mathcal{I}_0} \Theta(m) & \dots & \sum_{m \in \mathcal{I}_0} W^{-mL} \Theta(m) \\ \sum_{m \in \mathcal{I}_0} W^m \Theta(m) & \dots & \sum_{m \in \mathcal{I}_0} W^{-m(L-1)} \Theta(m) \\ \vdots & \ddots & \vdots \\ \sum_{m \in \mathcal{I}_0} W^{mL} \Theta(m) & \dots & \sum_{m \in \mathcal{I}_0} \Theta(m) \end{bmatrix}$$
(7)

If $\Theta(m) = I \forall m \in \mathcal{I}_0$ were to hold, the set of conditions established in Lemma 5 would have made $\mathcal{T}^H \mathcal{T}$ identity.

Lemma 5 $\mathcal{T}^{H}\mathcal{T}$ is identity if and only if we select:

- $M = m_0^* Q$ where Q is an integer.
- the spacing of the pilot tones to satisfy $q_n = q_0 + Qp$ ($q_n \in \mathcal{I}_0$) for some integer $q_0 \in [0, Q 1]$ and integer $p \in [0, m_0^* 1]$.

Under Lemma 5 the block components of $\mathcal{T}^{H}\mathcal{T}$ reduces to:

$$\sum_{m \in \mathcal{I}_0} W^{ml} \mathbf{\Theta}(m) = \sum_{n=0}^{m_0^* - 1} w^{nl} \mathbf{\Theta}(n) \quad l \in [-L, L]$$
(8)

where $w = e^{j2\pi/m_0^*}$ and $\Theta(n) := \tilde{t}^*(w^n)\tilde{t}^T(w^n)$. However $\Theta(m) \neq I$, indeed its rank is one. We assume $\tilde{t}(W^m) \ m \in \mathcal{I}_0$ to be a $K \times 1$ canonical vector, e.g., at each pilot tone *only* one transmitter transmits the training. To achieve the minimum redundancy we assume $m_0^* = K(L+1)$. We divide the m_0^* pilot tones into K groups, namely $\mathcal{G}_1, \ldots, \mathcal{G}_K$, such that each group \mathcal{G}_k contains L+1 tones. Let \mathcal{I}_k be the set of ordered indices of the tones which belong to \mathcal{G}_k :

$$\mathcal{I}_k := \left\{ u_n^{(k)} | w^n \in \mathcal{G}_k, \ u_n^{(k)} < u_{n+1}^{(k)}, \ n = 0, \dots, L \right\}$$

Furthermore, suppose the spacing of the tones belonging to \mathcal{G}_k satisfy $u_n^{(k)} = u_0^{(k)} + Kp$ for some integer $u_0^{(k)} \in [0, K-1]$ and integer $p \in [0, L]$. We can simplify (8) to:

$$\sum_{n=0}^{m_0^*-1} w^{nl} \boldsymbol{\Theta}(n) = \sum_{k=1}^K \boldsymbol{\Delta}_{kk} \sum_{n \in \mathcal{I}_k} w^{nl} = \begin{cases} L \boldsymbol{I}_k & l = 0\\ \boldsymbol{0} & l \neq 0 \end{cases}$$



Fig.1 training design for $m_0^* = 10$ and K = 5

which is the desired result. An alternative design for $\tilde{t}(w^n)$ is:

$$\tilde{\boldsymbol{t}}(\boldsymbol{w}^n) = \begin{bmatrix} 1 \ e^{jn\Omega} \ e^{j2n\Omega} \ \dots \ e^{j(K-1)n\Omega} \end{bmatrix}^T \tag{9}$$

where $\Omega := 2\pi/K$. It is not difficult to verify:

$$\sum_{n=0}^{m_0^*-1} w^{nl} \Theta(n) = m_0^* \delta(l) \boldsymbol{I}_K \quad l = 0, \pm 1, \dots, \pm L$$

The design proposed in (9) has an interesting interpretation which is illustrated through Fig.1 in which we assume $m_0^* = 10$ and K = 5. The x and y axes represent respectively n, the pilot tone index, and $n\Omega$, which we refer to it as spatial frequency. A shaded square indicates that all transmitters transmit a particular pilot tone simultaneously (indexed by n) focusing spatially the beam at angle $n\Omega$, the corresponding spatial frequency. Hence each pilot tone is transmitted at an increasing orthogonal angle. Note that for the canonical design, each pilot tone is transmitted only by one transmitter and the training signal is transmitted over all angles.

7. PRECODER DESIGN

Since Lemma 3 is inconclusive as to wether \overline{F} and \overline{t} can share pilot tones, in this Section we investigate whether it is possible to load \overline{F} and \overline{t} on overlapping subcarriers, while we maintain (*i*) the orthogonality between \overline{F} and \overline{t} [c.f.(4)], (*ii*) the proposed design for the training in (9) (It can be shown that there is a tradeoff between using the canonical design for the training and satisfying the orthogonality constraint when the training and precoder share the pilot tones). We may rewrite the block components of $\mathcal{T}^H \mathcal{F}_i = \mathbf{0}$ as:

$$\sum_{n=0}^{m_0^*-1} w^{nl} \Psi(w^n) = \mathbf{0}_{K \times K} \ l = 0, \pm 1, \dots, \pm L$$
 (10)

where $\Psi(w^n) := \tilde{\boldsymbol{t}}^*(w^n) \tilde{\boldsymbol{f}}_i^T(w^n)$. We choose the $K \times 1$ vector $\tilde{\boldsymbol{f}}_i$ to be such that its k component is $\{\tilde{\boldsymbol{f}}_i\}_k = e^{jn\Omega_{i,k}}$. It can be verified that (10) is satisfied if and only if:

• $\Omega_{i,q} = 2\pi l_{i,q}/m_0^*;$

$$l - (L+1)p + l_{i,q} \neq 0$$

• $-m_0^* < l - (L+1)p + l_{i,q} < m_0^*$

We select $l_{i,q}$ to be positive integer. Therefore the above set of conditions is satisfied if $0 < l_{i,q} < m_0^* - L$. The above argument verifies the existence of an orthogonal design for \overline{F} and \overline{t} in which symbols and training are loaded on overlapping subcarriers.



Fig. 1. MSE of the LS against the channel CRB for the nonorthogonal design



Fig. 2. MSE of the LS against the channel tr(CRB) normalized by the number of channel coefficients as ζ varies. for the orthogonal design

8. NUMERICAL RESULTS

We set K = 2, R = 2, L = 3, $\sigma_{ss}^2 = 1 \sigma_{hh}^2 = 1/(L + 1)$ 1). The simulation results are averaged over 100 sets of independent Rayleigh fading channels. Without loss of generality, we assume that $\mathcal{P} = 1$ and therefore the signal-to-noise ratio (SNR) is $SNR := -10 log_{10} \sigma_{nn}^2$. For each SNR, t and F are scaled such that the power constraint is satisfied. We set M = 24, N = 4. To satisfy the orthogonality constraint in (4) we simply load t and F on non-overlapping subcarriers. In particular, we select the design (*iii*) in [4]. Fig.1,2 compare the MSE of the initial and the improved LS channel estimate against CRB (in dB) as a function of SNR at $\zeta = 0.3$ for two designs namely the orthogonal and the non-orthogonal designs (to obtain the non-orthogonal case we only change $\bar{f}_1(1)$ and $\bar{f}_2(1)$ to e_1 and e_2 respectively). While the effect of violating the orthogonality constraint in not significant for the CRB, the MSE for the LS estimators have floor. Fig.3 compares the performance of the initial and the improved LS against the CRB as a function of ζ at SNR =10dB; whereas Fig.4 shows the symbol MSE (in dB). While by increasing the training power we improve tr(CRB), the symbol MSE achieves its minimum around $\zeta = 0.6.$

In summary we showed that the initial LS channel estimate can



Fig. 3. The performance of the LS and the proposed iterative estimator against CRB



Fig. 4. The symbol MSE in dB

be improved using the LMMSE of the symbols. We proposed two training designs which both minimize the MSE of the LS channel estimator. In MIMO systems, we showed the existence of orthogonal designs for \overline{F} and \overline{t} in which symbols and training are loaded on overlapping subcarriers.

9. REFERENCES

- A. Scaglione, B. Giannakis, S. Barbarossa, "Redundant Filterbank precoders and equalizers part II: blind channel estimation, synchronization, and direct equalization", *IEEE Trans. Sig. Proc.*, Vol. 47, No.7, pp. 2007-2022, July 1999.
- [2] P. Hoeher and F. Tufvesson," Channel estimation with superimposed pilot sequence," *Proc. GLOBECOM Conf.* pp. 2162-2166, Rio de Janeiro, Brazil, 1999.
- [3] J. H. Manton, I. Y. Mareels and Y.Hua, "Affine precoders for reliable communications," *Proc. ICASSP*, vol.5, pp. 2741-2744, Istanbul, Turkey, June 2000.
- [4] A. Vosoughi and A. Scaglione "The Best Training Depends on The Receiver Architecture," Proc. ICASSP Canada, May 2004
- [5] S.Ohno and G. B. Giannakis, "Optimal training and redundant precodering for block transmissions with application to wirlesss OFDM," *IEEE Trans. Communications*, vol.50, No.12, pp. 2113 -2123, Dec. 2002.