TRACE-ORTHOGONAL SPACE-TIME CODING FOR MULTIUSER SYSTEMS

Sergio Barbarossa, Antonio Fasano

INFOCOM Dpt., University of Rome “La Sapienza”, Via Eudossiana 18, 00184 Rome, Italy
E-mail: `<sergio,fasano>@infocom.uniroma1.it`

ABSTRACT

In this paper we prove that trace-orthogonal space-time coding provides a necessary and sufficient condition for information lossless coding in a multiple access system, where each user encodes its own symbols independently of the other users. Then, we show that the sub-optimal MMSE decoder can be implemented very simply as a set of scalar decoders and we prove that, under the assumption of using the (sub-optimal) MMSE decoder, the trace-orthogonal design with scaled unitary coding matrices yields minimum BER for each user.

1. INTRODUCTION

The major challenge in the design of space-time encoders is the search for the (possibly best) trade-off between three fundamental aspects: bit rate, bit error rate, and receiver complexity. Orthogonal space-time coding (OSTC) is known for achieving maximum diversity gain and minimum receiver complexity, but at the expenses of rate [1]. BLAST techniques are known for achieving maximum rate, but at the expenses of diversity. Recently, space-time encoding strategies guaranteeing full rate and full diversity were established in [2] and [3][1]. The detector guaranteeing the full-diversity performance is the maximum likelihood (ML) detector. However, ML for full rate systems can be too complicated to implement, especially when the MIMO dimensions or the symbol constellation size are not too small. Sphere decoding can be used instead, but its complexity is still quite high. In an effort to simplify the detector as much as possible, but still guaranteeing appreciable BER performance, in [4], [5] it was proposed a trace-orthogonal design (TOD). In this paper we consider the multiple access channel and we prove that *necessary and sufficient* condition for having an information lossless transmission consists in encoding the information symbols using TOD, for any given channel realization. Furthermore, we show that, under the assumption of using the (suboptimal) MMSE detector, *TOD with scaled unitary matrices minimizes the BER for each user*.

2. MULTI-USER TRACE-ORTHOGONAL DESIGN

We consider a multiple access system composed of N users, each with n_T transmit antennas, and an access point (AP), with n_R receive antennas. Let us assume that user k encodes its own n_s (complex) symbols $s_k(j), j = 1, \ldots, n_s$, through the following space-time linear encoder:

$$\boldsymbol{X}_k = \sum_{j=1}^{n_s} \boldsymbol{A}_k(j) s_k(j) \tag{1}$$

where $\{\boldsymbol{A}_k(j), j = 1, \ldots, n_s\}$ is the set of $n_T \times Q$ complex matrices assigned to user k. A multi-user (MU) space-time encoder is an MU *Trace-Orthogonal Design* (TOD), if the matrices $\boldsymbol{A}_k(j)$ satisfy, for each user k,

$$\langle \boldsymbol{A}_k(j), \boldsymbol{A}_k(m) \rangle := \operatorname{tr}(\boldsymbol{A}_k^H(j)\boldsymbol{A}_k(m)) = \delta_{jm}, \tag{2}$$

where δ_{jm} denotes the Kronecker symbol. Furthermore, we say that the encoding is a Unitary TOD (UTOD) if the following additional condition holds true

$$\boldsymbol{A}_k(j)\boldsymbol{A}_k^H(j) = \frac{1}{n_T}\boldsymbol{I}_{n_T}, \quad j = 1, \ldots, n_s. \tag{3}$$

The *shift and multiply bases* proposed in [4] are an example of UTOD.

Denoting by $\boldsymbol{H}_k$ the $n_R \times n_T$ channel matrix characterizing the link between the k-th user and the access point (AP), and by $\tilde{\boldsymbol{x}}_k$ the vector transmitted from the k-th user, the received vector, at the AP, is

$$\tilde{\boldsymbol{y}} = \sum_{k=1}^{N} \boldsymbol{H}_k \tilde{\boldsymbol{x}}_k + \tilde{\boldsymbol{v}} := \boldsymbol{\mathcal{H}}\tilde{\boldsymbol{x}} + \tilde{\boldsymbol{v}}, \tag{4}$$

where $\boldsymbol{\mathcal{H}} := [\boldsymbol{H}_1, \ldots, \boldsymbol{H}_N]$, $\tilde{\boldsymbol{x}} := [\tilde{\boldsymbol{x}}_1^T, \ldots, \tilde{\boldsymbol{x}}_N^T]^T$, and $\tilde{\boldsymbol{v}}$ is the noise vector, assumed to be zero mean, circularly symmetric complex Gaussian, with covariance matrix $\sigma_v^2\boldsymbol{I}$. We will refer to (4) as the *uncoded system.*

Let us consider now a system with space-time encoding, where the channels $\boldsymbol{H}_k$ are assumed to be constant over Q successive channel uses (quasi-static fading). If each user transmits the matrix $\boldsymbol{X}_k$, built as in (1), the received matrix is

$$\boldsymbol{Y} = \sum_{k=1}^{N} \boldsymbol{H}_k \boldsymbol{X}_k + \boldsymbol{V} := \boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{X}} + \boldsymbol{V}, \tag{5}$$

where $\boldsymbol{\mathcal{H}}$ is as in (4), $\boldsymbol{\mathcal{X}} := [\boldsymbol{X}_1^T, \ldots, \boldsymbol{X}_N^T]^T$, and $\boldsymbol{V}(n_R \times Q)$ is the received noise matrix. We will refer to (5) as the *coded system.* Applying the $\operatorname{vec}(\cdot)$ operator to (1), we obtain

$$\boldsymbol{x}_k = \operatorname{vec}(\boldsymbol{X}_k) = \sum_{j=1}^{n_s} \operatorname{vec}(\boldsymbol{A}_k(j)) s_k(j) := \boldsymbol{F}_k \boldsymbol{s}_k, \tag{6}$$

where $\boldsymbol{F}_k(Q \cdot n_T \times n_s)$ is the matrix whose j-th column is $\operatorname{vec}(\boldsymbol{A}_k(j))$ and $\boldsymbol{s}_k = [s_k(1) \cdots s_k(n_s)]^T$. To guarantee symbol recovery, the

This work has been supported by the project IST-2001-32549 - ROMANTIK - funded by the European Community.

[1]Maximum rate here is meant to be a symbol rate equal to n_T symbols per channel use (pcu). This should not be confused with the information rate used in [6] to derive the optimal rate-diversity trade-off.

matrices $\boldsymbol{F}_k$ must be full column rank, i.e. $\mathrm{rank}(\boldsymbol{F}_k) = n_s$. This means that the following inequality must be satisfied

$$n_s \leq Q \cdot n_T. \tag{7}$$

Applying the $\mathrm{vec}(\cdot)$ operator[2] to (5), and using (1) and (6), we get

$$\boldsymbol{y} = \mathrm{vec}(\boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{X}}) + \mathrm{vec}(\boldsymbol{V}) = (\boldsymbol{I}_Q \otimes \boldsymbol{\mathcal{H}})\boldsymbol{\Pi}\boldsymbol{\mathcal{F}}\boldsymbol{s} + \boldsymbol{v}, \tag{8}$$

where $\boldsymbol{\Pi}$ is the following permutation matrix

$$\boldsymbol{\Pi} = [\boldsymbol{I}_Q \otimes \boldsymbol{P}_1 \;\; \boldsymbol{I}_Q \otimes \boldsymbol{P}_2 \;\; \cdots \;\; \boldsymbol{I}_Q \otimes \boldsymbol{P}_N], \tag{9}$$

with $\boldsymbol{P}_k$ defined as

$$\boldsymbol{P}_k = \boldsymbol{u}_k \otimes \boldsymbol{I}_{n_T}, \tag{10}$$

where $\boldsymbol{u}_k$ is the k-th unit vector[3] in $\mathbb{C}^N$, that is $\boldsymbol{u}_k(j) = \delta_{kj}$ $(j = 1, \ldots, N)$, and $\boldsymbol{\mathcal{F}}$ is the following block diagonal matrix

$$\boldsymbol{\mathcal{F}} := \begin{pmatrix} \boldsymbol{F}_1 & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{F}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \boldsymbol{0} & \cdots & \cdots & \boldsymbol{F}_N \end{pmatrix} \tag{11}$$

where blocks $\boldsymbol{F}_k$ are defined in (6). Finally, vector $\boldsymbol{s}$ is defined as $\boldsymbol{s} = [\boldsymbol{s}_1^T \cdots \boldsymbol{s}_N^T]^T$.

3. INFORMATION LOSSLESS MULTI-USER CODING

In this section we prove under which conditions the multi-user space-time coded system (5) is information lossless, i.e. it yields the same capacity as (4). More specifically, we prove the following important result:

Theorem 1. *Let us assume that: i) the receiver has perfect knowledge of all channels; ii) each user has no channel knowledge; iii) the transmit power of each user is upper bounded by P_T; iv) different users encode their symbols independently of each other; v) the transmitted symbol vectors $\boldsymbol{s}_k$ have covariance matrix $\frac{P_T}{n_T}\boldsymbol{I}$. Under the previous assumptions, the coded system (5) has the same mutual information as the uncoded system (4), for any channel realization $\boldsymbol{\mathcal{H}}$, if and only if $\boldsymbol{\mathcal{F}}\boldsymbol{\mathcal{F}}^H = \boldsymbol{I}$.*

Proof. (*Sufficiency*) Under the previous assumptions, the aggregate mutual information exchanged between the whole set of users and the receiver of the uncoded system is

$$\mathrm{I}(\tilde{\boldsymbol{x}};\tilde{\boldsymbol{y}}|\boldsymbol{\mathcal{H}}) = \log\left|\boldsymbol{I} + \gamma\boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{H}}^H\right| = \log\left|\boldsymbol{I} + \gamma\boldsymbol{\mathcal{H}}^H\boldsymbol{\mathcal{H}}\right|, \tag{12}$$

where $\gamma = P_T/(n_T\sigma_v^2)$ is the average SNR per receiving antenna. Let us compute now the aggregate mutual information of the coded system (5) or, equivalently, (8), that for any given realization of the channel $\boldsymbol{\mathcal{H}}$, is

$$\mathrm{I}(\boldsymbol{\mathcal{X}};\boldsymbol{Y}|\boldsymbol{\mathcal{H}})^{\mathrm{cod}} = \frac{1}{Q}\log\left|\boldsymbol{I} + \gamma(\boldsymbol{I}\otimes\boldsymbol{\mathcal{H}})\boldsymbol{\Pi}\boldsymbol{\mathcal{F}}\boldsymbol{\mathcal{F}}^H\boldsymbol{\Pi}^H(\boldsymbol{I}\otimes\boldsymbol{\mathcal{H}}^H)\right|, \tag{13}$$

where factor $1/Q$ accounts for the Q channel uses. Since $\boldsymbol{\mathcal{F}}\boldsymbol{\mathcal{F}}^H = \boldsymbol{I}$, (13) becomes (note that $\boldsymbol{\Pi}\boldsymbol{\Pi}^H = \boldsymbol{I}$)

$$\mathrm{I}(\boldsymbol{\mathcal{X}};\boldsymbol{Y}|\boldsymbol{\mathcal{H}})^{\mathrm{cod}} = \frac{1}{Q}\log\left|\boldsymbol{I} + \gamma(\boldsymbol{I}\otimes\boldsymbol{\mathcal{H}})(\boldsymbol{I}\otimes\boldsymbol{\mathcal{H}}^H)\right|. \tag{14}$$

Using the property $(\boldsymbol{A}\otimes\boldsymbol{B})(\boldsymbol{C}\otimes\boldsymbol{D}) = (\boldsymbol{AC}\otimes\boldsymbol{BD})$, (14) becomes

$$\begin{aligned}\mathrm{I}(\boldsymbol{\mathcal{X}};\boldsymbol{Y}|\boldsymbol{\mathcal{H}})^{\mathrm{cod}} &= \frac{1}{Q}\log\left|\boldsymbol{I} + \gamma(\boldsymbol{I}\otimes\boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{H}}^H)\right| \\ &= \frac{1}{Q}\log\left|\boldsymbol{I}\otimes(\boldsymbol{I} + \gamma\boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{H}}^H)\right| = \log\left|\boldsymbol{I} + \gamma\boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{H}}^H\right|\end{aligned} \tag{15}$$

where we used the identity $|\boldsymbol{I}_n \otimes \boldsymbol{M}| = |\boldsymbol{M}|^n$. Comparing (15) with (12), we realize that $\mathrm{I}(\boldsymbol{\mathcal{X}};\boldsymbol{Y}|\boldsymbol{\mathcal{H}})^{\mathrm{cod}} = \mathrm{I}(\tilde{\boldsymbol{x}};\tilde{\boldsymbol{y}}|\boldsymbol{\mathcal{H}})$. Since this identity holds true for *any* channel realization $\boldsymbol{\mathcal{H}}$, this proves the sufficiency statement.

(*Necessity*) Due to space limitations, the proof of necessity will be carried out only in the case $n_R \geq N \cdot n_T$. We start now from

$$\mathrm{I}(\boldsymbol{\mathcal{X}};\boldsymbol{Y}|\boldsymbol{\mathcal{H}})^{\mathrm{cod}} = \mathrm{I}(\tilde{\boldsymbol{x}};\tilde{\boldsymbol{y}}|\boldsymbol{\mathcal{H}}) \qquad \forall\,\boldsymbol{\mathcal{H}} \in \mathbb{C}^{n_R\times N\cdot n_T} \tag{16}$$

From (12) and (13), using the identity $|\boldsymbol{I} + \boldsymbol{AB}| = |\boldsymbol{I} + \boldsymbol{BA}|$, (16) can be rewritten as

$$\left|\boldsymbol{I}_{Q\cdot N\cdot n_T} + \gamma(\boldsymbol{I}_Q\otimes\boldsymbol{\mathcal{H}}^H\boldsymbol{\mathcal{H}})\boldsymbol{\Pi}\boldsymbol{\mathcal{F}}\boldsymbol{\mathcal{F}}^H\boldsymbol{\Pi}^H\right| = \left|\boldsymbol{I}_{N\cdot n_T} + \gamma\boldsymbol{\mathcal{H}}^H\boldsymbol{\mathcal{H}}\right|^Q \tag{17}$$

$\forall\,\boldsymbol{\mathcal{H}} \in \mathbb{C}^{n_R\times N\cdot n_T}$, where the dimensions of the identity matrices involved are explicitly indicated. Since (17) holds for *any* realization of the channel, it is satisfied, in particular, for channel matrices $\hat{\boldsymbol{\mathcal{H}}}_\lambda$ such that

$$\gamma\hat{\boldsymbol{\mathcal{H}}}_\lambda^H\hat{\boldsymbol{\mathcal{H}}}_\lambda = \lambda\boldsymbol{I}_{N\cdot n_T} \qquad \lambda \in \mathbb{R}^+ \tag{18}$$

where $\mathbb{R}^+ \equiv [0, +\infty)$. Note that such realizations always exist since we have assumed $n_R \geq N \cdot n_T$. Substituting $\hat{\boldsymbol{\mathcal{H}}}_\lambda$ in (17), we get

$$\left|\boldsymbol{I}_{Q\cdot N\cdot n_T} + \lambda\boldsymbol{\Pi}\boldsymbol{\mathcal{F}}\boldsymbol{\mathcal{F}}^H\boldsymbol{\Pi}^H\right| = |\boldsymbol{I}_{N\cdot n_T} + \lambda\boldsymbol{I}_{N\cdot n_T}|^Q \quad \lambda \in \mathbb{R}^+. \tag{19}$$

Since $\boldsymbol{\Pi}\boldsymbol{\mathcal{F}}\boldsymbol{\mathcal{F}}^H\boldsymbol{\Pi}^H$ is Hermitian positive semidefinite, let us indicate with φ_k $(k = 1, \ldots, N\cdot Q\cdot n_T)$ its eigenvalues. Note that $\varphi_k \in \mathbb{R}^+$ and this property will be useful later. Hence, (19) can be written

$$\prod_{k=1}^{N\cdot Q\cdot n_T}(1 + \lambda\varphi_k) = (1+\lambda)^{N\cdot Q\cdot n_T} \qquad \lambda \in \mathbb{R}^+. \tag{20}$$

Equation (20) is an identity between polynomials of degree NQn_T in the indeterminate λ (it holds for any $\lambda \in \mathbb{R}^+$). But the two polynomials coincide if and only if the coefficients of the corresponding powers of λ are equal. Using this fact, equating in particular the coefficients of $\lambda^{N\cdot Q\cdot n_T}$ and λ form (20) we get

$$\prod_{k=1}^{N\cdot Q\cdot n_T}\varphi_k = 1 \quad \text{and} \quad \sum_{k=1}^{N\cdot Q\cdot n_T}\varphi_k = N\cdot Q\cdot n_T. \tag{21}$$

Since $\varphi_k \in \mathbb{R}^+$ it is easy to verify[4] that (21) jointly imply

$$\left(\prod_{k=1}^{N\cdot Q\cdot n_T}\varphi_k\right)^{\frac{1}{N\cdot Q\cdot n_T}} = \frac{1}{N\cdot Q\cdot n_T}\sum_{k=1}^{N\cdot Q\cdot n_T}\varphi_k, \tag{22}$$

[2] If $\boldsymbol{A}(r\times t)$, $\boldsymbol{X}(t\times p)$, and $\boldsymbol{B}(p\times s)$ are matrices, we have $\mathrm{vec}(\boldsymbol{AXB}) = (\boldsymbol{B}^T\otimes\boldsymbol{A})\,\mathrm{vec}(\boldsymbol{X})$, where $\otimes$ is the Kronecker product.

[3] It is a column vector.

[4] If $x \in \mathbb{R}^+$ then $x = 1 \iff x^{1/n} = 1$ for $n \in \mathbb{N}$.

but the well know relation between geometric and arithmetic means implies that (22) holds if and only if all the φ_k are equal. This implies, using (21), that $\varphi_1 = \cdots = \varphi_{N\cdot Q\cdot n_T} = 1$, which is equivalent to state[5] that $\boldsymbol{\Pi}\boldsymbol{\mathcal{F}}\boldsymbol{\mathcal{F}}^H\boldsymbol{\Pi}^H = \boldsymbol{I}_{N\cdot Q\cdot n_T}$. Since $\boldsymbol{\Pi}$ is invertible, this last condition is equivalent to

$$\boldsymbol{\mathcal{F}}\boldsymbol{\mathcal{F}}^H = \boldsymbol{I}_{N\cdot Q\cdot n_T}, \tag{23}$$

which proves the necessity for the subclass of channels satisfying (18). But (23) is also the condition that guarantees (16) for *all* channel matrices, as it follows from the proof of sufficiency. Hence, we conclude, a fortiori, that (23) is a necessary condition for all channel matrices. The proof is thus complete. □

Taking into account the structure of $\boldsymbol{\mathcal{F}}$ in (11), it is evident that (23) is equivalent to say

$$\boldsymbol{F}_k\boldsymbol{F}_k^H = \boldsymbol{I}_{Q\cdot n_T}, \qquad k = 1,\ldots,N. \tag{24}$$

Condition (24) holds true only if $\boldsymbol{F}_k(Q\cdot n_T\times n_s)$ is full row rank, that is only if

$$n_s \geq Q\cdot n_T\ . \tag{25}$$

Combining (25) with (7), we arrive at the following equality

$$n_s = Q\cdot n_T, \tag{26}$$

which is equivalent to say that the rate of the code, defined as $R = n_s/Q$ is equal to n_T, for all the users. This means that all users must use a *full-rate* code. Moreover, condition (26) forces $\boldsymbol{F}_k$ to be square and this, together with (24), implies that $\boldsymbol{F}_k$ is unitary. This is a strong result that allows us to fully characterize the encoding matrices $\boldsymbol{A}_k(j)$ $(j = 1,\ldots,n_s)$. In fact, the generic element $\{\boldsymbol{F}_k^H\boldsymbol{F}_k\}_{i,j}$ of the matrix product $\boldsymbol{F}_k^H\boldsymbol{F}_k$ can be written as

$$\mathrm{vec}^H(\boldsymbol{A}_k(i))\,\mathrm{vec}(\boldsymbol{A}_k(j)) = \mathrm{tr}(\boldsymbol{A}_k^H(i)\boldsymbol{A}_k(j)) = \delta_{ij} \tag{27}$$

where $\mathrm{vec}^H(\boldsymbol{A})\,\mathrm{vec}(\boldsymbol{B}) = \mathrm{tr}(\boldsymbol{A}^H\boldsymbol{B})$ was used. The last equality leads to the trace-orthogonality condition (2).

Thus, combining (26) and (27) we can summarize the results of this section in the following statement: *a space-time coding strategy for multiple access systems is information lossless if and only if each user uses a full-rate Trace-Orthogonal Design.*

4. MMSE DECODING

Maximum Likelihood decoding is the optimal procedure, but it suffers from exponential complexity. Sphere-Decoding could be used, but its complexity is still quite high. It is then important to analyze the performance of sub-optimal decoders. We consider the minimum mean square error (MMSE) receiver, which derives first a soft estimate of the transmitted symbols and then it takes a hard decision, based on those estimates. In the sequel, we will show that TOD leads to very simple scalar MMSE decoding.

We derive the MMSE estimator from vector model (8) assuming that $\boldsymbol{s}\sim\mathcal{CN}(\boldsymbol{0},\sigma_s^2\boldsymbol{I})$, $\boldsymbol{v}\sim\mathcal{CN}(\boldsymbol{0},\sigma_v^2\boldsymbol{I})$, and noise and signals are independent. Under these assumptions, the MMSE estimator is linear and assumes the following simple expression

$$\hat{s}_k(j) = \mathrm{tr}\left(\boldsymbol{A}_k^H(j)\,\boldsymbol{P}_k^H\,\boldsymbol{\mathcal{W}}\,\boldsymbol{Y}\right), \tag{28}$$

[5]The identity matrix is the only Hermitian matrix with all the eigenvalues equal to 1.

where $\hat{s}_k(j)$ is the estimate of the j-th symbol transmitted by the k-th user and $\boldsymbol{P}_k$ is defined as in (10); the matrix $\boldsymbol{\mathcal{W}}$ is the same for all the users and is equal to

$$\boldsymbol{\mathcal{W}} = \left(\boldsymbol{\mathcal{H}}^H\boldsymbol{\mathcal{H}} + \frac{1}{\gamma}\boldsymbol{I}_{N\cdot n_T}\right)^{-1}\boldsymbol{\mathcal{H}}^H, \tag{29}$$

with $\gamma = \sigma_s^2/\sigma_v^2$. The role of matrix $\boldsymbol{P}_k$ is to extract from $\boldsymbol{\mathcal{W}}$ the subset of rows that is relevant for the k-th user. In particular, using Matlab notation, $\boldsymbol{P}_k\boldsymbol{\mathcal{W}} = \boldsymbol{\mathcal{W}}\left((k-1)\cdot n_T + 1 : k\cdot n_T, :\right)$.
The main feature of (28) is low complexity. In fact, if we neglect the computation of the inverse in (29), which has to be done only once every channel coherence time, the complexity of the estimator is comparable to the ML detector used for orthogonal design [1], but with the advantage of Trace-Orthogonal Design of being full rate. The price paid for having this low complexity is that MMSE is a sub-optimal detector for TOD.

Let us consider now the BER performance of the MMSE detector, in case of BPSK modulation. We prove that TOD, with unitary matrices, minimizes the BER. Similar results were proven, in the single user case, in [7], but considering the MSE averaged over the channel realizations. Here, we do not take any channel average and we prove that TOD with unitary matrices minimizes the BER *for any channel realization and for every user*.

The MMSE estimate is affected by inter-symbol interference (ISI) and multiuser interference (MUI) and thus the exact BER derivation is not simple. However, invoking the central limit theorem, when n_s and/or N are sufficiently large, one can get a fairly good approximation of the final BER by modeling ISI and MUI as additive complex Gaussian noise. Within the limit of such an approximation the error probability for the j-th symbol of the k-th user, conditioned on the channel realization, can be expressed as

$$P_{\epsilon,j}^{(k)} = \frac{1}{2}\,\mathrm{erfc}\left(\sqrt{\frac{\mathrm{SINR}_j^{(k)}}{2}}\right), \tag{30}$$

where $\mathrm{SINR}_j^{(k)}$ is the signal-to-interference-plus-noise ratio on the j-th symbol for the k-th user, defined as

$$\mathrm{SINR}_j^{(k)} = \frac{\sigma^2_{signal_j^{(k)}}}{\sigma^2_{int_j^{(k)}+noise_j^{(k)}}} \tag{31}$$

where $\sigma^2_{signal_j^{(k)}}$ is the variance of the useful component in $\hat{s}_k(j)$, whereas $\sigma^2_{int_j^{(k)}+noise_j^{(k)}}$ is the variance of the ISI, MUI and noise contained in $\hat{s}_k(j)$. Considering now the average probability of error for the k-th user

$$\overline{P_\epsilon^{(k)}} = \frac{1}{2n_s}\sum_{j=1}^{n_s}\mathrm{erfc}\left(\sqrt{\frac{\mathrm{SINR}_j^{(k)}}{2}}\right), \tag{32}$$

it is interesting to investigate if there is any subclass of TOD that minimizes (32). With this objective in mind, let us consider the covariance matrix of the estimation errors for the k-th user, that is

$$\boldsymbol{K}_\epsilon^{(k)} = \sigma_s^2\left[\boldsymbol{I}_{n_s} - \boldsymbol{F}_k^H(\boldsymbol{I}_Q\otimes\boldsymbol{P}_k^H\boldsymbol{\mathcal{W}}\boldsymbol{\mathcal{H}}\boldsymbol{P}_k)\boldsymbol{F}_k\right], \tag{33}$$

where $\mathcal{W}$ is defined in (29). In (33), the diagonal entries are the mean square errors $\sigma_{\epsilon,j}^{2(k)}$ $(j = 1, \ldots, n_s)$ and are given by

$$\sigma_{\epsilon,j}^{2(k)} = \sigma_s^2 - \sigma_s^2 \operatorname{tr}\left(\boldsymbol{A}_k^H(j)\boldsymbol{P}_k^H \mathcal{W}\mathcal{H}\boldsymbol{P}_k \boldsymbol{A}_k(j)\right). \quad (34)$$

Note that $\sigma_{\epsilon,j}^{2(k)}$ depends only on the encoding matrix with the same indices k and j. It is possible to prove that $\sigma_{\epsilon,j}^{2(k)}$ is related to $\mathrm{SINR}_j^{(k)}$ through the following relation

$$\mathrm{SINR}_j^{(k)} = \frac{\sigma_s^2}{\sigma_{\epsilon,j}^{2(k)}} - 1 . \quad (35)$$

From (33), exploiting again the identity $\boldsymbol{F}_k\boldsymbol{F}_k^H = \boldsymbol{I}$, the sum of all MSE's is

$$\sum_{j=1}^{n_s} \sigma_{\epsilon,j}^{2(k)} = \operatorname{tr}(\boldsymbol{K}_\epsilon^{(k)}) = \sigma_s^2 \left[n_s - Q \cdot \operatorname{tr}(\boldsymbol{P}_k^H \mathcal{W}\mathcal{H}\boldsymbol{P}_k)\right]. \quad (36)$$

The last term in (36) does not depend on the encoding matrices. Thus, (36) shows the invariance of the sum of all MSE's with respect to the choice of the matrices $\boldsymbol{A}_k(j)$, provided that TOD is used. This result is important since allows us to characterize the minimizers for (32). In fact, substituting (35) in (32), we get

$$\overline{P_\epsilon^{(k)}} = \frac{1}{2n_s}\sum_{j=1}^{n_s} \operatorname{erfc}\left(\sqrt{\frac{\sigma_s^2}{2\sigma_{\epsilon,j}^{2(k)}} - \frac{1}{2}}\right). \quad (37)$$

Let us consider now the function $\operatorname{erfc}\left(\sqrt{\frac{\sigma_s^2}{2x} - \frac{1}{2}}\right)$ that appears in (37). This is a convex function of x. Hence, applying Jensen's inequality to (37) leads to

$$\overline{P_\epsilon^{(k)}} \geq \frac{1}{2}\operatorname{erfc}\left(\sqrt{\frac{\sigma_s^2}{\frac{2}{n_s}\sum_{j=1}^{n_s}\sigma_{\epsilon,j}^{2(k)}} - \frac{1}{2}}\right), \quad (38)$$

with equality if and only if all the $\sigma_{\epsilon,j}^{2(k)}$ (for $j = 1, \ldots, n_s$) are equal. But, thanks to (36), the term in the denominator is constant over the encoding matrices $\boldsymbol{A}_k(j)$. This implies that, for each channel realization, (38) is the minimum achievable average BER for the k-th user. In particular it is reached if and only if $\sigma_{\epsilon,1}^{2(k)} = \cdots = \sigma_{\epsilon,n_s}^{2(k)}$, that is, if and only if (see (34)), for any realization of the channel

$$\operatorname{tr}\left(\boldsymbol{P}_k^H \mathcal{W}\mathcal{H}\boldsymbol{P}_k\boldsymbol{A}_k(j)\boldsymbol{A}_k^H(j)\right) = C_k \quad j = 1, \ldots, n_s \quad (39)$$

where C_k is a constant independent of j. It is possible to prove that a necessary and sufficient condition for (39) to be true, for any channel realization, is

$$\boldsymbol{A}_k(1)\boldsymbol{A}_k^H(1) = \cdots = \boldsymbol{A}_k(n_s)\boldsymbol{A}_k^H(n_s) = \frac{1}{n_T}\boldsymbol{I}_{n_T}, \quad (40)$$

that is, the encoding is a Unitary TOD. In particular, (40) is satisfied if $\boldsymbol{A}_k(j)$ $(j = 1, \ldots, n_s)$ are scaled unitary matrices.
Since MMSE is clearly sub-optimal, in Fig.1 we compare the BER, averaged over $10,000$ independent channel realizations, obtained with full rate Unitary TOD, using the shift-and-multiply bases of [4], setting $n_T = 2$ and $n_R = 4, 5$ and 6 for a two-user system. Every user is transmitting at full-rate, i.e. n_T symbols per channel use so that the aggregate rate is $N \cdot n_T$. We can see that, for

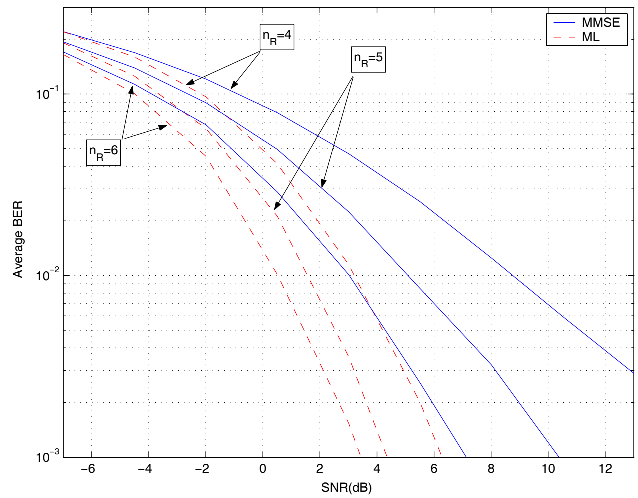

Fig. 1. Average BER obtained with full rate unitary TOD, using shift and multiply bases and ML (dashed line) or MMSE (solid line) decoding.

$n_R = 2n_T$, the MMSE decoder has a considerable loss with respect to ML, but, as soon as n_R increases with respect to $N \cdot n_T$, the loss decreases. Considering the huge difference in complexity between the MMSE and the ML decoders, TOD is then an interesting candidate for multi-user systems, provided that the access point has a sufficient number of receive antennas.

5. REFERENCES

[1] V. Tarokh, H. Jafarkhani, and A. R. Calderbank,"Spacetime block codes from orthogonal designs",*IEEE Trans. Inform. Theory*, vol. 45, pp. 1456-1467, July 1999.

[2] Ma, X., Giannakis, G.B., "Full-diversity full-rate complex-field space-time coding", *IEEE Transactions on Signal Processing*, Vol. 49, Nov. 2003, pp. 2917–2930.

[3] El Gamal, H., Damen, M., O., "Universal space-time coding", *IEEE Transactions on Information Theory*, Vol. 49, May 2003, pp. 1097–1119.

[4] Barbarossa, S., "Trace-orthogonal design of MIMO systems with simple scalar detectors, full diversity and (almost) full rate", *Proc. of SPAWC 2004*, Lisbon, Portugal, July 11–14, 2004.

[5] J. K. Zhang, K. M. Wong, T. N. Davidson, "Information Lossless Full Rate Full Diversity Cyclotomic Linear Dispersion Codes", *ICASSP 2004*, Montreal, Canada, May 17-21, 2004.

[6] Zheng, L., Tse, D.N.C., "Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels,", *IEEE Trans. on Information Theory*, Vol. 49, May 2003, pp. 1073–1096.

[7] J. Liu, J. K. Zhang, K. M. Wong, "Design of Optimal Orthogonal Linear Codes in MIMO Systems for MMSE Receiver", *ICASSP 2004*, Montreal, Canada, May 17-21, 2004.