A MINIMAX OPTIMAL DECODER FOR OFDM OVER UNKNOWN FREQUENCY-SELECTIVE FADING CHANNELS

Ofer Shayevitz and Meir Feder

Tel Aviv University, Dept. of EE-Systems
Tel Aviv 69978, Israel
Email: {ofersha,meir}@eng.tau.ac.il

ABSTRACT

We address the problem of decoding in unknown frequency selective fading channels, using an OFDM signaling scheme and adopting a block fading model. For a given codebook, we seek a decoder independent of the channel fading, whose worst case performance relative to a Maximum Likelihood (ML) decoder that knows the channel, is optimal. Specifically, the decoder is selected from a family of quadratic decoders, and the optimal decoder is referred to as a Quadratic Minimax (QMM) decoder for that family. The intuitively appealing QMM decoding procedure is derived for the case where the fading is unknown, and also for the case where the fading coefficients satisfy some general constraints. The QMM decoder is also shown to outperform the Generalized Likelihood Ratio Test (GLRT), while maintaining a comparable complexity. Simulations verify the superiority of the proposed decoder over the GLRT and over the practically used training sequence approach.

1. INTRODUCTION

Orthogonal frequency-division multiplexing (OFDM) signaling has gained much attention as an effective multicarrier technique for wireless transmissions over frequency selective fading channels [1]. By using the Fast Fourier Transform (FFT) and its inverse (IFFT) and adding a cyclic prefix to each data block, OFDM converts a frequency selective fading channel with additive white Gaussian noise (AWGN) into parallel independent subchannels (bands) with AWGN [2], and therefore greatly simplifies the equalization stage at the receiver, when the bands fading coefficients are known.

In many situations, the receiver is not familiar with the specific channel over which communication takes place, and thus the fading coefficients are unknown. An important example for such a situation is found in mobile wireless communication, where variations of the transmitter location in a dense urban environment leads to constantly changing scattering scenarios, which in turn result in a varying channel law. In such cases, decoding methods that relay on the knowledge of fading values, such as Maximum Likelihood (ML) decoding, become completely useless.

The most common approach to this problem is the use of a training sequence to estimate the channel fading [3], typically followed by a ML decoder tuned to the estimated channel. This approach has some drawbacks. On the one hand, a good estimate of the fading in a time-varying environment requires a long training sequence and frequent retransmissions, resulting in a rate penalty. On the other hand, using a shorter training sequence increases the estimate mismatch, thus increasing error rates and decreasing capacity [4]. Another approach is the Generalized Likelihood Ratio Test (GLRT), which performs a joint ML estimate of the channel and the codewords [5]. The GLRT is a heuristic generalization of the ML decoding rule, and in many cases there exist other decoders with uniformly better performance [6].

In this paper, we present a novel decoder for the unknown OFDM setting, based on a competitive minimax universality criterion [7],[8],[9]. This criterion seeks a decoder that is optimal in the sense of attaining an error probability uniformly as close as possible to that of the ML decoder that knows the channel.

2. SYSTEM MODEL

We consider an OFDM signaling scheme over an unknown frequency selective fading channel. It is assumed that there are L frequency bands, where each band suffers unknown fading and AWGN. The FFT, IFFT and the cyclic prefix associated with the OFDM signaling scheme will be disregarded here, being a constant part of the transmission scheme. Adopting a block fading model, we assume that the unknown fading coefficients are constant throughout a block of K consecutive OFDM symbols, and only change from block to block. We also assume that a given codebook of M codewords is used, where codewords are selected with equal probability. Each codeword is assumed to occupy a single block, and can be therefore represented by a $L \times K$ matrix.

Specifically, the output of the channel when transmitting the i-th codeword $X^{(i)}$ is given by

$$Y = AX^{(i)} + Z$$

$$A = diag\{\boldsymbol{a}\}, \quad X^{(i)}, Y, Z \in \mathbb{R}^{L \times K}, \quad Z_{\ell k} \sim N(0,1)$$

where the elements of the *fading vector* $\boldsymbol{a}$ are the L unknown fading coefficients, $X^{(i)}_{\ell k}$ is the component of the i-th codeword transmitted on the k-th symbol in band ℓ, and the elements of the matrix Z are i.i.d normal random variables with unit variance. All the results derived herein are also valid for complex OFDM channels; see [7] for details.

Denote the ℓ-th row of $X^{(i)}$ by $\boldsymbol{x}^{(i)}_\ell$. We define $P^{(i)}_\ell$ and $\rho^{(i,j)}_\ell$ to be the (transmitted) power of codeword i on band ℓ and the correlation coefficient between codewords i and j on band ℓ respectively:

$$P^{(i)}_\ell = \| \boldsymbol{x}^{(i)}_\ell \|^2 , \quad \rho^{(i,j)}_\ell = \frac{\langle \boldsymbol{x}^{(i)}_\ell , \boldsymbol{x}^{(j)}_\ell \rangle}{\sqrt{P^{(i)}_\ell P^{(j)}_\ell}}$$

The probability of error associated with a decoder Ω for a specific fading vector $\boldsymbol{a}$ is denoted by $P_e(\boldsymbol{a}, \Omega)$. Since a precise analysis of $P_e(\boldsymbol{a}, \Omega)$ for general decoders is usually hard, we resort to high SNR approximations. We define the decoder's *power error exponent* as the asymptotic slope of the error probability as a function of the SNR on a logarithmic scale:

$$E^{\Omega}(\boldsymbol{a}) = \lim_{r \to \infty} -\frac{1}{r} \log P_e(\sqrt{r}\boldsymbol{a}, \Omega)$$

As this quantity is easier to determine, we will later use it instead of the error probability. Further, a decoder Ω can be decomposed (though not uniquely) into *pairwise components* Ω^{ij} that decides only between codewords i and j, where the decision rule for Ω favors the codeword that is selected by all its corresponding pairwise components. If no such codeword exists, then decision can be made by some inconsistency resolving rule, but we disregard this event here as it does not affect our results. Decoders that uniquely assign a metric to different codewords (such as the ML and the GLRT) can be decomposed into pairwise components simply by comparing pairwise metrics. Inversely, a decoder Ω can also be defined by stating its pairwise components Ω^{ij}.

The *pairwise error probability* $P_e^{i \to j}(\boldsymbol{a}, \Omega^{ij})$ is defined as the probability that Ω^{ij} favors codeword j while codeword i was transmitted, over a channel with a fading vector $\boldsymbol{a}$. The *pairwise power error exponent* is defined accordingly as

$$E^{\Omega}_{ij}(\boldsymbol{a}) = \lim_{r \to \infty} -\frac{1}{r} \log P_e^{i \to j}(\sqrt{r}\boldsymbol{a}, \Omega^{ij})$$

Notice that generally $E^{\Omega}_{ij}(\boldsymbol{a}) \neq E^{\Omega}_{ji}(\boldsymbol{a})$. For a constant number of codewords, the decoder's error probability is dominated by the worst pair, and so $E^{\Omega}(\boldsymbol{a})$ equals the minimal pairwise exponent:

$$E^{\Omega}(\boldsymbol{a}) = \min_{i \neq j} E^{\Omega}_{ij}(\boldsymbol{a}) \tag{1}$$

Denote by $d^{\Omega}_{ij}(\boldsymbol{a})$ the square of the minimal distance of codeword i, transmitted over a channel with a fading vector $\boldsymbol{a}$, from the separation surface of Ω^{ij}. Since the noise is AWGN, it is easy to verify [7] that generally

$$E^{\Omega}_{ij}(\boldsymbol{a}) = \frac{1}{2} d^{\Omega}_{ij}(\boldsymbol{a}) \tag{2}$$

In our setting, the ML decoding rule is a minimum Euclidian distance rule. Using (2) and a pairwise metrics decomposition, the power error exponent for the ML decoder is easily verified to be

$$E^*(\boldsymbol{a}) = \frac{1}{8} \min_{i \neq j} \left\{ \sum_{\ell=0}^{L-1} \left(P_\ell^{(i)} + P_\ell^{(j)} - 2\rho_\ell^{(ij)} \sqrt{P_\ell^{(i)} P_\ell^{(j)}} \right) a_\ell^2 \right\} \tag{3}$$

The corresponding probability of error for the ML decoder will be denoted by $P_e^*(\boldsymbol{a})$.

A decoder Ω will be called a *Quadratic Decoder*, if there exists a set of symmetric matrices H_{ij} so that Ω can be decomposed into

$$\Omega^{ij}(\boldsymbol{y}) = \begin{cases} i & \boldsymbol{y}^T H_{ij}\, \boldsymbol{y} > 0 \\ j & \boldsymbol{y}^T H_{ij}\, \boldsymbol{y} < 0 \\ arbitrary & o.w. \end{cases}$$

Notice that there is an implicitly assumed dependence between the matrices H_{ij} and H_{ji}, needed to ensure that the pairwise components Ω^{ij} and Ω^{ji} describe the same decoding rule. In the OFDM setting, the GLRT turns out to be a quadratic decoder, which may imply that such decoders are interesting. Therefore, selecting an optimal decoder from families of quadratic decoders will be in the focus of our interest in the following section.

3. THE QUADRATIC MINIMAX (QMM) DECODER

3.1. The Minimax Criterion

In this section, a competitive minimax criterion for decoders will be suggested and applied in the OFDM setting. This approach follows on the work presented in [8], with two main differences: First, a fixed codebook is assumed and the optimal decoder is considered in the limit of high SNR rather than in the limit of increasing block size, by using the power error exponents in lieu of the error probability. Second, structure is introduced by selecting the decoder from a family of quadratic decoders, so that determining the optimal decoder becomes more tractable.

Specifically, given a family $\mathcal{F}$ of quadratic decoders, we seek an optimal decoder $\Omega \in \mathcal{F}$ in the competitive minimax sense:

$$\inf_{\Omega \in \mathcal{F}} \sup_{\boldsymbol{a}} \frac{P_e(\boldsymbol{a}, \Omega)}{(P_e^*(\boldsymbol{a}))^{\xi}} \tag{4}$$

for some $0 \leq \xi \leq 1$. For $\xi = 1$, this ratio represents the relative loss in performance incurred by employing a decoder ignorant of the channel in use, and therefore the proposed criterion seeks a decoder whose worst case relative loss is minimal. Minimax optimality w.r.t the ML may sometimes be too ambitious, and so by using $\xi < 1$, optimality is sought w.r.t a fraction ξ of the ML error probability, on an exponential scale. Now, in the limit of high SNR, the minimax criterion above may be approximated by replacing the error probabilities with their corresponding power error exponents:

$$\inf_{\Omega \in \mathcal{F}} \sup_{\boldsymbol{a}} \frac{P_e(\boldsymbol{a}, \Omega)}{(P_e^*(\boldsymbol{a}))^{\xi}} \approx \inf_{\Omega \in \mathcal{F}} \sup_{\|\boldsymbol{a}\|=1} \sup_{r \gg 0} \left(\xi E^*(\boldsymbol{a}) - E^{\Omega}(\boldsymbol{a}) \right) r$$

A decoder that is optimal by this criteria, must achieve a power error exponent at least as good as a fraction ξ of the ML exponent, uniformly over all fading values. For a high value of ξ, such a decoder may not exist at all, since the supremum will always diverge. For a low value of ξ, there may be many optional decoders. Therefore, we would be interested in the maximal value $\xi = \xi^*$ for which such a decoder exists. Defining ξ^{Ω} as the guaranteed fraction of the ML exponent attained by a decoder Ω, we have

$$\xi^{\Omega} \triangleq \inf_{\|\boldsymbol{a}\|=1} \frac{E^{\Omega}(\boldsymbol{a})}{E^*(\boldsymbol{a})}, \quad \xi^* = \sup_{\Omega \in \mathcal{F}} \xi^{\Omega} = \sup_{\Omega \in \mathcal{F}} \inf_{\|\boldsymbol{a}\|=1} \frac{E^{\Omega}(\boldsymbol{a})}{E^*(\boldsymbol{a})} \tag{5}$$

The decoder attaining ξ^* will be termed the *Quadratic Minimax (QMM)* decoder for the family $\mathcal{F}$.

3.2. The Decoding Procedure

We now specify the family $\mathcal{F}$ by describing the pairwise decoding procedure for decoders $\Omega \in \mathcal{F}$. It can be shown that when making a pairwise decision, it is sufficient to consider, in each band, only the projection of the observation onto the subspace spanned by the two codewords in that band [7]. This is also intuitively reasonable since anything orthogonal is essentially noise.

With that in mind, we choose to consider decoders with the following pairwise decoding procedure: Let $\boldsymbol{y}_\ell$ be the observation in band ℓ (the ℓ-th row of the matrix Y) and let $\boldsymbol{r}_\ell$ be the projection of $\boldsymbol{y}_\ell$ onto $span\{\boldsymbol{x}_\ell^{(i)}, \boldsymbol{x}_\ell^{(j)}\}$. Now, express $\boldsymbol{r}_\ell$ as a linear combination of $\boldsymbol{x}_\ell^{(i)}$ and $\boldsymbol{x}_\ell^{(j)}$

$$\boldsymbol{r}_\ell = \alpha_\ell\, \boldsymbol{x}_\ell^{(i)} + \beta_\ell\, \boldsymbol{x}_\ell^{(j)}$$

The pairwise component Ω^{ij} favors codeword i over j if

$$\sum_\ell w_\ell^{ij}\alpha_\ell^2 \;\geq\; \sum_\ell w_\ell^{ji}\beta_\ell^2$$

for some set of coefficients w_ℓ^{ij}. The family $\mathcal{F}$ is therefore the family of all decoders with pairwise components following the above decoding rule. However, it turns out [7] that for our purposes it is sufficient to consider a much smaller family $\mathcal{F}^* \subset \mathcal{F}$ without any loss of generality. Decoders $\Omega \in \mathcal{F}^*$ have pairwise components each dependent only upon a selection of a **single** weight parameters $\lambda_{ij} > 0$, and favoring codeword i over j if

$$\sum_\ell \left[P_\ell^{(i)}(1-|\rho_\ell^{(ij)}|)\right]\alpha_\ell^2 \;\geq\; \lambda_{ij}\sum_\ell \left[P_\ell^{(j)}(1-|\rho_\ell^{(ij)}|)\right]\beta_\ell^2 \tag{6}$$

A decoder $\Omega \in \mathcal{F}^*$ is therefore described by $\frac{M(M-1)}{2}$ weights λ_{ij} of its pairwise components (notice that $\lambda_{ij} = \frac{1}{\lambda_{ji}}$ for consistency). Next, we seek the weights of the QMM decoder for the family $\mathcal{F}^*$.

3.3. Weights Selection

The power error exponent of a decoder $\Omega \in \mathcal{F}^*$ is the minimum of all its pairwise power error exponents. These in turn are related to the pairwise minimal distance according to (2), and are given by a the solutions of a set of quadratic optimization problems with quadratic constraints. Unfortunately, these problems lack an analytical solution in general, and so the power error exponent does not have an analytical expression. However, a lower bound can be found by transforming the pairwise optimization problem, using a linear transformation, into a problem that can be solved analytically [7]. This bound can be optimized by selecting the transformation matrix to have a minimal condition number, and the resulting bound $\hat{E}^\Omega(\boldsymbol{a})$ for the power error exponent is

$$E^\Omega(\boldsymbol{a}) \geq \frac{1}{2}\min_{i\neq j}\left\{\frac{1}{1+\lambda_{ij}}\sum_{\ell=0}^{L-1} P_\ell^{(i)}(1-|\rho_\ell^{(ij)}|)a_\ell^2\right\} \triangleq \hat{E}^\Omega(\boldsymbol{a}) \tag{7}$$

A corresponding lower bound $\hat{\xi}^\Omega$ for ξ^Ω is now derived by replacing $E(\boldsymbol{a})$ with $\hat{E}(\boldsymbol{a})$ in (5):

$$\xi^\Omega = \inf_{\|\boldsymbol{a}\|=1}\frac{E^\Omega(\boldsymbol{a})}{E^*(\boldsymbol{a})} \geq \inf_{\|\boldsymbol{a}\|=1}\frac{\hat{E}^\Omega(\boldsymbol{a})}{E^*(\boldsymbol{a})} \triangleq \hat{\xi}^\Omega \tag{8}$$

By definition, the QMM decoder for the family $\mathcal{F}^*$ is the one maximizing ξ^Ω. Since we do not have an explicit expression for ξ^Ω, we shall find a decoder $\Omega \in \mathcal{F}^*$ maximizing the lower bound $\hat{\xi}^\Omega$ instead. Although suboptimal, this decoder will still be referred to as the QMM decoder, and its guaranteed fraction of the ML exponent will still be denoted ξ^*, so to avoid too many notations.

In order to derive the QMM decoder, we first establish an expression dependent on the weights λ_{ij} for $\hat{\xi}^\Omega$, and then seek weights maximizing it. For a given decoder, the channels with a fading vector $\boldsymbol{a}$ for which the infimum in the right hand side of (8) is attained will be referred to as *critical channels*, and if $\boldsymbol{a}_0$ happens to be a critical channel, then the lower bound $\hat{\xi}^\Omega$ is just the ratio $\hat{E}^\Omega(\boldsymbol{a}_0)/E^*(\boldsymbol{a}_0)$. Therefore, for any specific decoder, the lower bound we seek can be found by determining any one of the critical channels. Generally, the critical channels are difficult to determine in closed form, and may vary from one quadratic decoder to another, as they depend on the selection of the weights λ_{ij}. However, it turns out that there exists a finite set of channels independent of the weights, that always includes at least one critical channel. Considering only vectors $\boldsymbol{a}$ with unity norm, introduce the change of variables $b_\ell = a_\ell^2$ for $\ell = 0,\ldots,L-2$. With a slight abuse of notations, we now use $\hat{E}^\Omega(\boldsymbol{b})$ and $E^*(\boldsymbol{b})$ instead of $\hat{E}^\Omega(\boldsymbol{a})$ and $E^*(\boldsymbol{a})$ respectively. The domain of $\hat{E}^\Omega(\boldsymbol{b})$ and $E^*(\boldsymbol{b})$ is the $L-1$ dimensional simplex defined by

$$S = \left\{\boldsymbol{b} \;:\; b_\ell \geq 0\,,\, \sum_{\ell=0}^{L-2} b_\ell \leq 1\right\} \tag{9}$$

It is easily seen that both the ML power error exponent $E^*(\boldsymbol{b})$ and the lower bound $\hat{E}^\Omega(\boldsymbol{b})$ are polyhedral functions over the simplex S, being the minima of affine functions (hyperplanes). Polyhedral functions are characterized by *extreme points*, also called *corners*. We state our next result without proof; see [7] for details.

Theorem. *For any selection of weights λ_{ij}, the finite set of the extreme points of $E^*(\boldsymbol{b})$ over the simplex S includes at least one critical channel.*

We thus see that $\hat{\xi}^\Omega$ for any decoder $\Omega \in \mathcal{F}^*$, can be found by searching over a finite a-priori determined set of channels, independent of the decoder itself. Let $\boldsymbol{b}_0,\ldots,\boldsymbol{b}_{N-1}$ be the set of extreme points of $E^*(\boldsymbol{b})$ over S. Then the bound $\hat{\xi}^\Omega$ is given by

$$\hat{\xi}^\Omega = \min_n \frac{\hat{E}^\Omega(\boldsymbol{b}_n)}{E^*(\boldsymbol{b}_n)} \tag{10}$$

and we now seek weights λ_{ij} that maximize it. Denote the channels corresponding to the extreme points by $\boldsymbol{a}_0,\ldots,\boldsymbol{a}_{N-1}$. Using (7) and (10) we have

$$\hat{\xi}^\Omega = \frac{1}{2}\min_{i\neq j}\left\{\frac{1}{1+\lambda_{ij}}\min_n\left(\frac{\sum P_\ell^{(i)}(1-|\rho_\ell^{(ij)}|)a_\ell^2}{E^*(\boldsymbol{a}_n)}\right)\right\}$$

Now, using the weights dependency $\lambda_{ij} = \frac{1}{\lambda_{ji}}$ and defining

$$f_{ij}(\boldsymbol{a}) \triangleq \frac{1}{2}\sum P_\ell^{(i)}(1-|\rho_\ell^{(ij)}|)a_\ell^2\,,\quad s^{(ij)} \triangleq \min_n \frac{f_{ij}(\boldsymbol{a_n})}{E^*(\boldsymbol{a}_n)}$$

the optimal selection of weights maximizing $\hat{\xi}^\Omega$ is

$$\lambda_{ij}^* = \frac{s^{(ij)}}{s^{(ji)}} \tag{11}$$

where $E^*(\boldsymbol{a})$ is given in (3). Notice that if the number of ML extreme points is relatively low, it may be preferable to calculate the weights online, so to avoid storing them in memory.

3.4. Utilizing Relations Between Fading Values

In some cases, a-priori information regarding the interdependency of the fading values may be available, and can be incorporated into our QMM decoding scheme. For instance, if the channel suffers from a multi-path distortion with a time spread of T_m, and the OFDM bands are δW wide, then if $\delta W \cdot T_m \ll 1$, adjacent frequency bands suffer fading that is highly dependent. Specifically, assume it is known that the fading vector $\boldsymbol{a}$, rather than taking arbitrary values in $\mathbb{R}^L$, takes only values in a *constraint set* $A \subseteq \mathbb{R}^L$. In this case, the minimax criterion (4) becomes

$$\inf_{\Omega\in\mathcal{F}}\sup_{\boldsymbol{a}\in A}\frac{P_e(\boldsymbol{a},\Omega)}{(P_e^*(\boldsymbol{a}))^\xi} \tag{12}$$

Under some conditions on the set A, a suboptimal QMM decoder for the modified optimality criterion above can be derived. In order for our SNR-asymptotic approach to apply, the set A must be scale-invariant, so that any channel in it can be found with all possible gains. If additionally the set

$$S \triangleq \left\{ \boldsymbol{b} \mid \boldsymbol{b} = (a_0^2, \ldots, a_{L-2}^2)\,,\, \boldsymbol{a} \in A\,, \|\boldsymbol{a}\| = 1 \right\} \qquad (13)$$

is a polyhedra, then the QMM optimal weights for the constrained case can be found in a manner similar to that of finding the optimal weights in the unconstrained case, where the only difference is that the set S defined in (13) replaces the simplex of (9) as the set over which ML extreme points are sought.

3.5. Comparison with the GLRT

The power error exponent for the GLRT can be upper bounded by finding, for each pairwise component, the distance to a point on its separation surface, which attains the minimal distance per-band, and the resulting bound is

$$E^{GLRT}(\boldsymbol{a}) \leq \frac{1}{4} \min_{i \neq j} \left\{ \sum_{\ell=0}^{L-1} P_\ell^{(i)} (1 - |\rho_\ell^{(ij)}|) a_\ell^2 \right\} \qquad (14)$$

Now, notice that this bound coincides with the lower bound $\hat{E}^{\Omega}(\boldsymbol{a})$ of (7) for a uniform selection of weights $\lambda_{ij} = 1$. Therefore, since the QMM decoder selects weights λ_{ij} maximizing the lower bound, it is guaranteed to have a higher value of ξ than that attained by the GLRT, thus outperforming it in the minimax sense.

4. SIMULATION RESULTS

The coding scheme used for simulation is based on the *Complex Field Coding (CFC)* notion, which is basically linear precoding with redundancy, suggested for OFDM over Rayleigh fading channels [10]. Simulations were performed over randomly selected channels, comparing the QMM to the GLRT (figures 1,2) and to the training approach (figure 2) in terms of the attained fraction of the ML exponent ξ. As can be seen, the QMM outperform both in terms of the worst case value of ξ attained over those channels.

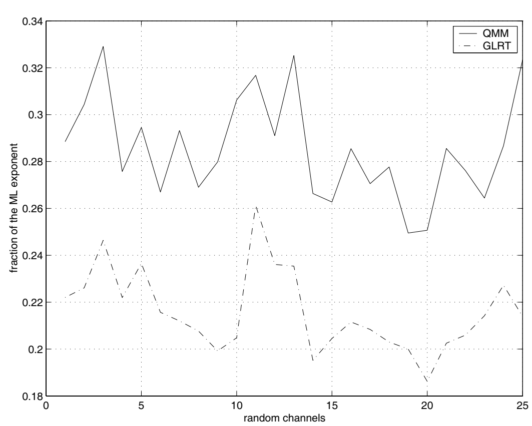

Fig. 1. QMM vs. GLRT, 120 codewords CFC, 3×4 channel

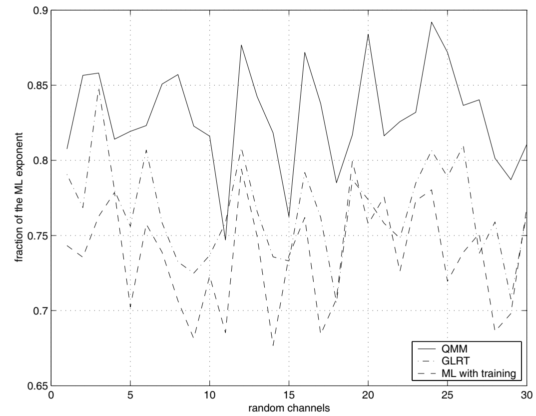

Fig. 2. QMM vs. GLRT and ML-training, 28 codewords CFC, 3×5 channel, 2 training symbols per block

5. REFERENCES

[1] "IEEE standard 802.11a-1999, adopted by the ISO/IEC and redesignated as ISO/IEC 8802-11:1999/Amd 1:2000(E)," .

[2] S. B. Weinstein and P. M. Ebert, "Data transmission by frequency-division multiplexing using the discrete Fourier transform," *IEEE Trans. Commun. Technol.*, vol. 19, pp. 628–634, Oct. 1971.

[3] F. Tufvesson and T. Maseng, "Pilot assisted channel estimation for OFDM in mobile cellular systems," Phoenix, U.S.A, May 1997, pp. 1639–1643.

[4] M. Medard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Trans. Inform. Theory*, vol. 46, pp. 933–946, May 2000.

[5] J. Choi, "A joint channel estimation and detection for frequency-domain equalization using an approximate EM algorithm," *Signal Processing*, vol. 84, pp. 865–880, May 2004.

[6] E. Erez and M. Feder, "Improving the generalized likelihood ratio test for unknown linear Gaussian channels," *IEEE Trans. Inform. Theory*, vol. 49, pp. 919–936, Apr. 2003.

[7] O. Shayevitz and M. Feder, "Universal decoding for frequency-selective fading channels," *Submitted*, Sept. 2004.

[8] M. Feder and N. Merhav, "Universal composite hypothesis testing: A competitive minimax approach," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1504–1517, June 2002.

[9] A. Lapidoth and P. Narayan, "Reliable communications under channel uncertainty," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2148–2177, Oct. 1998.

[10] Z. Wang and G. B. Giannakis, "Complex-field coding for OFDM over fading wireless channels," *IEEE Trans. Inform. Theory*, pp. 707–720, Mar. 2003.