OFDM Channel Estimation with Optimal Threshold-Based Selection of CIR Samples

3 authors:

Stefano Rosati
École Polytechnique Fédérale de Lausanne
15 PUBLICATIONS   **37** CITATIONS

SEE PROFILE

Giovanni Emanuele Corazza
University of Bologna
292 PUBLICATIONS   **2,282** CITATIONS

SEE PROFILE

Alessandro Vanelli-Coralli
University of Bologna
166 PUBLICATIONS   **795** CITATIONS

SEE PROFILE

OFDM Channel Estimation with Optimal Threshold-Based Selection of CIR Samples

Stefano Rosati, Giovanni E. Corazza, and Alessandro Vanelli-Coralli

DEIS/ARCES - University of Bologna, Bologna, Italy,
e-mail: { srosati, gecorazza, avanelli }@arces.unibo.it

Abstract—In this paper, we address the problem of OFDM data-aided channel estimation based on the selection of the most significant samples (MSS) of the Channel Impulse Response (CIR), i.e. those samples which contain most of the useful energy. We provide a novel and complete analytical characterization for MSS selection based on threshold crossing, which yields a closed form for the estimate mean-square error (MSE), that we use to derive analytically the optimum threshold in the minimum MSE sense. The optimum threshold value is matched to the specific channel power profile, but this information is hardly available to the receiver. For these reason, we also propose a sub-optimal method for threshold setting that does not require any knowledge of channel statistics. We show that the performance of this sub-optimal method is very close to the optimum case, as well as to Wiener channel estimation in the case of sparse multipath channels. Our proposed method outperforms previous approaches based on heuristically set thresholds.

I. INTRODUCTION

One of the main advantages of Orthogonal Frequency Division Multiplexing (OFDM) is its possibility to cope with strongly dispersive channels using low-complexity equalizers, with a single tap per sub-carrier. This is due to the fact that propagation channels which are frequency selective over the entire OFDM bandwidth appear as non-selective on each narrowband sub-carrier. In particular, this is essential for the case of OFDM single-frequency networks (SFNs), whereby the same signal is transmitted from widely separated sites to achieve seamless radio coverage. In fact, this is the most critical case in terms of time dispersion, since interfering signal replica may come with very large differential delays, giving rise to very sparse channel profiles.

Clearly, the equalizer will perform adequately if and only if accurate estimation of the channel transfer function (CTF), or equivalently of the channel impulse response (CIR), is performed at the receiver. In other words, channel estimation becomes *the* critical function which largely determines the overall receiver performance. Data-aided (DA) channel estimation is the subject of the present paper, whereby known pilots are multiplexed into OFDM symbols, drawing a regular pattern of known sub-carriers with inter-pilot spacing dependent on the channel coherence bandwidth.

The problem of DA estimation of frequently selective time-varying wireless channels for OFDM communications has been widely investigated in recent years, due to the strong rise of OFDM onto the wireless scene. Generally, DA channel estimation is performed in two steps: first, the CTF is punctually estimated on pilot sub-carriers; then, punctual estimates are interpolated or filtered throughout the OFDM sub-carriers comb. DA channel estimation methods differ on the way they accomplish the second step. Two-dimensional (2D) time-frequency Wiener filtering [1] is optimal in the Minimum Mean Squared Error (MMSE) sense. On the other hand, 2D Wiener filtering requires perfect knowledge of the channel statistics and large complexity. In order to reduce the computational complexity, it is possible to separate the time and frequency domains, applying in each case a one-dimensional Wiener filter. In [2] P. Hoeher at al. showed that this segregated approach reduces significantly the computational complexity with only a slight/moderate performance decrease as compared to 2D Wiener filtering;Alternately, Channel Estimation can be accomplished treating raw estimates in the time-domain using a Discrete Fourier Transform (DFT) based scheme [3]. In fact, effective interpolation can be performed by applying an inverse Discrete Fourier Transform (IDFT) to produce a CIR estimate, zero-padding the result and then returning to the CTF estimate through a DFT. Considering this scheme, in [4], the MMSE channel estimator working in the the domain, which takes into account channel correlation in both time and frequency domain, has been proposed. Also in this method the knowledge of channel statistics is still required.

In practice, it is very useful to adopt methods which require the minimum possible amount of information about the channel statistics, but that are applicable to any kind of power delay profile, and in particular to sparse channels, as those encountered in SFNs. The simplest agnostic approach is to interpolate punctual Data-Aided Least Squares (DA-LS) using DFT-based scheme without any elaboration in the time-domain. Unfortunately, simplicity here goes along with mediocre performance [5], and much research work is being devoted to improving this approach without increasing complexity in any significant way. The main idea to achieve this goal is the following: after the IDFT, not all the CIR samples are significant, as many may correspond to delays where no propagation channel path is actually present. Therefore, if one can devise a technique to retain only the significant samples, performance can be improved without complexity increase. The first idea, discussed in [5], is that by estimating the maximum channel delay, we can discard all samples that exceed this maximum, thus reducing noise. A further step,

introduced by Minn et al [6], is to discard also any sample that is within the maximum delay, but is likely not to correspond to a channel path. In particular, they proposed to select only the J strongest samples, identified here as the Most Significant Samples (MSS) of the CIR estimate, J being an important design parameter. It can be shown that in the ideal case where J equals the actual number of channel paths, N_m, very good performance can be achieved; but when J differs from N_m, performance degrades rapidly. Instead of pre-determining a-priori the total MSS number, Kang et al. [8] proposed to select them by comparing to a threshold ξ. In this way, a dynamic number of MSS is selected per OFDM symbol. It is clear that the threshold value is critical to the algorithm performance. While in [8] the threshold was set according to heuristics, in [9] a genie-aided approach was followed, based again on the knowledge of channel statistics, then analysed by means of simulation.

In this paper, we produce a complete analytical characterization of OFDM channel estimation based on LS-DFT interpolation with threshold-based MSS selection. Our analytical framework allows to determinate in closed form the MSE; in turn, this can be used to determine analytically the optimum threshold value by minimizing the MSE itself. This provides a benchmark for the performance of all these algorithms. As expected, the optimal threshold depends on the actual channel power delay profile. As an example, we report the optimal threshold in the case of a sparse uniform power profile. To avoid falling back into the trap of requiring a-priori knowledge of channel statistics, we also introduce a pragmatic approach for threshold selection, based on a specification for the overall false alarm probability. Even if this method provides a sub-optimal of threshold, we show that it yields very close performance to the optimal case, and that it is also very robust to various power delay profiles. Furthermore, in the case of sparse channels, performance is shown to be close to Wiener channel estimation.

II. System Model

We consider an OFDM signal with N total subcarriers; the ℓ-th OFDM symbol, $\bar{s}_\ell = (s_{0,\ell}, \ldots, s_{N-1,\ell})$, is obtained as the N-point Inverse Discrete Fourier Transform (IDFT) of the vector of complex symbols $\bar{x}_\ell = (x_{0,\ell}, \ldots, x_{N-1,\ell})$, according to

$$s_{i,\ell} = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} x_{k,\ell} e^{j\,2\pi ki/N} \qquad i = 0, \ldots, N-1 \quad (1)$$

As anticipated in the Introduction, the complex symbols $x_{k,\ell}$ carry either data information, $a_{k,\ell}$, or pilot reference symbols, $p_{k,\ell}$, used for synchronization and channel estimation. Let N_p be the number of pilot sub-carriers, which we assume to be uniformly scattered in each OFDM symbol (constant pilot spacing), but may shift periodically from symbol to symbol to improve coverage of the entire frequency comb. Let $P(\ell)$ be the set of pilot subcarrier indices in the ℓ-th ODFM symbol,

identified as *pilot pattern*, of size N_p. Therefore, we can write

$$x_{k,\ell} = \begin{cases} p_{k,\ell} & \text{if } k \in P(\ell) \\ a_{k,\ell} & \text{if } k \notin P(\ell) \end{cases} \quad (2)$$

To improve synchronization and estimation performance, pilots can be transmitted with energy amplified by a factor β^2 with respect to data symbols (i.e. $E[p_{k,\ell}^2] = \beta^2 E_s$, where $E[a_{k,\ell}^2] = E_s$).

In order to avoid intersymbol interference and maintain subcarrier orthogonality in multipath, a *cyclic prefix* of length N_g samples is inserted at the beginning of each OFDM symbol. This is followed by digital to analog conversion at sample rate $R = 1/T$, so that the time continuous signal can be written as

$$s(t) = \frac{1}{\sqrt{T_u}} \sum_{\ell=-\infty}^{\infty} \text{rect}\left(\frac{t}{T_L} - \frac{1}{2} - \ell\right) \sum_{k=0}^{N-1} x_{k,\ell} e^{j\,2\pi \frac{1}{T_u} k(t-T_g)} \quad (3)$$

where $T_u = NT$ represents the OFDM useful symbol duration, $T_g = N_g T$ represents the duration of the guard interval associated to the cyclic prefix. Therefore $T_L = T_u + T_g$ is the total OFDM symbol duration, and $f_u = 1/T_u$ is the subcarrier spacing. Notably, the normalization factor $\frac{1}{\sqrt{T_u}} = \frac{1}{\sqrt{N}} \cdot \frac{1}{\sqrt{T}}$ accounts for both the IDFT and the D/A normalization factors.

The OFDM signal is transmitted over a time-varying frequency selective fading channel, under the assumption that the channel coherence time exceeds T_L. The baseband equivalent channel impulse response is modelled as a tapped delay line:

$$h(t) = \sum_{j=0}^{N_m-1} h_j(t)\delta\left(t - \tau_j\right) \quad (4)$$

where $h_j(t)$ and τ_j are respectively the gain and delay of the j-th path, and N_m represents the number of multiple propagation paths. In Rayleigh fading, at any time instant $h_j(t)$ can be modeled as a complex Gaussian random variable with zero mean and variance $\gamma_j^2/2$ per branch. The total channel energy is normalized to one, i.e. $\sum_j E[h_j^2] = \sum_j \gamma_j^2 = 1$, and the maximum delay is assumed to be smaller than the guard interval duration, i. e. $\tau_{max} = \max_j \tau_j < T_g$. The received signal can be written as

$$r(t) = h(t) * s(t) + n(t) \quad (5)$$

where $n(t)$ represents a Complex Additive White Gaussian Noise (AWGN) random process, with two-sided power spectral density equal to N_0. Filtering and sampling the received signal every T seconds yields

$$r(uT) = \sum_j h_j(uT)\, s(uT - \tau_j) + n(uT) \quad (6)$$

Removing the guard interval and re-arranging the vector at the input of the FFT, the samples belonging to the ℓ-th OFDM symbols can be collected into a vector $\bar{r}_\ell$ with components:

$$r_{i,\ell} = r(((\ell-1)(N + N_g) + N_g + i)T)$$
$$i = |u|_{N+N_g} - N_g \qquad \ell = \lceil u/(N + N_g) \rceil \quad (7)$$

Having assumed that $h_j(t)$ remains constant over a OFDM symbol duration, at the output of the FFT, the received vector in the frequency domain is:

$$
\begin{aligned}
y_{k,\ell} &= \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} r_{i,\ell}\, \mathrm{e}^{-\jmath\, 2\pi ki/N} \\
&= x_{k,\ell} H_{k,\ell} + n_{k,\ell} \quad k = 0, \dots, N-1 \quad (8)
\end{aligned}
$$

where $n_{k,\ell}$ is the complex AWGN sample in frequency domain with zero-mean and variance N_0, and $H_{k,\ell}$ is the Channel Transfer Function (CTF) sampled at the subcarrier k, in the ℓ-th symbol, which can be expressed equivalently as:

$$
H_{k,\ell} = \sum_{j=0}^{N_m-1} h_{j,\ell}\, \mathrm{e}^{-\jmath\, 2\pi k\tau_j/NT} \quad k = 0, \dots, N-1 \quad (9)
$$

III. DA-LS ESTIMATION AND DFT INTERPOLATION

As explained in the Introduction, we focus on channel estimation in the frequency domain starting from LS punctual estimates, followed by interpolation via IDFT, padding, and DFT. Time interpolation through several OFDM symbols would follow but is out of our scope here.

1) Punctual Least Squares Estimation: Considering a generic pilot tone positioned at the k-th subcarrier, and in the ℓ-th OFDM symbol, the punctual LS estimate is given by:

$$
\hat{H}_{k,\ell}^{R} = \frac{y_{k,\ell}}{p_{k,\ell}} = H_{k,\ell} + \frac{n_{k,\ell}}{p_{k,\ell}} \qquad k \in P(\ell) \quad (10)
$$

DA channel estimation methods differ on the way they interpolate punctual observations, $\hat{H}_{k,\ell}^{R}$, over data subcarriers. We consider here the DFT interpolation approach.

2) DFT interpolation: Interpolation can be performed by applying an IDFT to the vector of punctual estimates to produce a CIR estimate, $\hat{h}_{i,\ell}$, as follows

$$
\hat{h}_{i,\ell} = \frac{1}{N_p} \sum_{k=0}^{N_p-1} \hat{H}_{k,\ell}^{R} \mathrm{e}^{-\jmath\, 2\pi ki/N_p} \quad i = 0, \dots, N_p-1 \quad (11)
$$

then, assuming that $N_p > \lceil \tau_{max}/T \rceil$, CTF estimation over the entire frequency comb can obtained by zero padding the CIR over the entire OFDM symbol duration and then applying a DFT:

$$
\hat{H}_{\ell}^{LS} = DFT\{\hat{h}_{\ell}^{ZP}\}; \ \hat{h}_{\ell}^{ZP} = \{\underbrace{\hat{h}_{0,\ell}, \dots, \hat{h}_{N_p-1,\ell}}_{N_p}, \underbrace{0, \dots, 0}_{N-N_p}\}
$$
$$(12)$$

3) Statistical Characterization and MSE: Our aim is to model statistically the CIR estimates, $\hat{h}_{i,\ell}$, and compute the DA-LS MSE value. Substituting (10) in (11), we obtain

$$
\hat{h}_{i,\ell} = \frac{1}{N_p} \left(\sum_{k=0}^{N_p-1} H_{k,\ell}\mathrm{e}^{-\jmath\, 2\pi ki/N_p} + \sum_{k=0}^{N_p-1} \frac{n_{k,\ell}}{p_{k,\ell}}\mathrm{e}^{-\jmath\, 2\pi ki/N_p} \right)
$$
$$(13)$$

Thus, we can write the CIR estimate as a sum of the actual value, $h_{i,\ell}$, and a noise component, $d_{i,\ell}$:

$$
\hat{h}_{i,\ell} = h_{i,\ell} + d_{i,\ell} \quad (14)
$$

where:

$$
d_{i,\ell} = \frac{1}{N_p} \sum_{k=0}^{N_p-1} \frac{n_{k,\ell}}{p_{k,\ell}}\mathrm{e}^{-\jmath\, 2\pi ki/N_p} \quad (15)
$$

As a consequence of the central limit theorem, the noise components, $d_{i,\ell}$, are complex Gaussian random variables distributed as:

$$
d_{i,\ell} \sim \mathcal{N}_c \left(0, \frac{1}{\rho N_p} \right) \quad (16)
$$

where $\rho = \beta^2 E_s/N_0$ represents the pilot energy to noise ratio. Considering (14) and (16), the CIR estimate samples, $\hat{h}_{i,\ell}$ are also complex Gaussian random variables distributed as:

$$
\hat{h}_{i,\ell} \sim \begin{cases} \mathcal{N}_c \left(0, \frac{1}{\rho N_p} \right) & \text{if } i \notin \mathcal{C} \\ \mathcal{N}_c \left(0, \frac{1}{\rho N_p} + \gamma_i^2 \right) & \text{if } i \in \mathcal{C} \end{cases} \quad (17)
$$

where $\mathcal{C}$ is the set of indices corresponding to those time samples where propagation paths are actually present, i.e. $i \in \mathcal{C}$ iff $\gamma_i^2 \neq 0$. Considering (14) and (17), and exploiting the linearity of the IDFT, following lines similar to [5] we can show that the MSE of DA-LS channel estimation is given by:

$$
MSE^{LS} = \frac{1}{N} \sum_{k=0}^{N-1} E\left[|H_{k,\ell} - \hat{H}_{k,\ell}^{LS}|^2 \right] =
$$
$$(18)$$
$$
= \sum_{i=0}^{N_p-1} E\left[|h_{i,\ell} - \hat{h}_{i,\ell}|^2 \right] = \sum_{i=0}^{N_p-1} E\left[|d_i|^2 \right] = \frac{N_0}{\beta^2 E_s} = \frac{1}{\rho}
$$

It is useful to compare this result to Wiener filtering (WF). Under the assumption of uniformly scattered pilots, and considering a uniform power delay profile, the MSE for a Wiener channel estimator is given by [7], [5]:

$$
MSE^{WF} = \frac{N_m}{N_m + \rho N_p} \quad (19)
$$

It is clear that, the larger N_p, the larger the gap between LS and WF. We must try to close this gap without significant complexity increases.

IV. MSS THRESHOLD BASED SELECTION

As anticipated in the introduction, not all the CIR estimate samples are significant; in fact, many of them may correspond to delays where no propagation channel path is actually present, and consequently they contain only noise. By retaining only the MSS samples after the IDFT, it is possible reduce drastically the noise presence, especially when the CIR is very sparse, as is the case for SFNs. Obviously, the critical aspect of this method is the MSS selection strategy.

The Threshold Based Selection (TBS) strategy to identify the MSS, which has been treated in [6], [8], [9], is based on the concept that only those samples which overcome the threshold ξ in absolute value are retained:

$$
\hat{h}_i^{Th} = \begin{cases} 0 & \text{if } |\hat{h}_i| \leq \xi \\ \hat{h}_i & \text{if } |\hat{h}_i| > \xi \end{cases} \quad (20)
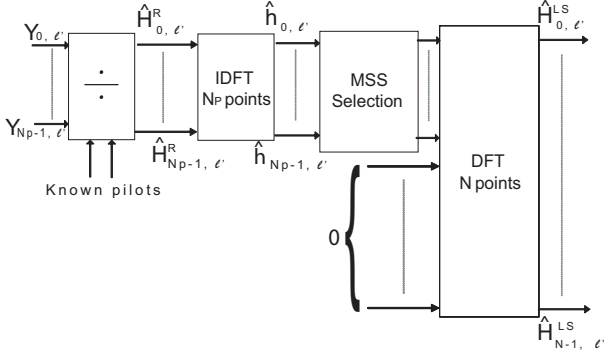$$

Fig. 1. Block diagram of MSS-based Channel Estimator.

Where for sake of notation, we drop the symbol index without any loss of generality. In [8, eq. (21)] the threshold, ξ, has been set equal to twice the estimated noise level:

$$\xi = \frac{2}{\rho N_p} \tag{21}$$

In [9, eq. (6)], the authors proposed a criterion for setting the threshold value that requires the knowledge of CIR power profile:

$$\xi_i = \frac{1}{\rho N_p} \ln \left(2 + \rho N_p \gamma_i^2 \right) \tag{22}$$

Here, we propose to derive the optimal threshold by deriving a closed form expression for the TBS MSE performance, and then minimizing it with respect to ξ.

A. Analysis of the TBS MSE

Considering (14), the estimation errors $\varepsilon_i = |h_i - \hat{h}_i^{Th}|$ are given by:

$$\varepsilon_i = \begin{cases} 0 & \text{if } |\hat{h}_i| \leq \xi \quad \text{and} \quad i \notin \mathcal{C} \\ |d_i| & \text{if } |\hat{h}_i| > \xi \quad \text{and} \quad i \notin \mathcal{C} \\ |h_i| & \text{if } |\hat{h}_i| \leq \xi \quad \text{and} \quad i \in \mathcal{C} \\ |d_i| & \text{if } |\hat{h}_i| > \xi \quad \text{and} \quad i \in \mathcal{C} \end{cases} \tag{23}$$

Note that, for each decision over any generic sample four events are possible:

Correct Rejection, (CR): A non-significant sample is correctly rejected, because the absolute value of its noise component $|d_i|$ is smaller than the threshold. This case does not contribute to MSE.

False Alarm, (FA): A sample that does not contain channel energy is reckoned as belonging to MSS, since its noise component overcame the threshold. In this case, it contributes to MSE with noise components that is distributed as:

$$\varepsilon_i \mid FA \sim \mathcal{K}' \mathcal{R} \left(\frac{1}{2\rho N_p} \right) \quad \varepsilon_i > \xi \tag{24}$$

Where $\mathcal{R}(\cdot)$ represents Rayleigh distribution, and $\mathcal{K}'$ is a normalization factor given by:

$$\mathcal{K}' = e^{-\xi^2 \rho N_p} \tag{25}$$

Missed Detection, (MD): A sample that actually contains channel energy is rejected, since it does not overcome the threshold. This case contributes to MSE with the neglected channel energy that is distributed as:

$$\varepsilon_i \mid MD \sim \mathcal{K}'' \mathcal{R} \left(\frac{\gamma_i^2}{2} \right) \quad \varepsilon_i \leq \xi \tag{26}$$

Where $\mathcal{K}''$ is given by:

$$\mathcal{K}'' = \frac{1}{1 - e^{-\xi^2/\gamma_i^2}} \tag{27}$$

Correct Detection, (CD): A sample that contains channel energy is correctly considered as MSS. This case contributes to MSE only with the noise component of CIR estimates, that is distributed as:

$$\varepsilon_i \mid CD \sim \mathcal{R} \left(\frac{1}{2\rho N_p} \right) \tag{28}$$

The CD probability, P^{cd}, and its complementary MD Probability, P^{md} are given by:

$$P_i^{cd} = Prob\{|\hat{h}_i| > \xi \mid i \in \mathcal{C}\} = e^{-\frac{\rho N_p \xi^2}{1 + \gamma_i^2 \rho N_p}} \tag{29}$$
$$P_i^{md} = 1 - P_i^{cd} \quad \forall i \in \mathcal{C} \tag{30}$$

Similarly, the FA probability, P^{fa}, and its complementary CR Probability, P^{cr} are given by:

$$P_i^{fa} = Prob\{|\hat{h}_i| > \xi \mid i \notin \mathcal{C}\} = e^{-\xi^2 \rho N_p} \tag{31}$$
$$P_i^{cr} = 1 - P_i^{fa} \quad \forall i \notin \mathcal{C} \tag{32}$$

Evidently FA and CR probability do not depend on i hence in following we refer to P^{fa} and P^{cr} without index i.

The MSE can be written as sum of three terms that represent the different contributions from, respectively, missed detection, false alarm and correct detection:

$$MSE^{TBS} = \sum_{i=0}^{N_p} E\left[\varepsilon_i^2\right] = \sum_{i \in \mathcal{C}} E\left[\varepsilon_i^2\right] + \sum_{i \notin \mathcal{C}} E\left[\varepsilon_i^2\right] =$$
$$= \sum_{i \in \mathcal{C}} \left(P_i^{cd} E\left[\varepsilon_i^2 | CD\right] + P_i^{md} E\left[\varepsilon_i^2 | MD\right] \right) + \sum_{i \notin \mathcal{C}} P^{fa} E\left[\varepsilon_i^2 | FA\right] =$$
$$= \sum_{i \in \mathcal{C}} \frac{P_i^{cd}}{\rho N_p} + P_i^{md} \left(\gamma_i^2 - \frac{\xi^2}{e^{\xi^2/\gamma_i^2} - 1} \right) + \sum_{i \notin \mathcal{C}} P^{fa} \frac{\left(1 + \rho N_p \xi^2\right)}{\rho N_p} =$$
$$= \sum_{i \in \mathcal{C}} \left(\frac{e^{-\frac{\rho N_p \xi^2}{1 + \gamma_i^2 \rho N_p}}}{\rho N_p} + \left(1 - e^{-\frac{\rho N_p \xi^2}{1 + \gamma_i^2 \rho N_p}} \right) \right.$$
$$\left. \left(\gamma_i^2 - \frac{\xi^2}{e^{\xi^2/\gamma_i^2} - 1} \right) \right) + (N_p - N_m) \frac{e^{-\xi^2 \rho N_p} \left(1 + \rho N_p \xi^2\right)}{\rho N_p} \tag{33}$$

B. Optimal Threshold: Analytical Expression

We aim at an analytical derivation of the optimal threshold, ξ_{opt}, which minimizes the MSE. In other words, we focus on the search of the global minimum of (33):

$$\xi_{opt} = \arg \min_{\xi} MSE^{TBS}(\xi) \tag{34}$$

$$\sum_{i\in\mathcal{C}} 2\xi \left[-\frac{e^{-\frac{\rho N_p \xi^2}{1+\gamma_i^2 \rho N_p}}}{1+\gamma_i^2 \rho N_p} + \frac{1+e^{\frac{\xi^2}{\gamma_i^2}}\left(\frac{\xi^2}{\gamma_i^2}-1\right)}{\left(e^{\frac{\xi^2}{\gamma_i^2}}-1\right)^2}\left(1-e^{-\frac{\rho N_p \xi^2}{1+\gamma_i^2 \rho N_p}}\right) + \frac{\rho N_p e^{-\frac{\rho N_p \xi^2}{1+\gamma_i^2 \rho N_p}}}{1+\gamma_i^2 \rho N_p}\left(\gamma_i^2 - \frac{\xi^2}{e^{\frac{\xi^2}{\gamma_i^2}}-1}\right)\right]$$
$$-2\rho N_p \xi^3 e^{-\rho N_p \xi^2}(N_p - N_m) = 0 \qquad (35)$$

Minimum is computed by setting the first derivative of (33) equal to zero, as reported in (35). As expected the root of (35) depend on the power profile of the channel γ_i^2, as well as on the pilot energy to noise ratio, ρ. This means that in order to derive ξ_{opt} the knowledge of the statistical of channel is required. [1] Even if, under the hypothesis of perfect channel statistics knowledge the optimal channel estimator in term of Minimum MSE is the Wiener filter, it is worthwhile to investigate also the statistics-aware TBS approach. Assuming a uniform power profile, where all the N_m paths have equal energy, in this case:

$$\gamma_i^2 = \begin{cases} \frac{1}{N_m} & i \in \mathcal{C} \\ 0 & i \notin \mathcal{C} \end{cases} \qquad (36)$$

Substituting (36) in (33), we obtain the analytical expression of the MSE for uniform power profile channel:

$$MSE^{TBS} = N_m \left[\frac{e^{-\frac{\rho N_m N_p \xi^2}{N_m + \rho N_p}}}{\rho N_p} + \left(1 - e^{-\frac{\rho N_m N_p \xi^2}{N_m + \rho N_p}}\right)\right.$$
$$\left. \left(\frac{1}{N_m} - \frac{\xi^2}{e^{N_m \xi^2}-1}\right)\right] + (N_p - N_m)\frac{e^{-\xi^2 \rho N_p}\left(1+\rho N_p \xi^2\right)}{\rho N_p} \qquad (37)$$

Thus, considering (36), the equation (34) becomes (38). In order to simplify the expressions in (37) and (38), we assume that $N_m + \rho N_p \approx \rho N_p$. This assumption is justified even in low SNR regions, since the number of pilot N_p is typically large enough to ensure that $\rho N_p >> N_m$. Under this assumption we can rewrite (37) as:

$$MSE^{TBS} \approx 1 + \frac{N_m e^{-N_m \xi^2}}{\rho N_p} + e^{-N_m \xi^2}\left(1+N_m \xi^2\right)$$
$$+ (N_p - N_m)\frac{e^{-\xi^2 \rho N_p}\left(1+\rho N_p \xi^2\right)}{\rho N_p} \qquad (39)$$

Setting the first derivative of (39) equal to zero we obtain:

$$e^{(-N_m + \rho N_p)\xi^2} = \frac{N_p - N_m}{\rho N_p \xi^2 - 1}\left(\frac{\rho N_p \xi}{N_m}\right)^2 \qquad (40)$$

Note that both (38) and (40) are transcendental equations, containing exponential terms and polynomial terms, thus they can be solved numerically. Alternatively we approximate the

polynomial term, (i.e. the right term in (40)), with its asymptotic value[2] in order to obtain a closed form solution:

$$\xi_{opt} \approx \sqrt{\frac{\ln\left(\frac{(N_p - N_m)\rho N_p}{N_m^2}\right)}{\rho N_p - N_m}} \qquad (41)$$

As figure 2 shows, the threshold values obtained using formula (41), and those obtained solving the equation (38) are very close for all SNR values. This ensures that the approximations made in (39) and in (41) are extremely tight.

C. Sub-Optimal Threshold

As discussed previously, ξ_{opt} can derived only with the awareness of channel statistics. For this reason we aim at a more pragmatic method to set threshold which, will not depend on the power profile of the channel, γ_i^2. In any case this method can still require SNR estimation, that can be easily obtained, also jointly with MSS channel estimation, as illustrated in [8].

Define the *Overall Correct Rejection* probability, P_{ocr}, as the probability of CR for all the samples $i \notin \mathcal{C}$.

$$P_{ocr} = Prob(|\hat{h}_i| \leq \xi \ \forall i \notin \mathcal{C}) = (P^{cr})^{N_p - N_m} \qquad (42)$$

Its complementary $P_{ofa} = 1 - P_{ocr}$, that we have named *Overall False Alarm* Probability, is given by:

$$P_{ofa} = 1 - (1 - P^{fa})^{N_p - N_m} =$$
$$= 1 - \sum_{k=0}^{N_p - N_m} (-1)^k \binom{N_p - N_m}{k}(P^{fa})^k \qquad (43)$$

Under the assumption on $P^{fa} << 1$, which is an essential condition for the proper functioning of the algorithm, we can approximate the expression in (43) with its linear term. Moreover, since the number of paths, N_m, is usually negligible compared to the number of pilots N_p, we can approximate $N_p - N_m \approx N_p$. As a consequence of these approximations (43) becomes:

$$P_{ofa} \approx N_p P^{fa} = N_p e^{-\xi^2 \rho N_p} \qquad (44)$$

We can extract the threshold ξ and obtain the design formula where $\overline{P_{ofa}}$ is now a parameter:

$$\xi = \sqrt{\frac{\ln\left(N_p / \overline{P_{ofa}}\right)}{\rho N_p}} \qquad (45)$$

In figure 2 the threshold values obtained using sub-optimal strategy are compared with the optimal threshold obtained

$$2N_m\xi\left[-\frac{e^{-\frac{N_m\rho N_p\xi^2}{N_m+\rho N_p}}N_m}{N_m+\rho N_p}+\frac{1+e^{\xi^2 N_m}\left(\xi^2 N_m-1\right)}{\left(e^{\xi^2 N_m}-1\right)^2}\left(1-e^{-\frac{N_m\rho N_p\xi^2}{N_m+\rho N_p}}\right)+\frac{\rho N_p e^{-\frac{N_m\rho N_p\xi^2}{N_m+\rho N_p}}}{N_m+\rho N_p}\left(\frac{1}{N_m}-\frac{\xi^2}{e^{\xi^2 N_m}-1}\right)\right] \quad (38)$$
$$-2(N_p-N_m)\rho N_p\xi^3\mathrm{e}^{-\rho N_p\xi^2}=0$$

by solving the equation (38), and also with those obtained adopting some other setting strategies. It is worthwhile to note that using this sub-optimal strategy the resulting thresholds are very close the optimal solution for all the considered SNR valus.
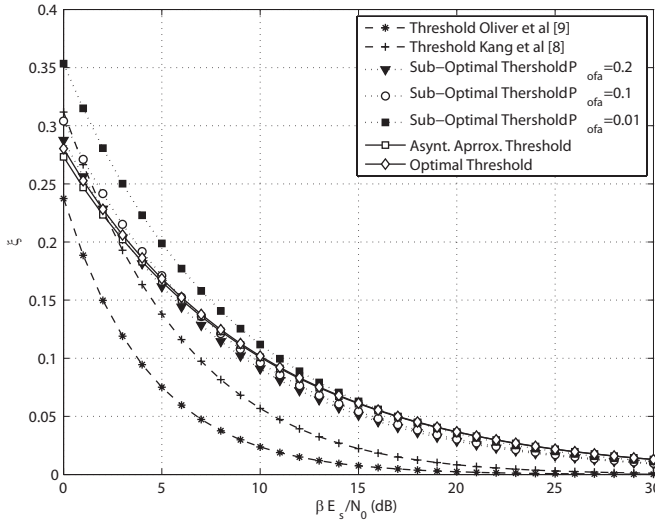

Fig. 2. Comparison between sub-optimal thresholds obtained with several values of $\overline{P_{ofa}}$; the optimal threshold obtained by solving the equation (38)); the asymptotic approximation defined in (41); Considering an Uniform channel with $N_m=6$

V. NUMERICAL RESULTS

In this section our intent is first to validate the analytical model through numerical results obtained by simulations, and second to assess the performance and the robustness against channel with different power delay profiles, of the considered TBS channel estimation algorithms. We use Monte Carlo simulations to evaluate the performance of channel estimators in term of MSE using two different channel model: a uniform channel and a more realistic ITU-TU6 channel, whose power delay profile are illustrated in Table I. We consider an OFDM system compliant with DVB-SH standard [10], which uses $N=1024$ carriers and provides $N_p=71$ scattered pilots uniformly distributed over $N_a=852$ active subcarriers, thus the pilot frequency spacing is equal to $12f_u$. In particular, in our results we compare TBS channel estimation algorithms using: the optimal threshold obtained by solving the equation (38); its asymptotical approximation provided in equation (40); the sub-optimal threshold obtained according equation (41) corresponding to several values of $\overline{P_{ofa}}$; the criterion for

TABLE I
POWER DELAY PROFILE OF THE CONSIDERED MULTIPATH CHANNELS

	Path	#1	#2	#3	#4	#5	#6
Uniform	Delay[μs]	0.0	1.0	2.0	3.0	4.0	5.0
Channel	Power[dB]	0	0	0	0	0	0
ITU	Delay[μs]	0.0	0.2	0.5	1.6	2.3	5.0
TU6	Power[dB]	-3	0	-2	-6	-8	-10

threshold setting proposed in [8, eq. (21)] and also reported in (21), as well as the criterion proposed in [9, eq. (6)] and also reported in (22). Moreover we also report as benchmarks the performance of LS and Wiener filter channel estimation. As first outcome we observe that, as shown in figure 3, the numerical results validate the analytical expression for the generic TBS channel estimator reported in (37). Comparing methods that do not require channel statistics knowledge, TBS using the proposed sub-optimal threshold reaches the best performance, with a gain larger than 9 dB and 6 dB respect to LS and TBS using Kang criterion respectively. Even Oliver criterion, which takes advantage of the channel statistics knowledge, is outperformed by the proposed criterion. Moreover, TBS with sub-optimal threshold performance is very close to the TBS using the optimal threshold, and also close to Wiener filter channel estimation, which is the Minimum MSE method. In figure 4 it is shown that also considering a more realistic channel such as ITU-TU6 the MSE performance is nearly unvaried. Again TBS using sub-optimal threshold offers the best performance for both power delay profiles considered.

Even if we observe that the better performance has been obtained when $\overline{P_{ofa}}=0.1$, also using $\overline{P_{ofa}}=0.2$ and $\overline{P_{ofa}}=0.01$ the algorithm maintains its performance very close to the best case. For this reason we can affirm that this approach is not sensitive to different values of $\overline{P_{ofa}}$, different power delay profiles.

VI. CONCLUSIONS

In this paper, we have addressed the problem OFDM Channel Estimation keeping into consideration that channel statics knowledge is hardly available at the receiver, therefore Wiener Filtering, which is the best Channel estimator in MSE sense, is not always applicable. For this reason, we have treated the OFDM data-aided channel estimation based on the MSS threshold-crossing selection of the CIR. With an appropriate choice of threshold, which is clearly a crucial aspect, this approach can significantly reduce noise effects by taking advantage of typical sparseness of CIR of multipath fading channels, especially for SFNs. We have provided a novel analytical characterization for this approach, through which we derive a closed form for the estimate MSE, and
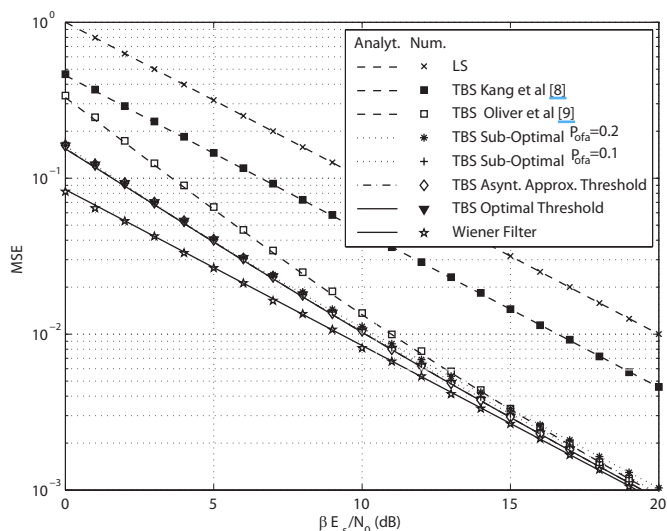
Fig. 3. Comparison between the performance of several threshold based MMS selection with Uniform channel
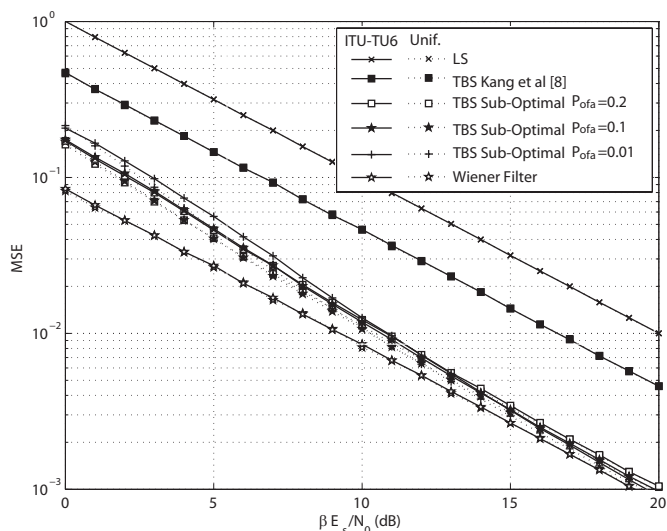

Fig. 4. Comparison between the performance of several threshold based MMS selection with TU6 and uniform channel

consequently, the optimum threshold in the minimum MSE sense. As expected, the optimal threshold depends on the channel power delay profile, thus it can be derived only with the awareness of the channel statistics. For these reason, we propose a novel sub-optimal method for choice of the threshold that do not require any knowledge of channel statistic. We show that this sub-optimal method not only outperforms previous approaches based on heuristically set thresholds, but it also achieves performance very close to the optimal case, and at same time being robust to various power delay profiles. Furthermore, in the case of sparse channels, performance is shown to be close to Wiener channel estimation.

REFERENCES

[1] P. Hoeher, S. Kaiser, P. Robertson, "Two-dimensional pilot symbol aided channel estimation by Wiener filtering", *in Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Munich, Germany, vol 3, pp. 1845-1848, April 1997.

[2] P. Hoeher, S. Kaiser, P. Robertson, "Pilot-Symbol-Aided Channel Estimation in Time and Frequency", *in Proc. Comm. Theory Mini-Conf. (CTMC)* within *IEEE Global Telecommun. Conf. (GLOBECOM97)*, Phoenix, USA, pp. 90-96, 1997.

[3] J. J. van de Beek, O. Edfors, M. Sandell, S. K. Wilson, and P. O. Borjesson, "On Channel Estimation in OFDM Systems", *in Proc. 45th IEEE Vehicular Technology Conference (VTC)*, Chicago, USA, , vol. 2, pp. 815 - 819, July 1995.

[4] Y. Li, L. J. Cimini Jr., N. R. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels", *IEEE Trans. Commun.*, vol. 46, n. 7, pp. 902-915, July 1998.

[5] O. Edfors, M. Sandell, J. J. van de Beek, S. K. Wilson, and P. O. Borjesson, "Analysis of DFT-based channel estimators for OFDM", *Wireless Personal Communications: An International Journal*, vol .12 n. 1, p.55-70, January 2000.

[6] H. Minn, V.K. Bhargava, "An investigation into time-domain approach for OFDM channel estimation", *IEEE Trans. Broadcast.*, vol. 46, n. 4, pp. 240 - 248, December 2000.

[7] M. Morelli, U. Mengali, "A Comparison of Pilot-Aided Channel Estimation Methods for OFDM Systems", *IEEE Trans. on Signal Processing*, Vol. 49, n. 2, pp: 3065-3073, December 2001.

[8] Y. Kang, K. Kim, H. Park, "Efficient DFT-based channel estimation for OFDM systems on multipath channels" *IET Communications*, vol. 1, n. 2, pp. 197 - 202, April 2007.

[9] J. Oliver, R. Aravind, K.M.M. Prabhu, "Sparse channel estimation in OFDM systems by threshold-based pruning", *Electronics Letters*, vol. 44, n. 13, pp. 830-832, June 2008.

[10] "Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for satellite transmission to handheld. (DVB-SH)." *ETSI EN 302 583 V1.0.0*, June 2007.