ENHANCING SPECTRAL EFFICIENCY IN ADVANCED MULTICARRIER TECHNIQUES: A CHALLENGE

Leonardo G. Baltar, Tobias Laas, Michael Newinger, Amine Mezghani, Josef A. Nossek

Institute for Circuit Theory and Signal Processing
Technische Universität München
Arcisstr. 21, 80290 München, Germany
Email: leo.baltar@tum.de

ABSTRACT

Advanced multicarrier systems, like the Offset-QAM filter bank based (OQAM-FBMC) ones, are gaining importance as candidates for the physical layer of the 5-th generation of wireless communications. One of the main advantages of FBMC, when compared to traditional cyclic prefix based OFDM, is its higher spectral efficiency. However, this gain can be lost again if the problem of training based channel estimation is not tackled correctly. This is due to the memory inserted by the longer pulse shaping and the loss of orthogonality of overlapping subcarriers. In this paper we approach the problem of training based channel estimation for FBMC systems. We propose an iterative algorithm based on the expectation maximization (EM) maximum likelihood (ML) that reduces the overhead and consequently improves the spectral efficiency.

***Index Terms*—** OQAM; Filter Bank Multicarrier; Channel Estimation; ML estimation; Expectation Maximization

1. INTRODUCTION

We consider FBMC systems in wireless environments with multipath propagation. In contrast to CP-OFDM, where a rectangular pulse shaping is used, we take a finite impulse response (FIR) prototype filter with a duration greater than the symbol period, but as in CP-OFDM it is modulated by complex exponentials. Consequently, more spectrally concentrated subcarriers are obtained which only overlap with the two adjacent ones. Moreover, the FBMC system does not include any guard interval, which also improves spectral efficiency, at the cost of higher complexity.

In FBMC, orthogonality, i.e. inter-symbol interference (ISI) and inter-channel interference (ICI)-free received symbols, can only be guaranteed by the so called OQAM [1], where the symbols' real and imaginary parts are staggered by $T/2$ and T is the QAM symbol period in each subcarrier. Furthermore, the prototype filter can be designed according to different goals, but we restrict ourselves to an FIR approximation of the root raised cosine (RRC) with roll-off one. This choice of the prototype will indeed introduce some ISI and ICI, but if its degree is high enough, the interference is negligible compared to the other impairments, for example, ISI and ICI caused by a multipath channel.

This work is partially supported by the EU project EMPhAtiC (ICT-318362).

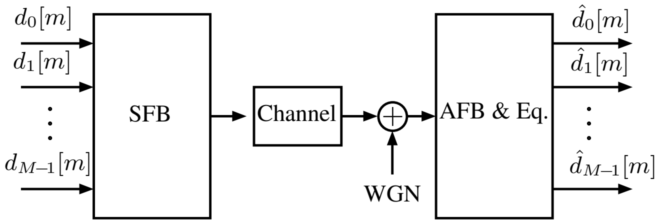

Fig. 1. FBMC System Overview.

In [2, 3], we have presented per-subcarrier fractionally spaced equalizers that almost remove all ISI and ICI caused by the multipath channel. We have assumed perfect channel impulse response (CIR) knowledge at the receiver side. In [4] we have presented a method for the estimation of the narrowband multipath channel viewed in each subcarrier using an EM based algorithm in order to increase the spectral efficiency. In this contribution, we first present how to estimate the broadband CIR without taking care of the spectral efficiency, then we extend the results of [4] for the broadband channel estimation to improve the spectral efficiency.

2. ADVANCED MULTICARRIER: OQAM BASED FILTER BANK MULTICARRIER

A high level model of the FBMC system is shown in Fig. 1. In this transmultiplex system, a synthesis filter bank (SFB) performs a frequency division multiplexing of the T seconds long QAM data symbols $d_k[m]$ at the transmitter. An analysis filter bank (AFB) at the receiver separates the data onto each subcarrier. We assume a slowly fading frequency selective channel. Usually M_u out of M subcarriers are used.

Here we regard exponentially modulated SFBs and AFBs, i.e. only one low-pass filter has to be designed and the other

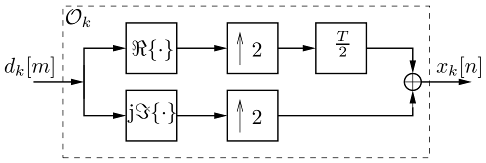

Fig. 2. O-QAM staggering for odd indexed subcarrier

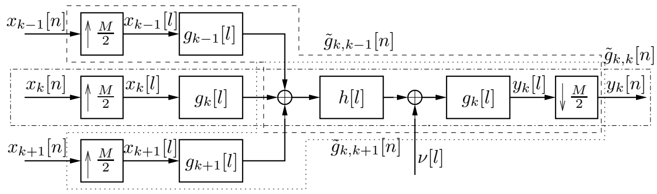

Fig. 3. Subchannel model for the FBMC/OQAM system.

sub-filters are obtained by modulating it as follows [5]

$$g_k[l] = g_{\mathrm{P}}[l] \exp\left(\mathrm{j}\frac{2\pi}{M}k\left(l - \frac{L_{\mathrm{P}}-1}{2}\right)\right),\ l = 0, \ldots, L_{\mathrm{P}}-1,$$

where $g_{\mathrm{P}}[l]$ is the impulse response of this prototype filter of degree $L_{\mathrm{P}} - 1$. The prototype chosen here is an RRC filter with roll-off factor one and consequently only the spectrum of contiguous subcarriers overlap. The non-contiguous subcarriers are separated by the high stop-band attenuation. We define $L_{\mathrm{P}} = KM+1$, where K is the time overlapping factor that determines how many symbols superimpose. K should be kept as small as possible not only to limit the complexity but also to reduce the time-domain spreading of the symbols and the transmission latency.

To maintain the orthogonality between all subcarriers and for all time instants, the complex QAM input symbols $d_k[m]$ are OQAM staggered. We illustrate the OQAM staggering for odd indexed subcarriers in Fig. 2. For even indexed subcarriers the $T/2$ delay is placed at the lower branch. The OQAM de-staggering is performed at the receiver by the application of flow-graph reversal, substitution of up-samplers by down-samplers and exchange of $\Re\{\cdot\}$ and $\mathrm{j}\Im\{\cdot\}$.

After the OQAM staggering, the subcarrier signals are up-sampled by $M/2$, filtered and added. A broadband signal is then generated and digital-to-analog converted into an IQ baseband signal that is analog processed and transmitted. At the receiver side the RF signal is amplified, brought to baseband, filtered and then analog-to-digital converted. The received signal is then filtered and down-sampled by $M/2$.

The fact that only contiguous subcarriers overlap, allows us to construct the model for one subcarrier shown in Fig. 3. The inputs $x_k[n]$ are OQAM symbols and the received subcarrier signals $y_k[n]$ still have to be equalized and de-staggered before further processing of the QAM symbols. As a consequence, in this model the input and output sampling rates are $2/T$. We assume here a multipath channel with perfect frequency synchronization (no carrier frequency offset or Doppler shift). A time offset can be incorporated in the CIR.

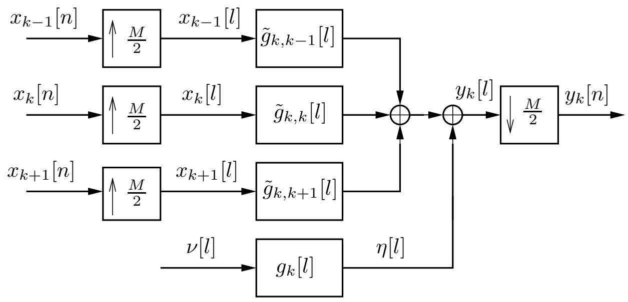

Fig. 4. Subcarrier model for broadband channel estimation.

3. SYSTEM MODEL AND CHANNEL IMPULSE RESPONSE ESTIMATION

Let us assume that per-subcarrier linear or decision feedback equalizers are employed. As shown in [2, 3], the CIR knowledge is necessary for their design. Here we further assume that the training sequences are not only employed in the observed subcarriers, but also in their adjacent ones. Later we will alleviate this and only the observed subcarriers will be reserved for training.

The subcarrier model of Fig. 3 can be simplified as shown in Fig. 4 for the purpose of estimating the broadband CIR $h[l]$. The output $y_k[n]$ contains the downsampled received samples at the OQAM symbol rate and is defined as

$$\begin{aligned} y_k[n] =& \tilde{g}_{k,k}[n]*x_k[n]+\tilde{g}_{k,k-1}[n]*x_{k-1}[n] \\ &+\tilde{g}_{k,k+1}[n]*x_{k+1}[n]+\eta_k[n], \end{aligned} \quad (1)$$

where $*$ represents linear convolution, and $\tilde{g}_{k,k}[n]$, $\tilde{g}_{k,k-1}[n]$ and $\tilde{g}_{k,k+1}[n]$ are the downsampled impulse responses of length $L_{\tilde{g}} = \lceil \frac{2L_{\mathrm{P}}+L_h-2}{M/2} \rceil$ that result from the convolution of tha transmit filters $g_k[l]$, $g_{k-1}[l]$ and $g_{k+1}[l]$, the receive filter $g_k[l]$ and the frequency selective channel $h[l]$. $\eta_k[n]$ is the downsampled narrowband colored noise.

One can stack the coefficients of the impulse responses $\tilde{g}_{k,k}[n]$, $\tilde{g}_{k,k-1}[n]$ and $\tilde{g}_{k,k+1}[n]$ in vectors $\tilde{\mathbf{g}}_{k,k}$, $\tilde{\mathbf{g}}_{k,k-1}$ and $\tilde{\mathbf{g}}_{k,k+1} \in \mathbb{C}^{L_{\tilde{g}}}$ and decompose them as the products of matrices $\bar{\mathbf{G}}_{k,k}$, $\bar{\mathbf{G}}_{k,k-1}$ and $\bar{\mathbf{G}}_{k,k+1} \in \mathbb{C}^{L_{\tilde{g}}\times L_h}$ with the CIR vector $\mathbf{h} \in \mathbb{C}^{L_h}$ as $\tilde{\mathbf{g}}_{k,k} = \bar{\mathbf{G}}_{k,k}\mathbf{h}$, $\tilde{\mathbf{g}}_{k,k-1} = \bar{\mathbf{G}}_{k,k-1}\mathbf{h}$, $\tilde{\mathbf{g}}_{k,k+1} = \bar{\mathbf{G}}_{k,k+1}\mathbf{h}$. We then have $\bar{\mathbf{G}}_{k,k} = \mathbf{J}_{\mathrm{DS}}^G\mathbf{G}_{k,k}$, $\bar{\mathbf{G}}_{k,k-1} = \mathbf{J}_{\mathrm{DS}}^G\mathbf{G}_{k,k-1}$, $\bar{\mathbf{G}}_{k,k+1} = \mathbf{J}_{\mathrm{DS}}^G\mathbf{G}_{k,k+1}$, where $\mathbf{J}_{\mathrm{DS}}^G$ is a downsampling matrix with its ℓ-th row given by $\mathbf{e}_q^T \in \{0,1\}^{L_h}$ for $q = (\ell-1)M/2+1$ and $\ell \in \{1,2,...,L_{\tilde{g}}\}$. $\mathbf{e}_q$ is a unity vector with 1 in the q-th position and 0s elsewhere. $\mathbf{G}_{k,k}$, $\mathbf{G}_{k,k-1}$ and $\mathbf{G}_{k,k+1} \in \mathbb{C}^{(2L_{\mathrm{P}}-1)\times L_h}$ are the convolution matrices generated by the impulse responses $(g_k*g_k)[l]$, $(g_k*g_{k-1})[l]$ and $(g_k*g_{k+1})[l]$. Moreover, we have that $\boldsymbol{\eta}_k = \boldsymbol{\Gamma}_k\boldsymbol{\nu}$, where $\boldsymbol{\Gamma}_k \in \mathbb{C}^{L_{\mathrm{o}}\times(L_{\mathrm{P}}+L_{\mathrm{o}}-1)}$ is the corresponding downsampled version of the convolution matrix generated from the impulse response $g_k[l]$.

Now we can stack the samples $y_k[n]$ in a vector to obtain

$$\begin{aligned} \mathbf{y}_k =& \mathbf{X}_k\tilde{\mathbf{g}}_{k,k}+\mathbf{X}_{k-1}\tilde{\mathbf{g}}_{k,k-1}+\mathbf{X}_{k+1}\tilde{\mathbf{g}}_{k,k+1}+\boldsymbol{\Gamma}_k\boldsymbol{\nu}[l], \\ =& (\mathbf{X}_k\bar{\mathbf{G}}_{k,k}+\mathbf{X}_{k-1}\bar{\mathbf{G}}_{k,k-1}+\mathbf{X}_{k+1}\bar{\mathbf{G}}_{k,k+1})\mathbf{h}+\boldsymbol{\Gamma}_k\boldsymbol{\nu}[l], \end{aligned}$$

where $\mathbf{X}_{k-1}, \mathbf{X}_k, \mathbf{X}_{k+1} \in \mathbb{C}^{L_o \times L_{\bar{g}}}$ are Hankel matrices containing $x_{k-1}[n], x_k[n]$ and $x_{k+1}[n]$, the training sequences of length $L_T = L_o + L_{\bar{g}} - 1$ each. By defining $\mathbf{S}_k = \mathbf{X}_k \bar{\mathbf{G}}_{k,k}$ and $\mathbf{U}_k = \mathbf{X}_{k-1}\bar{\mathbf{G}}_{k,k-1} + \mathbf{X}_{k+1}\bar{\mathbf{G}}_{k,k+1}$, we get

$$\mathbf{y}_k = (\mathbf{S}_k + \mathbf{U}_k)\mathbf{h} + \boldsymbol{\eta}_k \tag{2}$$

for each observed subcarrier. Finally, we stack the M_t vectors with the outputs of the observations subcarriers to obtain

$$\begin{bmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_{M_t-1} \end{bmatrix} = \left(\begin{bmatrix} \mathbf{S}_0 \\ \mathbf{S}_1 \\ \vdots \\ \mathbf{S}_{M_t-1} \end{bmatrix} + \begin{bmatrix} \mathbf{U}_0 \\ \mathbf{U}_1 \\ \vdots \\ \mathbf{U}_{M_t-1} \end{bmatrix} \right) \mathbf{h} + \begin{bmatrix} \boldsymbol{\Gamma}_0 \\ \boldsymbol{\Gamma}_1 \\ \vdots \\ \boldsymbol{\Gamma}_{M_t-1} \end{bmatrix} \boldsymbol{\nu},$$
$$\mathbf{y} = (\mathbf{S} + \mathbf{U})\mathbf{h} + \boldsymbol{\eta}. \tag{3}$$

It should be noted that the vectors $\mathbf{y}_k$ that are collected into $\mathbf{y}$ usually do not belong to contiguous subcarriers, i.e. the observations are sparsely taken on the subcarrier axis. This allows the use of training that is frequency multiplexed with data symbols. This way, a higher spectral efficiency is obtained by reducing the number of transmitted training symbols.

3.1. Broadband ML

We can see that in the linear model of (3) the noise $\boldsymbol{\eta}$ is Gaussian distributed with zero mean and covariance matrix $\mathbf{R}_\eta = \sigma_\nu^2 \boldsymbol{\Gamma}\boldsymbol{\Gamma}^H = \mathrm{diag}(\mathbf{R}_{\eta,0}, \mathbf{R}_{\eta,1}, ..., \mathbf{R}_{\eta,M_t-1})$ and the observation $\mathbf{y}$ given $\mathbf{h}$ is then Gaussian distributed. The maximum likelihood (ML) estimate of $\mathbf{h}$ in this case is given by

$$\hat{\mathbf{h}} = \arg \max_{\mathbf{h} \in \mathbb{C}^{L_h}} p(\mathbf{y}|\mathbf{h}) = \arg \min_{\mathbf{h}} J(\mathbf{h}), \tag{4}$$

where $J(\mathbf{h}) = (\mathbf{y} - (\mathbf{S}+\mathbf{U})\mathbf{h})^H \mathbf{R}_\eta^{-1} (\mathbf{y} - (\mathbf{S}+\mathbf{U})\mathbf{h})$.

Since $\mathbf{R}_\eta$ is independent of $\mathbf{h}$ and if $((\mathbf{S}+\mathbf{U})^H \mathbf{R}_\eta^{-1}(\mathbf{S}+\mathbf{U}))$ is invertible, as we will assume here, we just need to apply the derivative and make it equal to zero

$$\frac{\partial J(\mathbf{h})}{\partial \mathbf{h}^H} = (\mathbf{S}+\mathbf{U})^H \mathbf{R}_\eta^{-1} (\mathbf{S}+\mathbf{U})\hat{\mathbf{h}} - (\mathbf{S}+\mathbf{U})^H \mathbf{R}_\eta^{-1} \mathbf{y} = \mathbf{0}.$$

The ML estimate of $\mathbf{h}$ is then given by

$$\hat{\mathbf{h}} = \left((\mathbf{S}+\mathbf{U})^H \mathbf{R}_\eta^{-1} (\mathbf{S}+\mathbf{U})\right)^{-1} (\mathbf{S}+\mathbf{U})^H \mathbf{R}_\eta^{-1} \mathbf{y}. \tag{5}$$

The covariance matrix of the estimation error $\boldsymbol{\Delta}\hat{\mathbf{h}} = (\hat{\mathbf{h}} - \mathbf{h})$ of the broadband ML estimator is given by

$$\mathbf{R}_{\boldsymbol{\Delta}\hat{\mathbf{h}}} = \mathrm{E}\left[\boldsymbol{\Delta}\hat{\mathbf{h}}\boldsymbol{\Delta}\hat{\mathbf{h}}^H\right] = \left((\mathbf{S}+\mathbf{U})^H \mathbf{R}_\eta^{-1} (\mathbf{S}+\mathbf{U})\right)^{-1}. \tag{6}$$

As a consequence, the theoretical MSE of the broadband ML estimator is given by $\epsilon = \frac{\sigma^2 M_u}{M} \mathrm{tr}\left\{\mathbf{R}_{\boldsymbol{\Delta}\hat{\mathbf{h}}}\right\}$.

When multicarrier systems like FBMC or CP-OFDM are deployed, the number of subcarriers filled with data and training M_u is smaller than M, in order to allow for upsampling, filtering and D/A conversion. Even if all M_u subcarriers are filled with training symbols only, the estimation of the broadband CIR can only be performed in a fraction of its total frequency response. As a consequence $\mathbf{R}_{\boldsymbol{\Delta}\hat{\mathbf{h}}}$ will become ill conditioned or even singular. The reason is that the regions of the channel frequency response that are not excited cannot and need not be reliably estimated.

To solve this problem we perform the projection of the lower dimensional $\mathbf{h}_{DS} \in \mathbb{C}^{L_{h_{DS}}}$ into the higher dimensional $\mathbf{h}$ given by $\mathbf{h} = \mathbf{A}\mathbf{h}_{DS}$. This linear operation can be seen as $\mathbf{h}_{DS}$, the broadband CIR that can be estimated in the occupied spectrum, being fractionally upsampled by a factor of $L_{frac} = L_h / L_{h_{DS}}$, where $L_{h_{DS}} = \lfloor \frac{M_u}{M} L_h \rfloor$. This operation is performed in three steps: upsampling by a factor of L_h, low-pass filtering and downsampling by a factor $L_{h_{DS}}$. Mathematically, this can be described by

$$\mathbf{A} = \mathbf{J}_{DS}^A \left[\mathbf{0}_A \quad \mathbf{I}_{L_{h_{DS}} L_h} \quad \mathbf{0}_A\right] \mathbf{G}_{int} \mathbf{J}_{US}, \tag{7}$$

where $\mathbf{J}_{DS}^A$ is a downsampling matrix with its ℓ-th row given by $\mathbf{e}_q^T \in \{0,1\}^{(L_{h_{DS}} L_h)}$ for $q = (\ell - 1)L_{h_{DS}} + 1$ and $\ell \in \{1, 2, ..., L_h\}$, $\mathbf{J}_{US}$ is an upsampling matrix with its ℓ-th column given by $\mathbf{e}_q \in \{0,1\}^{(L_{h_{DS}} L_h)}$ for $q = (\ell-1)L_h + 1$ and $\ell \in \{1, 2, ..., L_{h_{DS}}\}$, $\mathbf{G}_{int} \in \mathbb{R}^{(L_{h_{DS}} L_h + 2(d_g - 1)) \times (L_{h_{DS}} L_h)}$ is a convolution matrix obtained from the interpolation filter $\mathbf{g}_{int} \in \mathbb{R}^{2d_g - 1}$, $\mathbf{0}_A \in \{0\}^{(L_{h_{DS}} L_h) \times (d_g - 1)}$. $g_{int}[n]$ is taken as an FIR approximation of a raised cosine filter with a sharp roll-off $\alpha = 0.001$, transfer function degree of $L_{g_{int}} = 10 L_{h_{DS}} L_h$ and group delay $d_g = 5L_{h_{DS}} + 1$.

By substituting $\mathbf{h}$ in the linear model (3) we obtain

$$\hat{\mathbf{h}}_{DS} = \left(\mathbf{A}^H (\mathbf{S}+\mathbf{U})^H \mathbf{R}_\eta^{-1} (\mathbf{S}+\mathbf{U})\mathbf{A}\right)^{-1} \mathbf{A}^H (\mathbf{S}+\mathbf{U})^H \mathbf{R}_\eta^{-1} \mathbf{y}.$$

The corresponding MSE is given by

$$\epsilon_{DS} = \frac{\sigma^2 M_u}{M} \mathrm{tr}\left\{\left(\mathbf{A}^H (\mathbf{S}+\mathbf{U})^H \mathbf{R}_\eta^{-1} (\mathbf{S}+\mathbf{U})\mathbf{A}\right)^{-1}\right\}. \tag{8}$$

3.2. EM-ML Estimator

We now assume that the $\mathbf{U}_k$s are unknown, this means that the subcarriers adjacent to the ones with training contain data. Moreover, in [4] we have employed a subcarrier model where narrowband propagation channels were estimated in the subcarriers. Now we modify the subcarrier observations vector to contain a mix of two models: for the training part we take the broadband model and for the unknown data we take the narrowband model.

Now, we can rewrite the observations vector in (2) obtained for each subcarrier at the receiver as

$$\mathbf{y}_k = \mathbf{S}_k \mathbf{h} + \mathbf{H}_k \mathbf{u}_k' + \boldsymbol{\eta}_k, \tag{9}$$

where $\mathbf{H}_k \in \mathbb{C}^{L_o \times L_{\bar{g}'}}$, with $L_{\bar{g}'} = L_{h_k} + L_{\bar{g}'} - 1$ and $L_{\bar{g}'} = \lceil \frac{2L_P - 1}{M/2} \rceil$, is a convolution matrix obtained from the narrowband propagation channel $\mathbf{h}_k \in \mathbb{C}^{L_{h_k}}$ observed in each subcarrier that is calculated from the broadband channel by the transformation $\mathbf{h}_k = \mathbf{B}_k\mathbf{h}$. Thereby, the following definition holds

$$\mathbf{B}_k = \begin{bmatrix}\mathbf{I}_{L_{h_k}} & \mathbf{0}_{B1}\end{bmatrix}\mathbf{F}^{\mathrm{H}}_{L_{h_k}M_i}\begin{bmatrix}\mathbf{0}_{B2} & \mathbf{I}_{L_{h_k}M_i} & \mathbf{0}_{B3}\end{bmatrix}\mathbf{F}_{M_f}\begin{bmatrix}\mathbf{I}_{L_h}\\ \mathbf{0}_{B4}\end{bmatrix},$$

with $\mathbf{F}_{M_f}$ being an M_f-DFT matrix, $M_f = ML_{h_k}M_i$, $\mathbf{0}_{B1} \in \{0\}^{L_{h_k}\times(L_{h_k}(M_i-1))}$, $\mathbf{0}_{B2} \in \{0\}^{(L_{h_k}M_i)\times(kL_{h_k}M_i)}$, $\mathbf{0}_{B3} \in \{0\}^{(L_{h_k}M_i)\times((M-1-k)L_{h_k}M_i)}$, $\mathbf{0}_{B4} \in \{0\}^{(M_f-L_h)\times L_h}$, M_i is a resolution factor for the calculation's precision of the $\mathbf{h}_k$s.

Moreover, the following equalities hold

$$\mathbf{U}_k\mathbf{h} = \mathbf{U}'_k\mathbf{B}_k\mathbf{h} = \mathbf{U}'_k\mathbf{h}_k = \mathbf{H}_k\mathbf{u}'_k = \mathbf{H}_k\mathbf{G}_{\mathrm{u},k}\mathbf{x}_{\mathrm{u},k}, \quad (10)$$

where $\mathbf{U}'_k$ is a convolution matrix obtained for the narrowband subcarrier model. Note that $\mathbf{G}_{\mathrm{u},k} \in \mathbb{C}^{L_{\tilde{g}'}\times 2L'_{\mathrm{t}}}$ and $\mathbf{x}_{\mathrm{u},k} \in \mathbb{C}^{2L'_{\mathrm{t}}}$, where $L'_{\mathrm{t}} = L_{\tilde{g}'} + L_{\mathrm{o}} - 1$.

Stacking all the subcarrier observations, we get

$$\begin{bmatrix}\mathbf{y}_0\\ \mathbf{y}_1\\ \vdots\\ \mathbf{y}_{M_{\mathrm{t}}-1}\end{bmatrix} = \begin{bmatrix}\mathbf{S}_0\\ \mathbf{S}_1\\ \vdots\\ \mathbf{S}_{M_{\mathrm{t}}-1}\end{bmatrix}\mathbf{h} + \mathbf{H}_{\mathrm{E}}\begin{bmatrix}\mathbf{u}'_0\\ \mathbf{u}'_1\\ \vdots\\ \mathbf{u}'_{M_{\mathrm{t}}-1}\end{bmatrix} + \begin{bmatrix}\boldsymbol{\Gamma}_0\\ \boldsymbol{\Gamma}_1\\ \vdots\\ \boldsymbol{\Gamma}_{M_{\mathrm{t}}-1}\end{bmatrix}\boldsymbol{\nu},$$

$$\mathbf{y} = \mathbf{S}\mathbf{h} + \mathbf{H}_{\mathrm{E}}\mathbf{u}' + \boldsymbol{\eta} = \mathbf{S}\mathbf{h} + \mathbf{U}\mathbf{h} + \boldsymbol{\eta}, \quad (11)$$

where $\mathbf{U} = [\mathbf{B}_0^{\mathrm{T}}\mathbf{U}_0'^{\mathrm{T}}, \mathbf{B}_1^{\mathrm{T}}\mathbf{U}_1'^{\mathrm{T}}, ..., \mathbf{B}_{M_{\mathrm{t}}-1}^{\mathrm{T}}\mathbf{U}_{M_{\mathrm{t}}-1}'^{\mathrm{T}}]^{\mathrm{T}}$ and $\mathbf{H}_{\mathrm{E}} = \mathrm{diag}(\mathbf{H}_0, \mathbf{H}_1, ..., \mathbf{H}_{M_{\mathrm{t}}-1})$.

Although the exact values of $\mathbf{u}'_k$ are unknown, their statistics are known. We then define the interference covariance matrices as $\mathbf{R}_{u,k} = \frac{\sigma_d^2}{2}\mathbf{G}_{\mathrm{u},k}\mathbf{G}_{\mathrm{u},k}^{\mathrm{H}}$ and $\mathbf{P}_{u,k} = \frac{\sigma_d^2}{2}\mathbf{G}_{\mathrm{u},k}\mathrm{diag}(1,-1,1,...)\mathbf{G}_{\mathrm{u},k}^{\mathrm{T}}$, where $\mathbf{x}_{\mathrm{u},k}$ was assumed to be i.i.d. and Gaussian distributed with zero mean and variance $\sigma_d^2/2$. This is usually a good approximation although $\mathbf{x}_{\mathrm{u},k}$ is composed of symbols taken from a finite constellation.

For the linear model of (11), the ML estimator has no closed form solution and one way to calculate it is by employing the iterative EM algorithm [6], that works here as follows: Before the first iteration, an initial channel estimate is performed by ignoring $\mathbf{U}$. This estimate is given by

$$\hat{\mathbf{h}}_0 = (\mathbf{S}^{\mathrm{H}}\mathbf{R}_\eta^{-1}\mathbf{S})^{-1}\mathbf{S}^{\mathrm{H}}\mathbf{R}_\eta^{-1}\mathbf{y}. \quad (12)$$

Then, the iterative process starts. For each iteration i, the algorithm is divided into two steps: the E-step and the M-step. In the E-step, an approximation of the ML function (here its derivative) is obtained by taking its expected value conditioned on the channel estimate in the iteration before and the observed sequence, as follows

$$\mathrm{E}\left[\frac{\partial J(\mathbf{h}_i)}{\partial\mathbf{h}_i^{\mathrm{H}}}\bigg|\,\mathbf{y},\mathbf{h}_i\right] = \Big[\mathbf{S}^{\mathrm{H}}\mathbf{R}_\eta^{-1}(\mathbf{S}+\mathrm{E}[\mathbf{U}]) + \mathrm{E}[\mathbf{U}]^{\mathrm{H}}\mathbf{R}_\eta^{-1}\mathbf{S} + \mathrm{E}[\mathbf{U}^{\mathrm{H}}\mathbf{R}_\eta^{-1}\mathbf{U}]\Big]\,\mathbf{h}_i - (\mathbf{S}+\mathrm{E}[\mathbf{U}])^{\mathrm{H}}\mathbf{R}_\eta^{-1}\mathbf{y}.$$

The result is a function of $\mathrm{E}[\mathbf{u}]$ and $\mathrm{E}[\mathbf{U}^{\mathrm{H}}\mathbf{R}_\eta^{-1}\mathbf{U}]$. $\mathrm{E}[\mathbf{u}] = \hat{\mathbf{u}}_i$ is actually an estimate of the interference term $\mathbf{u}$ in the i-th iteration. Furthermore, one can show that

$$\mathrm{E}[\mathbf{U}^{\mathrm{H}}\mathbf{R}_\eta^{-1}\mathbf{U}] = \sum_{k=0}^{M_{\mathrm{t}}-1}\mathrm{E}[\mathbf{U}_k^{\mathrm{H}}\mathbf{R}_{\eta,k}^{-1}\mathbf{U}_k] = \sum_{k=0}^{M_{\mathrm{t}}-1}\mathbf{B}_k^{\mathrm{H}}\mathrm{E}[\mathbf{U}_k'^{\mathrm{H}}\mathbf{R}_{\eta,k}^{-1}\mathbf{U}'_k]\mathbf{B}_k.$$

To express the above expectation also in terms of the estimate $\hat{\mathbf{u}}_i$, one first has to write the matrix $\mathbf{U}'_k$ as a function of the vector $\mathbf{u}'_k$ as $\mathbf{U}'_k = \sum_{\ell=1}^{L_{h_k}}\mathbf{D}_\ell\mathbf{u}'_k\mathbf{e}_\ell^T$, where $\mathbf{D}_\ell = \begin{bmatrix}\mathbf{0}_{D1} & \mathbf{I}_{L_{\mathrm{o}}} & \mathbf{0}_{D2}\end{bmatrix}$ is a matrix that selects L_{o} rows of $\mathbf{u}'_k$, $\mathbf{0}_{D1} \in \mathbb{C}^{L_{\mathrm{o}}\times(\ell-1)}$, $\mathbf{0}_{D2} \in \mathbb{C}^{L_{\mathrm{o}}\times(L_{\mathrm{o}}-L_{\mathrm{t}}-\ell)}$ and $\mathbf{e}_\ell \in \{0,1\}^{L_{h_k}}$. We then obtain

$$\begin{aligned}\mathrm{E}\Big[\mathbf{U}_k'^{\mathrm{H}}\mathbf{R}_{\eta,k}^{-1}\mathbf{U}'_k\Big] &= \mathrm{E}\left[\sum_{\lambda=1}^{L_h}\mathbf{e}_\lambda\mathbf{u}_k'^{\mathrm{H}}\mathbf{D}_\lambda^{\mathrm{T}}\mathbf{R}_{\eta,k}^{-1}\sum_{\ell=1}^{L_h}\mathbf{D}_\ell\mathbf{u}'_k\mathbf{e}_\ell^{\mathrm{T}}\right]\\ &= \sum_{\lambda=1}^{L_h}\mathbf{e}_\lambda\sum_{\ell=1}^{L_h}\mathrm{E}\left[\mathbf{u}_k'^{\mathrm{H}}\mathbf{D}_\lambda^{\mathrm{T}}\mathbf{R}_{\eta,k}^{-1}\mathbf{D}_\ell\mathbf{u}'_k\right]\mathbf{e}_\ell^{\mathrm{T}}\\ &= \sum_{\lambda=1}^{L_h}\mathbf{e}_\lambda\sum_{\ell=1}^{L_h}\mathrm{tr}\left(\mathbf{D}_\lambda^{\mathrm{T}}\mathbf{R}_{\eta,k}^{-1}\mathbf{D}_\ell\mathrm{E}\big[\mathbf{u}'_k\mathbf{u}_k'^{\mathrm{H}}\big]\right)\mathbf{e}_\ell^{\mathrm{T}}.\end{aligned}$$

Furthermore,

$$\mathrm{E}\big[\mathbf{u}'_k\mathbf{u}_k'^{\mathrm{H}}\big] = \mathbf{R}_{\epsilon,k,i} + \mathrm{E}[\mathbf{u}'_k]\mathrm{E}[\mathbf{u}'_k]^{\mathrm{H}} = \mathbf{R}_{\epsilon,k,i} + \hat{\mathbf{u}}'_{k,i}\hat{\mathbf{u}}_{k,i}'^{\mathrm{H}}, \quad (13)$$

where $\mathbf{R}_{\epsilon,k,i}$ is the covariance matrix of the estimation error of $\mathbf{u}'_k$ in the i-th iteration. Consequently, we obtain

$$\mathrm{E}\Big[\mathbf{U}_k'^{\mathrm{H}}\mathbf{R}_{\eta,k}^{-1}\mathbf{U}'_k\Big] = \boldsymbol{\Psi}_{k,i} + \mathrm{E}[\mathbf{U}'_k]^{\mathrm{H}}\mathbf{R}_{\eta,k}^{-1}\mathrm{E}[\mathbf{U}'_k], \quad (14)$$

with $[\boldsymbol{\Psi}_{k,i}]_{\ell,\lambda} = \mathrm{tr}(\mathbf{D}_\ell^{\mathrm{T}}\mathbf{R}_{\eta,k}^{-1}\mathbf{D}_\lambda\mathbf{R}_{\epsilon,k,i})$. Then it follows that

$$\begin{aligned}\mathrm{E}[\mathbf{U}^{\mathrm{H}}\mathbf{R}_\eta^{-1}\mathbf{U}] &= \sum_{k=0}^{M_{\mathrm{t}}-1}\mathbf{B}_k^{\mathrm{H}}(\boldsymbol{\Psi}_{k,i} + \mathrm{E}[\mathbf{U}'_k]^{\mathrm{H}}\mathbf{R}_{\eta,k}^{-1}\mathrm{E}[\mathbf{U}'_k])\mathbf{B}_k,\\ &= \boldsymbol{\Psi}_i + \hat{\mathbf{U}}_i^{\mathrm{H}}\mathbf{R}_\eta^{-1}\hat{\mathbf{U}}_i, \quad (15)\end{aligned}$$

where $\hat{\mathbf{U}}_i^{\mathrm{H}}\mathbf{R}_\eta^{-1}\hat{\mathbf{U}}_i = \sum_{k=0}^{M_{\mathrm{t}}-1}\mathbf{B}_k^{\mathrm{H}}\hat{\mathbf{U}}_{k,i}'^{\mathrm{H}}\mathbf{R}_{\eta,k}^{-1}\hat{\mathbf{U}}'_{k,i}\mathbf{B}_k$, and $\boldsymbol{\Psi}_i = \sum_{k=0}^{M_{\mathrm{t}}-1}\mathbf{B}_k^{\mathrm{H}}\boldsymbol{\Psi}_{k,i}\mathbf{B}_k$.

Finally, we define $\mathbf{S}_{\hat{\mathbf{u}},i} = \mathbf{S} + \hat{\mathbf{U}}_i$ and write

$$\mathrm{E}\left[\frac{\partial J(\mathbf{h}_i)}{\partial\mathbf{h}_i^{\mathrm{H}}}\bigg|\,\mathbf{y},\mathbf{h}_i\right] = \left(\mathbf{S}_{\hat{\mathbf{u}},i}^{\mathrm{H}}\mathbf{R}_\eta^{-1}\mathbf{S}_{\hat{\mathbf{u}},i} + \boldsymbol{\Psi}_i\right)\mathbf{h}_i - \mathbf{S}_{\hat{\mathbf{u}},i}^{\mathrm{H}}\mathbf{R}_\eta^{-1}\mathbf{y}.$$

Given the estimate of $\mathbf{h}$ and the training, one can estimate $\mathbf{u}'$. It can be shown that the entries of $\mathbf{u}'$ are alternating purely real and purely imaginary, because in addition to the OQAM data signals in subcarriers $k-1$ and $k+1$ also the impulse response from the two adjacent subcarriers have alternating real and imaginary coefficients. As a consequence, the interference term $\mathbf{u}'_k$ has improper statistics and the following Widely Linear MMSE [7] estimator can be employed

$$\mathrm{E}[\mathbf{u}'_i] = \hat{\mathbf{u}}'_i = \mathbf{W}_1\mathbf{y}_u + \mathbf{W}_2\mathbf{y}_u^*, \quad (16)$$

where $\mathbf{y}_u = \mathbf{y} - \mathbf{S}\hat{\mathbf{h}}_i \approx \hat{\mathbf{H}}_{\mathrm{E},i}\mathbf{u}' + \boldsymbol{\eta}$,

$$\mathbf{W}_1 = (\mathbf{R}_{uy_u} - \mathbf{P}_{uy_u}\mathbf{R}_{y_u}^{-*}\mathbf{P}_{y_u}^{*})(\mathbf{R}_{y_u} - \mathbf{P}_{y_u}\mathbf{R}_{y_u}^{-*}\mathbf{P}_{y_u}^{*})^{-1},$$
$$\mathbf{W}_2 = (\mathbf{P}_{uy_u} - \mathbf{R}_{uy_u}\mathbf{R}_{y_u}^{-1}\mathbf{P}_{y_u})(\mathbf{R}_{y_u}^{*} - \mathbf{P}_{y_u}^{*}\mathbf{R}_{y_u}^{-1}\mathbf{P}_{y_u})^{-1},$$

while $\mathbf{R}_{y_u} = \hat{\mathbf{H}}_{\mathrm{E},i}\mathbf{R}_u\hat{\mathbf{H}}_{\mathrm{E},i}^{\mathrm{H}} + \mathbf{R}_\eta$, and $\mathbf{R}_{uy_u} = \mathbf{R}_u\hat{\mathbf{H}}_{\mathrm{E},i}^{\mathrm{H}}$. The pseudo-covariance matrices are given by $\mathbf{P}_{y_u} = \hat{\mathbf{H}}_{\mathrm{E},i}\mathbf{P}_u\hat{\mathbf{H}}_{\mathrm{E},i}^{\mathrm{T}}$ and $\mathbf{P}_{uy_u} = \mathbf{P}_u\hat{\mathbf{H}}_{\mathrm{E},i}^{\mathrm{T}}$, where $\mathbf{R}_u = \mathrm{diag}(\mathbf{R}_{u,0}, ..., \mathbf{R}_{u,M_t-1})$, and correspondingly for $\mathbf{P}_u$. One can see that all covariance matrices are block diagonal and the estimation of $\mathbf{u}'$ in (16) is equivalent to estimating the $\mathbf{u}_k'$s subcarrier-wise.

The corresponding error covariance is given by

$$\mathbf{R}_{\epsilon,i} = \mathbf{R}_u - \mathbf{W}_1\mathbf{R}_{uy_u}^{\mathrm{H}} - \mathbf{W}_2\mathbf{P}_{uy_u}^{\mathrm{H}}, \quad (17)$$

where $\mathbf{R}_{\epsilon,i} = \mathrm{diag}(\mathbf{R}_{\epsilon,0,i}, \mathbf{R}_{\epsilon,1,i}, ..., \mathbf{R}_{\epsilon,M_t-1}, i)$.

Finally, the M-step is performed, where $J(\mathbf{h}_i)$ is minimized, resulting in the new channel estimate

$$\hat{\mathbf{h}}_{\mathrm{DS},i+1} = (\mathbf{A}^{\mathrm{H}}(\mathbf{S}_{\hat{\mathbf{u}},i}^{\mathrm{H}}\mathbf{R}_\eta^{-1}\mathbf{S}_{\hat{\mathbf{u}},i} + \boldsymbol{\Psi}_i)\mathbf{A})^{-1}\mathbf{A}^{\mathrm{H}}\mathbf{S}_{\hat{\mathbf{u}},i}^{\mathrm{H}}\mathbf{R}_\eta^{-1}\mathbf{y}.$$

The estimation of $\mathbf{u}_i$ and $\mathbf{h}_i = \mathbf{A}\mathbf{h}_{\mathrm{DS},i}$ is then repeated N_{EM} times until convergence is achieved.

4. SIMULATION RESULTS

For the performance simulations, the parameters were $M = 256$, $M_{\mathrm{u}} = 210$, $K = 4$ and an RRC prototype with roll-off one. The total signal bandwidth is 12.6 MHz and the sampling rate is $M/T = 15.36$ MHz, giving a subcarrier bandwidth of 60 kHz and a symbol duration of $T = 16.67$ μs. The channel model was the ITU-Vehicular A without mobility. The CIR duration is $L_h = 36$ samples.

The observations were taken from every 4-th subcarrier. Then we can consider two cases: for known $\mathbf{u}$ 158 subcarriers are filled with training and the rest with data, while for unknown $\mathbf{u}$ only 53 subcarriers are filled with training and the rest with data. We have used random QPSK training symbols. The normalized MSE (NMSE) of the channel estimation was averaged over 100 channel realizations, each was also averaged over 10 training sequences and, for each training, averaged over 10 noise realizations. In Fig. 5, the theoretical and numerical NMSEs for the different estimators are depicted as a function of E_{s}/N_0. Further parameters are $L_{\mathrm{o}} = 4$ and for the ML-EM algorithm $L_{h_k} = 3$. The theoretical NMSE and the curve where $\mathbf{u}$ is known show a lower bound on the NMSE performance. The curve for $N_{\mathrm{EM}} = 0$ shows an upper bound, since there the interference from data carrying subcarriers degrades the estimation performance. The ML-EM curves show the performance for $N_{\mathrm{EM}} = 2, 5, 7$ and 10 iterations. One can see that for 10 and 20 iterations the performance is the same, showing that no more than 10 iterations are necessary. A great improvement compared to 0 iterations is achieved and a low level of MSE is achieved.

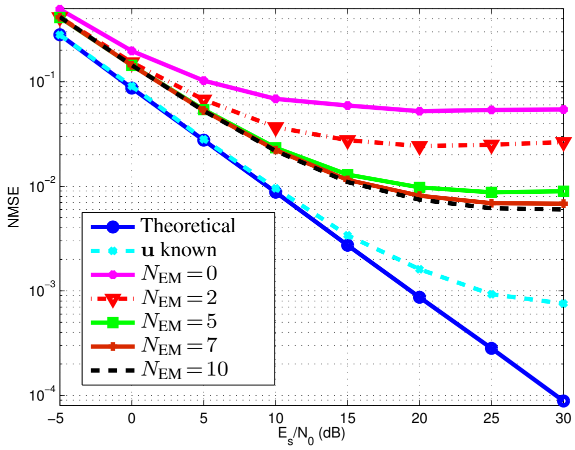

Fig. 5. NMSE as a function of E_{s}/N_0.

5. CONCLUSIONS

We have presented a novel method for the estimation of frequency selective channels in FBMC systems. It is an extension of an earlier method we have developed based on the EM algorithm, but this time for the broadband channel estimation. A three times improvement in the spectral efficiency can be achieved compared to the best possible ML estimator considering the same system model. We could show a great NMSE improvement for the same training burden.

REFERENCES

[1] B. R. Saltzberg, "Performance of an efficient parallel data transmission system," *IEEE Trans. Comm. Technology*, vol. COM-15, no. 6, pp. 805–811, Dec. 1967.

[2] D. S. Waldhauser, L. G. Baltar, and J. A. Nossek, "MMSE subcarrier equalization for filter bank based multicarrier systems," in *Proc. IEEE 9th Workshop Signal Proc. Advances in Wireless Comm. SPAWC 2008*.

[3] L. G. Baltar, D. S. Waldhauser, and J. A. Nossek, "MMSE subchannel decision feedback equalization for filter bank based multicarrier systems," in *Proc. IEEE Int. Symp. Circuits and Systems ISCAS 2009*.

[4] L. G. Baltar, A. Mezghani, and J. A. Nossek, "EM based per-subcarrier ML channel estimation for filter bank multicarrier systems," in *Proc. of the 10-th Int. Symposium on Wireless Comm. Systems ISWCS 2013*.

[5] T. Karp and N.J. Fliege, "Modified DFT filter banks with perfect reconstruction," *IEEE Trans. on Circuits and Systems II: Analog and Digital Signal Proc.*, vol. 46, no. 11, pp. 1404–1414, Nov. 1999.

[6] T.K. Moon, "The expectation-maximization algorithm," *IEEE SP Magazine*, vol. 13, no. 6, pp. 47–60, 1996.

[7] P. J. Schreier and L. L. Scharf, *Statistical Signal Processing of Complex-Valued Data*, Cambridge University Press, 2010.