MMSE Training Design for Filter Bank Multicarrier Systems with Per-Subcarrier Channel Estimation

Michael Newinger, Leonardo G. Baltar, Josef A. Nossek
Institute for Circuit Theory and Signal Processing
Technische Universität München
80290 Munich, Germany
Email: michael.newinger@gmx.de, leo.baltar@tum.de, josef.a.nossek@tum.de

***Abstract*—Filter bank based multicarrier (FBMC) systems present an alternative solution to cyclic prefix based orthogonal frequency division multiplexing (CP-OFDM) in wireless environments with multipath propagation. In this contribution we propose a novel training sequences design method to be employed in a per-subcarrier maximum likelihood (ML) narrowband channel estimation scheme, recently developed by the authors. In the training design we take into account the knowledge of transmitter and receiver side prototype filters to calculate the pilot sequences. The method targets at the minimum mean squared error (MMSE) of the channel estimation and uses a gradient projection algorithm to find the optimum sequences. In the numerical results we show a gain of up to 4 dB for the same MMSE value in a low to medium E_b/N_0 regime.**

I. Introduction

In this contribution we consider FBMC systems in wireless environments with multipath propagation. In contrast to CP-OFDM, where a rectangular pulse shaping is used, here a finite impulse response (FIR) prototype filter with a longer impulse response than the symbol period, i. e. the filter length is bigger than the total number of subcarriers M, is modulated by complex exponentials. As a consequence, we can achieve more spectrally concentrated subcarriers that overlap only with their neighbors. The prototype filter can be designed according to different goals, although here we concentrate our results on a finite length Root Raised Cosine (RRC) filter with unity roll-off. But to achieve subcarrier orthogonality, i. e. inter-symbol interference (ISI) and inter-channel interference (ICI)-free received symbols the real and imaginary parts of the input symbols must be staggered by $T/2$, resulting in the so called Offset-QAM (OQAM) modulation.

In [1] and [2] the authors have presented methods to analytically design per-subcarrier fractionally spaced equalizers that compensate ISI and ICI inserted by the multipath channel, but they have assumed perfect channel impulse response knowledge at the receiver. In [3] and [4] we have presented methods for the estimation of the narrowband multipath channel viewed in each subcarrier. In this contribution we extend those results by showing how to design training sequences that minimize the mean squared error (MMSE) of the channel estimation. For the optimization we employ a gradient projection algorithm.

II. FBMC System and Subcarrier Model

A high level model of the FBMC system is shown in Fig. 1. This filter bank configuration is known as transmultiplex in the signal processing literature [5]. A synthesis filter bank

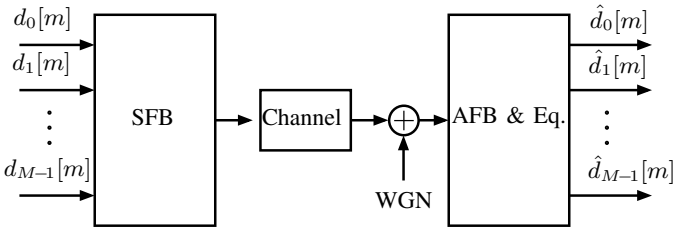

Figure 1. FBMC System Overview.

(SFB) performs a frequency division multiplexing of the QAM data symbols $d_k[m]$ on parallel subcarriers in a rate of $1/T$ at the transmitter. An analysis filter bank (AFB) at the receiver separates the data on each subcarrier. We assume a static frequency selective channel and AWGN between SFB and AFB. Usually some of the M subcarriers are left empty to limit the transmit signal spectrum.

Here we regard exponentially modulated SFB and AFB. This means that only one prototype low-pass filter has to be designed and the other sub-filters are obtained by modulating it as follows [6]

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

where $h_{\mathrm{P}}[l]$ is the impulse response of the prototype filter with length L_{P}. The prototype is an approximation of a Nyquist filter with roll-off factor one and as a consequence only the spectrum of contiguous subcarriers overlap and non-contiguous subcarriers are separated by a high stop-band attenuation. For example, an FIR RRC with length $L_{\mathrm{P}} = KM+1$ can be used, where K is the overlapping factor that determines how many symbols superpose each other in time. K should be kept as small as possible to limit the complexity and to reduce the time-domain spreading of the symbols, in particular in case time-variant propagation channels are considered.

Because the prototype filter is longer than the number of subcarriers ($L_{\mathrm{P}} > M$), to maintain the orthogonality between all of them and for all time instants, the complex QAM input symbols $d_k[m]$ need to have their real and complex parts staggered by $T/2$ resulting in the so-called OQAM modulation [7]. The OQAM staggering for even indexed subcarriers is illustrated in Fig. 2. For odd indexed subcarriers the delay of $T/2$ is located at the lower branch with purely imaginary numbers. The OQAM de-staggering is performed at the receiver by the application of flow-graph reversal [8] in Fig. 2,

978-1-4673-6187-3/13/$31.00 ©2013 IEEE

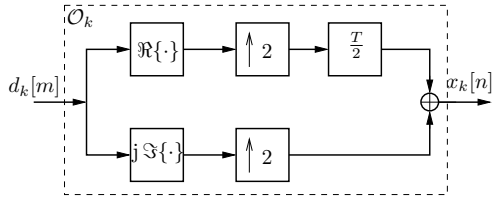

Figure 2. O-QAM staggering for odd indexed subcarrier

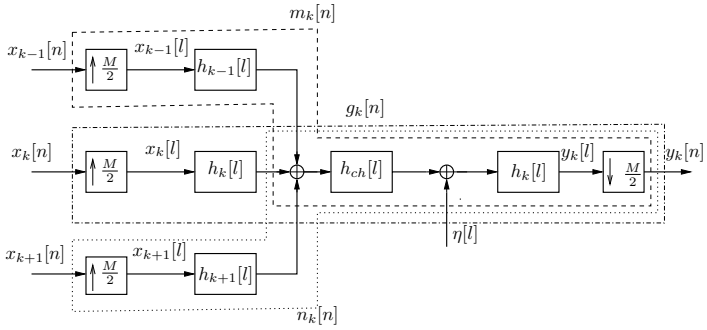

Figure 3. Subcarrier output model

substitution of up-samplers by down-samplers and exchange of blocks $\Re\{\cdot\}$ and $\mathrm{j}\,\Im\{\cdot\}$.

The signals of all subcarriers are up-sampled by $M/2$ after the OQAM staggering, filtered and added. A broadband signal is then generated and digital-to-analog (DA) converted into two IQ baseband signals that will be analog processed (filtered, up-converted to IF and RF, amplified, etc.) and transmitted. At the receiver side the RF signal is amplified, brought to baseband, filtered and then analog-to-digital (AD) converted. The digital received signal is then filtered by the analysis filters and down-sampled by $M/2$ to generate the OQAM staggered subcarrier signals.

The fact we have assumed only contiguous subcarriers overlap in the frequency domain, allows us to construct a model for the output of one subcarrier as 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 resulting QAM symbols. As a consequence, in this subcarrier model the input and output sampling rates are $2/T$. We have assumed here a multipath channel but with perfect time and frequency synchronization. In other words, no time or frequency shifts (Carrier frequency offset or Doppler shift or spread) are present. A more realistic model should involve this and other issues that are out of the scope of this contribution.

Now we can collect the subcarrier output symbols $y_k[n]$ into an observation vector $\mathbf{y}_k[n] \in {}^{L_\mathrm{o}}$. Then we can express it as a function of the k, (k–1) and (k+1) subcarrier input signals convolved with three impulse responses. They are a result of the convolution between transmit and receive filters, and the propagation channel with further down-sampling by $M/2$ (ref. Fig. 3). Of course a thermal noise term is also included. This leads us to the definition

$$\mathbf{y}_k[n] = \mathbf{G}_k\mathbf{x}_k[n] + \mathbf{M}_k\mathbf{x}_{k-1}[n] + \mathbf{N}_k\mathbf{x}_{k+1}[n] + \mathbf{\Gamma}_k\boldsymbol{\eta}[l], \quad (1)$$

where $\mathbf{x}_{k(\pm1)}[n] \in {}^{L_\mathrm{o}+L_\mathrm{g}-1}$ and we have employed the convolution (Toeplitz) matrices $\mathbf{G}_k$, $\mathbf{M}_k$, $\mathbf{N}_k \in {}^{L_\mathrm{o}\times(L_\mathrm{o}+L_\mathrm{g}-1)}$, composed by the L_g long impulse responses $g_k[n]$, $m_k[n]$ and $n_k[n]$, with $L_\mathrm{g} = \left\lceil \frac{2(L_\mathrm{P}-1)+L_\mathrm{ch}}{M/2} \right\rceil$ and L_ch is the impulse response length of the multipath channel. $\mathbf{\Gamma}_k \in {}^{L_\mathrm{o}\times(L_\mathrm{P}+\frac{M}{2}L_\mathrm{o})}$ is obtained by taking each $\frac{M}{2}$-th row of the convolution matrix composed by the analysis filter impulse response $h_k[l]$. This is the reason why the vector $\boldsymbol{\eta}[l] \in {}^{(L_\mathrm{P}+\frac{M}{2}L_\mathrm{o})}$ is defined in the high sampling rate M/T. The subchannel model in (1) and in Fig. 3 was first proposed in [1].

The multipath channel will introduce ISI and ICI in the received subcarrier symbols. To compensate for that a fractionally spaced equalizer has to be introduced at each subcarrier before the OQAM de-staggering. A per-subcarrier linear MMSE or DFE MMSE equalizer can be employed for this matter as in [1], [2]. But before the equalizers can be designed an estimation of the multipath channel impulse response needs to be available. Instead of estimating the broadband channel $h_{ch}[l]$, we will show in the next section a method of estimating many narrowband channels that are sufficient for the equalizer design.

We should mention here that there are many efficient structures in the literature for the subcarrier filters realization. For this matter, the modulation of the prototype filter is performed with the aid of a fast Fourier transform algorithm. In order to perform the filtering at a lower sampling rate, the polyphase decomposition [5] of the prototype filter is also employed. One example of efficient realization can be found in [6].

III. ML Structured Channel Estimation

In this contribution we assume that a per-subcarrier linear or DFE equalizer is employed and as a consequence a per-subcarrier estimator is sufficient for the equalizers design. We further assume that training sequences are not only employed in the subcarriers of interest, but also in their contiguous subcarriers. In the next section we show a method how to design those sequences appropriately in order to improve the estimation quality or reduce the training overhead.

To perform the structured channel estimation we first have to modify our subcarrier model. The idea is to model the multipath channel viewed at each subcarrier as a narrowband channel with a short impulse response and represent it in a lower sampling rate, namely the double of the QAM symbol rate $2/T$. For this, we first convolve the transmit filters $h_{k(\pm1)}[l]$ with the receive filter $h_k[l]$. Then we down-sample by $M/2$ the resulting impulse responses and obtain $h'_{k,k(\pm1)}[n]$. As a consequence the narrowband propagation channel experienced by subcarrier k has impulse response $c_k[n]$. The three new overall impulse responses for the contiguous subcarriers are now $g'_k[n]$, $m'_k[n]$ and $n'_k[n]$ and have length L'_g. This new subcarrier model is illustrated in Fig. 4. It is important to note here the difference to Fig. 3. In that model the impulse responses $g_k[n]$, $m_k[n]$ and $n_k[n]$ were obtained after the convolution between transmit filters, broadband propagation channel and receive filter, and then down-sampling the resulting impulse responses by $M/2$. Furthermore, it can be shown that for most multipath channels $L'_\mathrm{g} > L_\mathrm{g}$ is needed for an acceptable performance.

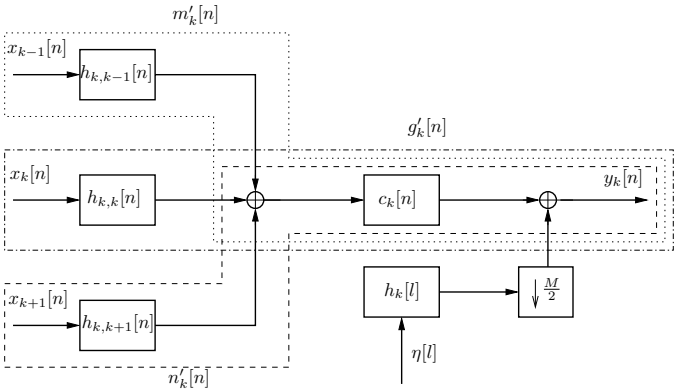

Figure 4. Modified subchannel model for structured approach

We can now express the subcarrier received signal as

$$\mathbf{y}_k[n] = \mathbf{G}'_k\mathbf{x}_k[n] + \mathbf{M}'_k\mathbf{x}_{k-1}[n] + \mathbf{N}'_k\mathbf{x}_{k+1}[n] + \mathbf{\Gamma}_k\boldsymbol{\eta}[l], \quad (2)$$

where now $\mathbf{x}_{k(\pm1)}[n] \in {}^{L_\mathrm{t}}$, the convolution matrices $\mathbf{G}'_k$, $\mathbf{M}'_k, \mathbf{N}'_k \in {}^{L_\mathrm{o}\times L_\mathrm{t}}$ contain the impulse responses $g'_k[n]$, $m'_k[n]$ and $n'_k[n]$, with $L_\mathrm{t} = L_\mathrm{o} + L'_\mathrm{g} - 1$, $L'_\mathrm{g} = L_{\mathrm{h}_k} + L_{\mathrm{c}_k} - 1$ and $L_{\mathrm{h}_k} = \left\lceil \frac{2(L_\mathrm{P}-1)}{M/2} \right\rceil$.

If we exchange the role of the matrices and vectors in equation (2), we get

$$\mathbf{y}_k[n] = \mathbf{X}_k[n]\mathbf{g}'_k + \mathbf{X}_{k-1}[n]\mathbf{m}'_k + \mathbf{X}_{k+1}[n]\mathbf{n}'_k + \boldsymbol{\nu}_k[n], \quad (3)$$

where $\mathbf{g}'_k$, $\mathbf{m}'_k$ and $\mathbf{n}'_k \in {}^{L'_\mathrm{g}}$ are vectors containing the impulse responses $g'_k[n]$, $m'_k[n]$ and $n'_k[n]$, plus $\mathbf{X}_{k(\pm1)}[n] \in {}^{L_\mathrm{o}\times L'_\mathrm{g}}$ are Hankel matrices obtained from the training sequences $x_{k(\pm1)}[n]$. Moreover, by assuming the training sequences have length $L_\mathrm{t} = L_\mathrm{o} + L'_\mathrm{g} - 1$, the observations contained in $\mathbf{y}_k[n]$ are free from time-overlap with subsequent data symbols. This requisite could be relaxed but some loss in performance would be expected. However, if one wants to reduce the training overhead and increase the spectral efficiency, this time-overlap should be considered in the channel estimation.

It is clear that the vectors $\mathbf{g}'_k$, $\mathbf{m}'_k$ and $\mathbf{n}'_k$ depend not only on the narrowband multipath channel, but also on the known transmit and receive filters. If we separate both impulse responses we can write $\mathbf{g}'_k = \mathbf{H}_{k,k}\mathbf{c}_k$, $\mathbf{m}'_k = \mathbf{H}_{k,k-1}\mathbf{c}_k$ and $\mathbf{n}'_k = \mathbf{H}_{k,k+1}\mathbf{c}_k$, where $\mathbf{H}_{k,k(\pm1)} \in {}^{L'_\mathrm{g}\times L_{\mathrm{c}_k}}$ are Toeplitz matrices containing the impulse responses $h_{k,k(\pm1)}[n]$, and $\mathbf{c}_k \in {}^{L_{\mathrm{c}_k}}$ is a vector containing the impulse response $c_k[n]$. Now the received signals are given by

$$\begin{aligned}\mathbf{y}_k[n] =& (\mathbf{X}_k[n]\mathbf{H}_{k,k} + \mathbf{X}_{k-1}[n]\mathbf{H}_{k,k-1} \\ &+ \mathbf{X}'_{k+1}[n]\mathbf{H}_{k,k+1})\mathbf{c}_k + \mathbf{\Gamma}_k\boldsymbol{\eta}[l] \\ =& \mathbf{S}_k\mathbf{c}_k + \boldsymbol{\nu}_k[n],\end{aligned} \quad (4)$$

where we define

$$\mathbf{S}_k = [\mathbf{X}_{k-1} \quad \mathbf{X}_k \quad \mathbf{X}_{k+1}] \begin{bmatrix} \mathbf{H}_{k,k-1} \\ \mathbf{H}_{k,k} \\ \mathbf{H}_{k,k+1} \end{bmatrix} = \breve{\mathbf{X}}_k\breve{\mathbf{H}}_k, \quad (5)$$

with $\breve{\mathbf{X}}_k \in \mathbb{C}^{L_\mathrm{o}\times 3L'_\mathrm{g}}$ and $\breve{\mathbf{H}}_k \in \mathbb{C}^{3L'_\mathrm{g}\times L_{\mathrm{c}_k}}$.

We can see that in the linear model of (4) the noise is zero mean Gaussian distributed with autocorrelation matrix $\mathbf{R}_{\nu_k\nu_k} = \sigma_\eta^2\mathbf{\Gamma}_k\mathbf{\Gamma}_k^\mathrm{H}$ and the observation $\mathbf{y}_k[n]$ given $\mathbf{c}_k$ is Gaussian distributed. The maximum likelihood (ML) estimate of $\mathbf{c}_k$ in this case is given by

$$\hat{\mathbf{c}}_{k,\mathrm{ML}} = \arg\max_{\mathbf{c}_k} p(\mathbf{y}_k[n]|\mathbf{c}_k). \quad (6)$$

Since $\mathbf{R}_{\nu_k\nu_k}$ is independent of $\mathbf{c}_k$ and $(\mathbf{S}_k^\mathrm{H}\mathbf{R}_{\eta,k}^{-1}\mathbf{S}_k)$ is invertible, the ML estimate of the narrowband multipath channel in each subcarrier is given by

$$\begin{aligned}\hat{\mathbf{c}}_k &= (\mathbf{S}_k^\mathrm{H}\mathbf{R}_{\nu_k\nu_k}^{-1}\mathbf{S}_k)^{-1}\mathbf{S}_k^\mathrm{H}\mathbf{R}_{\nu_k\nu_k}^{-1}\mathbf{y}_k[n] \\ &= (\mathbf{S}_k^\mathrm{H}(\mathbf{\Gamma}_k\mathbf{\Gamma}_k^\mathrm{H})^{-1}\mathbf{S}_k)^{-1}\mathbf{S}_k^\mathrm{H}(\mathbf{\Gamma}_k\mathbf{\Gamma}_k^\mathrm{H})^{-1}\mathbf{y}_k[n].\end{aligned} \quad (7)$$

We can note that L_{c_k} is a design parameter of the channel estimator. It can be different for different subcarriers depending on how frequency selective the channel is for a certain portion of the spectrum. Structured channel estimation is the name some authors use when known pulse shaping filters are located at transmitter and receiver side, like here, and this knowledge is used by the estimation algorithm [9].

IV. MMSE Training Sequence Design

Now that we have a channel estimation procedure, it remains the problem of finding optimal sequences to improve the quality of the estimation. First we have to write the Hankel matrix as a function of the vector that generates it as

$$\mathbf{X}_k = \sum_{j=1}^{L'_\mathrm{g}} \mathbf{D}_j\mathbf{O}_k\mathbf{x}_k^\mathrm{R}[n]\mathbf{e}_j^\mathrm{T} \quad (8)$$

where $\mathbf{x}_k^\mathrm{R}[n]$ is a purely real version of $\mathbf{x}_k[n]$, $\mathbf{D}_j = [\underbrace{\mathbf{0}}_{j-1}\mathbf{I}_{L_\mathrm{o}}\underbrace{\mathbf{0}}_{L'_\mathrm{g}-j}] \in \{0,1\}^{L_\mathrm{o}\times L_\mathrm{t}}$ selects the part of $\mathbf{x}_k^\mathrm{R}[n]$ that appear in each column of $\mathbf{X}_k$ and $\mathbf{O}_k = \mathrm{diag}([...,1,\mathrm{j},1,\mathrm{j},...])$ represents the alternation between purely real and purely imaginary symbols of the OQAM staggering. $\mathbf{e}_j$ is a unit vector of length L'_g with a 1 at position j.

We furthermore define a new purely real vector with training symbols of the three subcarriers

$$\breve{\mathbf{x}}_k = \begin{bmatrix}\mathbf{x}_{k-1}^{\mathrm{R,T}}[n] & \mathbf{x}_k^{\mathrm{R,T}}[n] & \mathbf{x}_{k+1}^{\mathrm{R,T}}[n]\end{bmatrix}^\mathrm{T} \in \mathbb{C}^{3L_\mathrm{t}} \quad (9)$$

and rewrite $\mathbf{S}_k$ as

$$\mathbf{S}_k = \sum_{j=1}^{3L'_\mathrm{g}} \breve{\mathbf{D}}_j\breve{\mathbf{O}}_k\breve{\mathbf{x}}_k\breve{\mathbf{e}}_j^\mathrm{T}\breve{\mathbf{H}}_k, \quad (10)$$

where now $\breve{\mathbf{e}}_j$ is a unit vector of length $3L'_\mathrm{g}$. $\breve{\mathbf{D}}_j \in \{0,1\}^{L_\mathrm{o}\times 3L_\mathrm{t}}$, however, has to be defined very carefully to select the correct part of $\breve{\mathbf{x}}_k$, namely

$$\breve{\mathbf{D}}_j = \begin{cases} [\underbrace{\mathbf{0}}_{j-1}\overbrace{\mathbf{I}_{L_\mathrm{o}}\underbrace{\mathbf{0}}_{L'_\mathrm{g}-j}}^{L_\mathrm{t}}\overbrace{\mathbf{0}}^{L_\mathrm{t}}\overbrace{\mathbf{0}}^{L_\mathrm{t}}], & 1 \le j \le L'_\mathrm{g}, \\ [\overbrace{\mathbf{0}}^{L_\mathrm{t}}\overbrace{\underbrace{\mathbf{0}}_{j-L'_\mathrm{g}-1}\mathbf{I}_{L_\mathrm{o}}\underbrace{\mathbf{0}}_{2L'_\mathrm{g}-j}}^{L_\mathrm{t}}\overbrace{\mathbf{0}}^{L_\mathrm{t}}], & L'_\mathrm{g}+1 \le j \le 2L'_\mathrm{g}, \\ [\overbrace{\mathbf{0}}^{L_\mathrm{t}}\overbrace{\mathbf{0}}^{L_\mathrm{t}}\overbrace{\underbrace{\mathbf{0}}_{j-2L'_\mathrm{g}-1}\mathbf{I}_{L_\mathrm{o}}\underbrace{\mathbf{0}}_{3L'_\mathrm{g}-j}}^{L_\mathrm{t}}], & 2L'_\mathrm{g}+1 \le j \le 3L'_\mathrm{g}, \end{cases}$$

and

$$\breve{\mathbf{O}}_k = \begin{bmatrix} \mathbf{O}_{k-1} & & \\ & \mathbf{O}_k & \\ & & \mathbf{O}_{k+1} \end{bmatrix}, \tag{11}$$

where $\mathbf{O}_{k-1} = \mathbf{O}_{k+1} = \text{diag}([..., \text{j}, 1, \text{j}, 1, ...])$.

With $\mathbf{S}_k$ described as a function of $\breve{\mathbf{x}}_k$, also $\breve{\mathbf{c}}_k$ can be expressed as a function of $\breve{\mathbf{x}}_k$ and we can finally formulate our optimization problem. One possible reliability measure is the mean squared error (MSE) of the narrowband channel estimation. We than state the minimization of the MSE (MMSE) as follows

$$\begin{aligned} \breve{\mathbf{x}}_k &= \arg\min_{\breve{\mathbf{x}}_k} \text{E}\{\|\mathbf{c}_k - \hat{\mathbf{c}}_k\|^2\} \\ &\text{s. t. } ||\mathbf{x}_k||^2 \leq E_\text{x},\ ||\mathbf{x}_{k\pm1}||^2 \leq \frac{E_\text{x}}{2}, \end{aligned} \tag{12}$$

where additionally to the constraint of real input sequences we have introduced the energy constraint with $E_\text{x} \in \mathbb{R}_+$. It is worth noting that the maximum energy allocated to the training sequences of the contiguous subcarriers is half of the energy allocated to the subcarrier of interest. In this way only every second subcarrier needs to have its inputs optimized and the training sequences of the subcarriers between them is obtained by adding the two resulting training sequences. At the end the training sequences in all subcarriers will obey to the same power constraint.

To further develop the objective function we first plug (4) into (7), secondly we plug (7) into the objective function, third perform some algebraic manipulation, fourth apply the property $\|\mathbf{a}\|^2 = \mathbf{a}^\text{H}\mathbf{a} = \text{tr}\{\mathbf{a}\mathbf{a}^\text{H}\}$, fifth exchange the trace with expectation operator and at the end perform more algebraic manipulation. Finally, we can find out that the objective function is given by

$$J(\breve{\mathbf{x}}_k) = \text{E}\{\|\mathbf{c}_k - \hat{\mathbf{c}}_k\|^2\} = \text{tr}\{(\mathbf{S}_k^\text{H}\mathbf{R}_{\nu_k\nu_k}^{-1}\mathbf{S}_k)^{-1}\}. \tag{13}$$

The optimization problem now looks as follows

$$\begin{aligned} \breve{\mathbf{x}}_k &= \arg\min_{\breve{\mathbf{x}}_k} \text{tr}\{(\mathbf{S}_k^\text{H}\mathbf{R}_{\nu_k\nu_k}^{-1}\mathbf{S}_k)^{-1}\} \\ &\text{s. t. } ||\mathbf{x}_k||^2 \leq E_\text{x},\ ||\mathbf{x}_{k\pm1}||^2 \leq \frac{E_\text{x}}{2}, \end{aligned} \tag{14}$$

where $\mathbf{S}_k = \sum_{j=1}^{3L_\text{g}'} \breve{\mathbf{D}}_j \breve{\mathbf{O}}_k \breve{\mathbf{x}}_k \mathbf{e}_j^\text{T} \breve{\mathbf{H}}_k$.

It can be shown that this objective function is not convex and this means that there is no global optimum. As a consequence there is no closed form solution for the optimal training sequence. Intuitively it is to expect that there is not only one, but infinitely many training sequences that provide the same MMSE of the channel estimation. Not so many of them also fulfill the energy constraints. Also for different starting points different local minima are reached, but we could observe that similar performance is achieved. In order to find a local optimum we propose to employ a gradient projection algorithm, because we have a constrained optimization problem [10]. This algorithm is like a gradient descent where in each iteration step a projection onto the feasible set is performed. In Fig. 5 we show the gradient projection algorithm that allows us to find optimum training sequences.

After the initialization of many variables and constants, the gradient of the objective function is given in line 11. The

1: **Initialization:**
2: Set $0 < \rho < 1$ and $0 < \beta < 1$
3: Set E_x
4: Let $\mathcal{K}$ be the set of all used subcarriers k
5: Initialize all sequences $\mathbf{x}_k$ with $k \in \mathcal{K}$ with L_t zeros
6: **for** $k = 2 : 2 : |\mathcal{K}|$ **do**
7: $\breve{\mathbf{x}}_k(0) \Leftarrow$ random vector inside the feasible set
8: $i \Leftarrow 0$
9: **repeat**
10: $\mathbf{S}_k(i) \Leftarrow \sum_{j=1}^{3L_\text{g}'} \breve{\mathbf{D}}_j \breve{\mathbf{O}}_k \breve{\mathbf{x}}_k(i) \mathbf{e}_j^\text{T} \breve{\mathbf{H}}_k$
11: $\nabla J(\breve{\mathbf{x}}_k(i)) \Leftarrow 2\sum_{j=1}^{3L_\text{g}'} \left(-\mathbf{e}_j^\text{T}\breve{\mathbf{H}}_k(\mathbf{S}_k^\text{H}(i)\mathbf{R}_{\nu_k\nu_k}^{-1}\mathbf{S}_k(i))^{-2}\mathbf{S}_k^\text{H}(i)\mathbf{R}_{\nu_k\nu_k}^{-1}\breve{\mathbf{D}}_j\breve{\mathbf{O}}_k\right)^\text{T}$
12: $\ell \Leftarrow 0$
13: **while true do**
14: $\alpha \Leftarrow \beta^\ell$
15: $\breve{\mathbf{x}}_k(i+1) \Leftarrow \Re\{\breve{\mathbf{x}}_k(i) - \alpha\nabla J(\breve{\mathbf{x}}_k(i))\}$
16: **if** $\|\mathbf{x}_{k-1}(i+1)\|^2 > \frac{E_\text{x}}{2}$ **then**
17: $\mathbf{x}_{k-1}(i+1) \Leftarrow \frac{\mathbf{x}_{k-1}(i+1)}{\|\mathbf{x}_{k-1}(i+1)\|}\sqrt{\frac{E_\text{x}}{2}}$
18: **end if**
19: **if** $\|\mathbf{x}_k(i+1)\|^2 > E_\text{x}$ **then**
20: $\mathbf{x}_k(i+1) \Leftarrow \frac{\mathbf{x}_k(i+1)}{\|\mathbf{x}_k(i+1)\|}\sqrt{E_\text{x}}$
21: **end if**
22: **if** $\|\mathbf{x}_{k+1}(i+1)\|^2 > \frac{E_\text{x}}{2}$ **then**
23: $\mathbf{x}_{k+1}(i+1) \Leftarrow \frac{\mathbf{x}_{k+1}(i+1)}{\|\mathbf{x}_{k+1}(i+1)\|}\sqrt{\frac{E_\text{x}}{2}}$
24: **end if**
25: $\breve{\mathbf{x}}_k(i+1) \Leftarrow \left[\mathbf{x}_{k-1}^\text{T}(i+1), \mathbf{x}_k^\text{T}(i+1), \mathbf{x}_{k+1}^\text{T}(i+1)\right]^\text{T}$
26: **if** $J(\breve{\mathbf{x}}_k(i)) - J(\breve{\mathbf{x}}_k(i+1)) \geq \rho\frac{\|\breve{\mathbf{x}}_k(i) - \breve{\mathbf{x}}_k(i+1)\|^2}{\beta^\ell}$ **then**
27: **break**
28: **else**
29: $\ell \Leftarrow \ell + 1$
30: **end if**
31: **end while**
32: $i \Leftarrow i + 1$
33: **until** Stopping criterion is met
34: $\mathbf{x}_{k-1} \Leftarrow \mathbf{x}_{k-1} + \mathbf{x}_{k-1}(i)$
35: $\mathbf{x}_k \Leftarrow \mathbf{x}_k(i)$
36: $\mathbf{x}_{k+1} \Leftarrow \mathbf{x}_{k+1} + \mathbf{x}_{k+1}(i)$
37: **end for**

Figure 5. Gradient Projection Algorithm for Training Sequences Design

derivation of this gradient is long and is not presented here. In our case we wish to have a purely real solution, so the projection in this case is obtained by taking the real part of the solution in each iteration step (see line 15). To fulfill the energy constraints, we limit the lengths of the solution vectors. If they do not fulfill them even after the projection onto the real field, we project them onto spheres with radius equal to the maximum lengths (see lines 17, 20 and 23). For the choice of the step size we have employed a generalized Amijo rule [11]. This rule can be seen in the internal loop between lines 13 and 31, where the step-size is reduced until the condition in line 26 is fulfilled. This condition guarantees that the objective function is always reduced.

V. Numerical results

To evaluate the performance of the narrowband channel estimation and the optimum design of the training sequences we have performed some numerical simulations. As broadband multipath channel model we have chosen static ITU vehicular A at a bandwidth of 10 MHz and sampling rate $M/T = 11.2$ MHz. The FBMC system was employed with $M = 256$ subcarriers from which 128 were occupied with training symbols. For the prototype an RRC filter with $K = 4$ and consequently $L_\text{P} = 1025$ was chosen. With this configuration and scenario a length of $L_{\text{c}_k} = 5$ for the narrowband multipath channels has been shown to be sufficient for all subcarriers. In the gradient

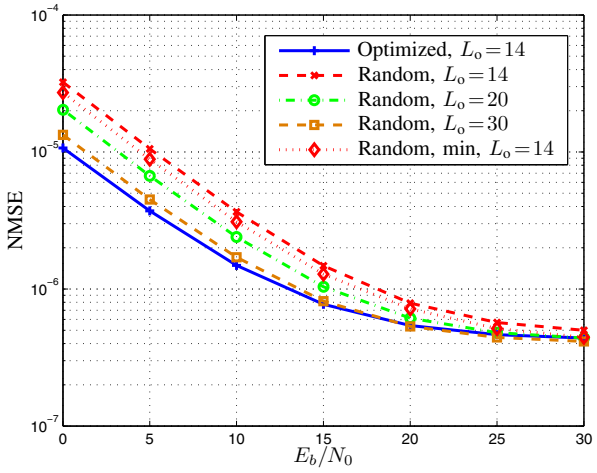

Figure 6. NMSE of the narrowband channel estimation

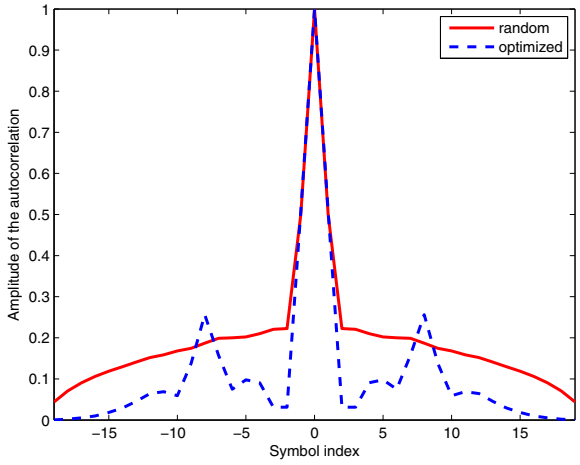

Figure 7. Autocorrelation of the narrowband channel input

projection algorithm we have used $\rho = 10^{-2}$ and $\beta = 0.5$.

In Fig. 6 the normalized MSE (NMSE) of the narrowband channel estimation is depicted for different number of observations and for both optimized training sequences and the ones generated by randomly chosen QPSK symbols. For each of the curves 250 channel realizations were averaged. In the case of the randomly generated, 50 sequences were generated for each channel realization and then averaged. We further show the best performance achieved by a random training with $L_o = 14$ and different sequences were used in different subcarriers. We can see from the results that for $L_o = 14$ a gain of up to 4 dB in low to medium E_b/N_0 regime for the same NMSE value is achieved when the optimized training is employed instead of random sequences. Alternatively, we can see that if we increase the number of observations to $L_o = 30$ and use random generated training sequences, we still do not achieve the same NMSE as the optimized sequences for $L_o = 14$. This shows that some training overhead can be saved with the optimal sequences, as we expected.

In addition we have depicted in Fig. 7 the autocorrelation function of the sequences at the input of the narrowband multipath channel in each subcarrier. It is clear that the autocorrelation for the optimized training sequences presents better time concentration than the obtained with random sequences. The side-lobes observed there have to do with the correlated noise that is added after the narrowband channel and that is also taken into account in the design of the sequences.

VI. Conclusion

In this contribution we have presented a method to design optimum training sequences for the narrowband channel estimation in FBMC systems. The method is based on the MMSE criterion and we have employed a gradient projection algorithm to find the optimum sequences. From the numerical results we could see that if the optimum sequences are used a better channel estimation quality is achieved for low to medium E_b/N_0 or, alternatively, a shorter training sequence is necessary for the same estimation quality.

Acknowledgment

The authors would like to thank Amine Mezghani for his suggestions and for the very fruitful discussions. This work is partially supported by the European project EMPhAtiC (ICT-318362).

References

[1] 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 Proces. Adv. in Wireless Comm, SPAWC 2008*, Recife, Brazil, pp. 525–529.

[2] 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*, Taipei, Taiwan, pp. 2802–2805.

[3] L. G. Baltar, M. Newinger, and J. A. Nossek, "Structured subchannel impulse response estimation for filter bank based multicarrier systems," in *Proc. 2012 Int. Symp. Wireless Comm. Systems, ISWCS*, Paris, France, pp. 191–195.

[4] L. G. Baltar and J. A. Nossek, "Multicarrier systems: A comparison between filter bank based and cyclic prefix based OFDM," in *Proc. 17th Int. OFDM Workshop 2012, InOWo'12*, Essen, Germany, pp. 1–5.

[5] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[6] T. Karp and N. Fliege, "Modified DFT filter banks with perfect reconstruction," *IEEE Trans. Circuits Syst. II*, vol. 46, no. 11, pp. 1404–1414, Nov. 1999.

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

[8] A. Oppenheim, R. Schafer, and J. Buck, *Discrete-Time Signal Processing*, 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 1997.

[9] B. C. Ng, M. Cedervall, and A. Paulraj, "A structured channel estimator for maximum likelihood sequence detection in multipath fading channels," *IEEE Trans. Veh. Technol.*, vol. 48, no. 4, pp. 1216–1228, Jul. 1999.

[10] A. Goldstein, "Convex programming in Hilbert space," *Bull. Amer. Math. Soc.*, vol. 70, no. 5, pp. 709–710, May 1964.

[11] D. Bertsekas, "On the Goldstein-Levitin-Polyak gradient projection method," *IEEE Trans. Autom. Control*, vol. 21, no. 2, pp. 174–184, Apr. 1976.