Bit Error Probability for Pilot-Symbol Aided Channel Estimation in FBMC-OQAM

Ronald Nissel[†,‡] and Markus Rupp[†]
[†] Institute of Telecommunications, Technische Universität Wien, Vienna, Austria
[‡] Christian Doppler Laboratory for Dependable Wireless Connectivity for the Society in Motion, ITC, TU Wien
Email: {rnissel, mrupp}@nt.tuwien.ac.at

***Abstract*—Filter Bank MultiCarrier (FBMC) might replace Orthogonal Frequency Division Multiplexing (OFDM) in 5G. For such FBMC system, we derive closed-form expressions for the Bit Error Probability (BEP) including channel estimation, whereas we focus on the comparison of FBMC with OFDM. We assume additive white Gaussian noise and a Rayleigh fading channel with low delay spread and low Doppler spread, so that the channel induced interference can be neglected compared to the noise. Our channel estimation is based on pilot symbols whereby the imaginary interference, inherently caused in FBMC, is canceled at the pilot positions either by auxiliary symbols or through coding. Moreover, we propose an optimal power allocation between pilot symbols and data symbols to minimize the BEP.**

I. Introduction

There is currently a lot of discussion [1] which modulation format will succeed Orthogonal Frequency Division Multiplexing (OFDM) in 5G, if at all. A promising candidate is Filter Bank MultiCarrier (FBMC) due to its superior spectral properties [2]. In FBMC, channel estimation becomes more challenging than in OFDM due to the imaginary interference. In order to straightforwardly apply pilot-symbol aided channel estimation in the same way as in OFDM, we have to cancel the imaginary interference either by auxiliary pilots [3] or by coding [4]. Besides pilot-symbol aided channel estimation, preamble based channel estimation [5] is also a possible method. However, the LTE standard employs pilot symbols because they allow a simple tacking of the channel, so that we will also focus on this method.

Measurements conducted with the Vienna Wireless Testbed show that the channel is highly correlated in the frequency domain, for our specific setup [6]. For a small number of subcarriers, it is even possible to model the channel as flat fading [7]. Moreover, we expect a further decrease in the size of the cells, decreasing also the delay spread. We will therefore focus on scenarios with low delay and low Doppler spreads, so that noise dominates the channel induced interference, which can thus be neglected. In particular, for our numerical example, we will consider the important special case of doubly block-fading, that is, for a short transmit block, the channel is time-invariant and flat. Of course, the channel of different blocks can vary and follows a Rayleigh distribution.

Novel contribution:

Firstly, we derive closed-form Bit Error Probability (BEP) expressions for FBMC systems which employ pilot-symbol aided channel estimation. Previous works [8]–[10] considered only the case of OFDM. Although the BEP differences between OFDM and FBMC are relatively small, our closed-form expressions give analytical insights and help to understand the differences between OFDM and FBMC in the context of channel estimation.

Secondly, our derivation allows to optimally allocate pilot power in order to minimize BEP. Again, previous works, for example [11], concentrated only on OFDM.

II. Multicarrier Modulation

In our transmission system model, data symbols $x_{l,k}$, at subcarrier-position l and time-position k, are modulated by the basis pulse $g_{l,k}(t)$, so that the transmitted signal $s(t)$ of a short transmit block can be written as:

$$s(t) = \sum_{k=0}^{K-1} \sum_{l=0}^{L-1} g_{l,k}(t)\, x_{l,k}, \tag{1}$$

with

$$g_{l,k}(t) = p(t - kT)\, \mathrm{e}^{\mathrm{j}2\pi\, lF\,(t-kT)}\, \underline{\mathrm{e}^{\mathrm{j}\frac{\pi}{2}(l+k)}}. \tag{2}$$

The basis pulse in (2) is, essentially, a time and frequency shifted version of the (unit energy) prototype filter $p(t)$. The time spacing T together with the frequency spacing F determines the spectral efficiency. We want to have T and F as small as possible, however, orthogonality of the basis pulses $g_{l,k}(t)$ require $TF \geq 1$. Beside orthogonality and $TF = 1$, another desired property is localization of the prototype filter $p(t)$ in both, time and frequency. Unfortunately, not all of these properties can be fulfilled at the same time according to the Balian-Low theorem. In OFDM, the prototype filter is not localized in the frequency domain, while in FBMC, the (complex) orthogonality condition is replaced by the less strict real orthogonality condition. Note that the underlined exponential expression in (2) guarantees exactly such real orthogonality in FBMC and is not required in OFDM.

To simplify analytical investigations, we transform our transmission model in the discrete time domain by sampling (2) with Δt and stacking these samples in a large matrix $\mathbf{G} \in \mathbb{C}^{\left(\frac{(K-1)T+6T_0}{\Delta t}+1\right)\times LK}$, so that the i-th row and the $l+kL$-th column of $\mathbf{G}$ is given by:

$$[\mathbf{G}]_{i,l+kL} = \sqrt{\Delta t}\; g_{l,k}(t)\Big|_{t=\Delta t\, i - 3T_0}. \tag{3}$$

The time shift $3T_0$ reflects the fact that the transmit signal starts before $t=0$, whereas T_0 will be defined later in this

978-1-4799-6664-6/16/$31.00 ©2016 IEEE

section. The discrete transmit signal in (1) can then be written in vector notation by:

$$\mathbf{s} = \mathbf{G}\mathbf{x}, \tag{4}$$

where the vector $\mathbf{x} \in \mathbb{C}^{LK\times 1}$ consists of all data symbols in vectorized form. Ignoring edge effects allows us to define the average transmit power P_S according to:

$$P_\mathrm{S} = \frac{1}{KT}\,\mathrm{tr}\left(\mathbb{E}\left\{\mathbf{s}\mathbf{s}^H\right\}\right)\Delta t. \tag{5}$$

For a fair comparison of different transmission techniques, we always consider the same average transmit power P_S.

We assume Additive White Gaussian Noise (AWGN) and a Rayleigh fading channel for which the Doppler spread and the delay spread are assumed to be so low, that inter-carrier interference and inter-symbol interference can be neglected. We then obtain the received data symbols by projection of the sampled received signal onto the sampled basis pulses, $\mathbf{y} = \mathbf{G}^H\mathbf{r}$, so that the whole transmission chain can be written as:

$$\mathbf{y} = \mathrm{diag}(\mathbf{h})\mathbf{D}\mathbf{x} + \mathbf{n}, \tag{6}$$

with

$$\mathbf{D} = \mathbf{G}^H\mathbf{G}. \tag{7}$$

The vector $\mathbf{h} \in \mathbb{C}^{LK\times 1}$ in (6) represents the channel, $\mathbf{h} \sim \mathcal{CN}(0, \mathbf{R}_\mathbf{h})$, and $\mathbf{n}$ the Gaussian distributed noise, $\mathbf{n} \sim \mathcal{CN}(0, P_\mathrm{n}\,\mathbf{D})$. According to our assumptions, only the received noise samples (each sample has a power of P_n) are white, while the demodulation step might color the noise $\mathbf{n}$. Only for a unitary matrix $\mathbf{G}$ the noise stays white.

A. OFDM

Most authors consider OFDM based on rectangular pulses which perform poorly in the frequency domain but allow efficient implementations by using fast Fourier transformations, making them very practical. The prototype filter can be written as:

$$p(t) = \begin{cases} \frac{1}{\sqrt{T_0}} & \text{if } \frac{T_0}{2} < t \le \frac{T_0}{2} \\ 0 & \text{otherwise} \end{cases}, \tag{8}$$

and leads to basis pulses $g_{l,k}(t)$ which are orthogonal for a time spacing of $T = T_0$ and a frequency spacing of $F = \frac{1}{T_0}$, achieving the maximum spectral efficiency of $TF = 1$. In practice, a Cyclic Prefix (CP) is usually included, extending the rectangular pulse at the transmitter by the length of the CP, so that the condition of orthogonality transform to bi-orthogonality. Many authors [4] then consider the same E_b/N_0 to compare OFDM with and without CP. However, we think that such approach gives not a fair comparison because the same E_b/N_0 requires a lower transmit power P_S for CP OFDM. The effects of a CP should be accounted for in the capacity as well as in the measured throughput, but not in terms of an artificial normalization of the transmit power. To avoid these issues, we consider OFDM without CP which allows a fair comparison to FBMC for the same average transmit power. In OFDM, the matrix $\mathbf{D}$ in (7) becomes an identity matrix, $\mathbf{D} = \mathbf{I}_{LK}$, so that our transmission model in (6) simplifies to:

$$y_{l,k} = h_{l,k}x_{l,k} + n_{l,k}. \tag{9}$$

Here, $h_{l,k}$ represents the channel (transfer function) at subcarrier position l and time position k.

B. FBMC

In contrast to OFDM, there is no specific prototype filter defined in FBMC, allowing us to choose from a large range of possible pulses. We consider a prototype filter which is based on Hermite polynomials $H_i(\cdot)$, because it offers good localization in both, time and frequency, and the pulse has the same shape in the frequency domain as in the time domain. The prototype filter can be written as [12]:

$$p(t) = \frac{1}{\sqrt{T_0}}\mathrm{e}^{-2\pi\left(\frac{t}{T_0}\right)^2}\sum_{\substack{i=\{0,4,8,\\12,16,20\}}} a_i H_i\left(2\sqrt{\pi}\frac{t}{T_0}\right), \tag{10}$$

for which the coefficients a_i can be found in [13]. The corresponding basis pulses $g_{l,k}(t)$ are orthogonal for a time spacing of $T = T_0$ and a frequency spacing of $F = \frac{2}{T_0}$, resulting in a time-frequency spacing of $TF = 2$. In order to achieve maximal spectral efficiency, that is, $TF = 1$, only real-valued data symbols are transmitted and the orthogonality condition is replaced by the real-orthogonality condition. The time spacing as well as the frequency spacing can then be reduced by a factor of two, $T = \frac{T_0}{2}$ and $F = \frac{1}{T_0}$, so that $TF = \frac{1}{2}$ (real symbols) which is equivalent to $TF = 1$ (complex symbols). As indicated in (2), the basis pulses have to include a phase shift of $\frac{\pi}{2}$ between adjacent symbols in order to keep the interference, caused by the time-frequency squeezing, purely imaginary. In FBMC, the matrix $\mathbf{D}$ has imaginary elements at the off-diagonal and only by taking the real part, we observe an identity matrix, $\Re\{\mathbf{D}\} = \mathbf{I}_{LK}$. Let us define the vector $\mathbf{d}_{l,k} \in \mathbb{C}^{1\times LK}$ as the $l + kL$-th row of $\mathbf{D}$. The transmission model in (6) can then be simplified to:

$$y_{l,k} = h_{l,k}\left(x_{l,k} + \mathrm{j}\,\Im\{\mathbf{d}_{l,k}\}\mathbf{x}\right) + n_{l,k}, \tag{11}$$

where $\Im\{\mathbf{d}_{l,k}\}\mathbf{x}$ represents the imaginary interference. Clearly, by taking the real part after channel equalization (perfect channel knowledge), we completely cancel the imaginary interference so that we end up with the same equation as in OFDM, compare to (9). Note that, if we consider the same transmit power P_S, see (5), the data symbol power $\mathbb{E}\{|x_{l,k}|^2\}$ in FBMC is only half as much as in OFDM due to the overlapping structure of the pulses in FBMC. However, we have to keep in mind that FBMC uses real symbols while OFDM employs complex-valued symbols.

III. Pilot-Symbol Aided Channel Estimation

In pilot-symbol aided channel estimation, certain data symbols, the so called pilots, are known a priory at the receiver. By dividing the received symbols by the pilot symbols, we obtain a Least Squares (LS) estimation of the channel $h_{l,k}$ at the pilot positions $(l,k) \in \mathcal{P}$. In vector notation, this reads as:

$$\hat{\mathbf{h}}_\mathcal{P}^\mathrm{LS} = \mathrm{diag}(\mathbf{x}_\mathcal{P})^{-1}\mathbf{y}_\mathcal{P}, \tag{12}$$

where $\mathbf{x}_\mathcal{P}$ represents all pilot symbols in vectorized form and $\mathbf{y}_\mathcal{P}$ the received data symbols at the pilot positions. Such LS

channel estimates can be interpreted as sampling of the transfer function at the pilot positions. In order to obtain the remaining channel coefficients, that is, the channel at the data positions, we interpolate and extrapolate:

$$\hat{\mathbf{h}} = \mathbf{A}\hat{\mathbf{h}}_{\mathcal{P}}^{\mathrm{LS}} \tag{13}$$

$$\hat{h}_{l,k} = \mathbf{a}_{l,k}\hat{\mathbf{h}}_{\mathcal{P}}^{\mathrm{LS}}. \tag{14}$$

Here, $\mathbf{A} \in \mathbb{C}^{LK\times|\mathcal{P}|}$ describes a large range of possible interpolation methods, such as, linear, spline, nearest neighbor or Minimum Mean Squared Error (MMSE) interpolation. The row interpolation vector $\mathbf{a}_{l,k} \in \mathbb{C}^{1\times|\mathcal{P}|}$ in (14) reflects the $l+kL$-th row of $\mathbf{A}$ and delivers the estimated channel $\hat{h}_{l,k}$ at subcarrier position l and time position k.

In OFDM, the channel estimation works exactly as described above but in FBMC, the estimation process becomes more challenging due to the imaginary interference described in (11). Taking the real part in order to get rid of this imaginary interference only works after channel equalization which is clearly not possible prior the channel estimation. We therefore have to operate in the complex domain. Note that the random imaginary interference term has the same power as the data symbols, leading to a signal-to-interference ratio of 0 dB, which is too low for accurate channel estimations.

In order to straightforwardly use the same estimation procedure as in OFDM, we therefore have to cancel the imaginary interference by modifying the transmitted data symbols by a matrix $\mathbf{C}$, so that the following condition is fulfilled:

$$\Im\{\mathbf{d}_{l,k}\}\,\mathbf{C}\begin{bmatrix}\mathbf{x}_{\mathcal{P}}\\ \mathbf{x}_{\mathcal{D}}\end{bmatrix} = 0 \quad \text{for} \quad (l,k)\in\mathcal{P}. \tag{15}$$

Note that the imaginary interference in (15) is only canceled at the pilot positions, whereas at the data positions we still observe imaginary interference which has to be canceled by taking the real part after channel equalization. In (15), we also distinguish between pilot symbols $\mathbf{x}_{\mathcal{P}}$, whereas each element has a power of $P_{\mathcal{P}}$, and data symbols $\mathbf{x}_{\mathcal{D}}$ with power $P_{\mathcal{D}}$. In literature, there are two methods proposed to cancel the imaginary interference: auxiliary pilots [3] and coding [4].

A. Auxiliary pilots

For each pilot symbol, one auxiliary symbol is used to cancel the imaginary interference at the corresponding pilot position. The idea is quite simple: several symbols close to the pilot symbol cause imaginary interference. One of these symbols is then used to cancel the interference of all other symbols. However, the interference weight between such symbol and the pilot symbol is below one, so that the auxiliary pilot power has to be increased, additionally reducing the available power for data and pilot symbols. Auxiliary pilots are only used to simplify the channel estimation at the receiver but they do not directly carry any information, so that they are "wasted" energy which constitutes the main drawback of this method. In terms of spectral efficiency, we have the same situation as in OFDM because two real symbols are used, which is equivalent to one complex symbol. In (15), we can describe such auxiliary pilots by a matrix $\mathbf{C} = \mathbf{C}_{\mathrm{a}} \in \mathbb{R}^{LK\times(LK-|\mathcal{A}|)}$. The diagonal elements of the matrix $\mathbf{C}\mathbf{C}^T$ consists mostly of ones except at auxiliary positions, where they describe the auxiliary-to-data power offset. The mean value over all auxiliary positions of such offset is denoted by $\bar{\kappa}_{\mathcal{A}\mathcal{D}}$.

B. Coding

By coding at both, transmitter and receiver, we can cancel the imaginary interference at the pilot positions while at the same time, we do not lose any useful transmit power which was the main drawback of auxiliary pilot symbols. For each pilot symbol, we spread $N-1$ data symbols over N closest time-frequency positions. This provides us with one additional degree of freedom which can be used to cancel the imaginary interference at the pilot position. After channel estimation we have to reassemble the data symbols, so that the reassembled received symbol vector $\tilde{\mathbf{y}}$ is given by:

$$\tilde{\mathbf{y}} = \mathbf{C}_{\mathrm{c}}^T\,\mathbf{y}. \tag{16}$$

Beside the cancellation condition in (15), matrix $\mathbf{C} = \mathbf{C}_{\mathrm{c}}$ must also be semi-orthogonal, so that $\mathbf{C}_{\mathrm{c}}^T\mathbf{C}_{\mathrm{c}}$ becomes an identity matrix of size $LK - |\mathcal{P}|$. A detailed algorithm of how to construct such matrix $\mathbf{C}_{\mathrm{c}}$ can be found in [13].

IV. Bit Error Probability

In order to express the BEP in closed-form, we apply the following Lemma:

Lemma 1: Let X and Y be zero mean, correlated complex-valued Gaussian random variables, then the probability that the real part of the complex Gaussian ratio $\frac{X}{Y}$ is smaller than a certain value z_{R}, reads:

$$\Pr\left(\Re\left\{\frac{X}{Y}\right\} < z_{\mathrm{R}}\right) = \frac{1}{2} - \frac{\Re\left\{\frac{\mathbb{E}\{XY^*\}}{\mathbb{E}\{|Y|^2\}}\right\} - z_{\mathrm{R}}}{2\sqrt{\left(\Re\left\{\frac{\mathbb{E}\{XY^*\}}{\mathbb{E}\{|Y|^2\}}\right\} - z_{\mathrm{R}}\right)^2 + \frac{\mathbb{E}\{|X|^2\}}{\mathbb{E}\{|Y|^2\}} - \left|\frac{\mathbb{E}\{XY^*\}}{\mathbb{E}\{|Y|^2\}}\right|^2}}. \tag{17}$$

Lemma 1 is a special case of the more general expression given in [14] which also includes the imaginary part. As shown in (17), the second order statistics play an important role. In order to keep the notation short, we therefore denote correlation vectors by $\mathbf{r}$, for example $\mathbf{r}_{\mathbf{h}_{\mathcal{P}},h_{l,k}} = \mathbb{E}\{\mathbf{h}_{\mathcal{P}}h_{l,k}^*\}$, and the correlation matrices by $\mathbf{R}$, for example, $\mathbf{R}_{\mathbf{h}_{\mathcal{P}}} = \mathbb{E}\{\mathbf{h}_{\mathcal{P}}\mathbf{h}_{\mathcal{P}}^H\}$. Although our equations can be extended to higher modulation orders, see for example [10] for the case of OFDM, we here derive explicit expressions only for a 4-Quadrature Amplitude Modulation (QAM) signal constellation in OFDM and a 2-Pulse-Amplitude Modulation (PAM) signal constellation in FBMC because the solutions are relatively compact. Because 4-QAM and 2-PAM are independent of the estimated amplitude, we normalize the interpolation vector in such a way that $\mathbf{a}_{l,k}\mathbf{1}_{|\mathcal{P}|\times 1} = 1$, that is, the row sum of $\mathbf{A}$ is one. Two important interpolation methods, which allow a compact expression of the BEP because they are independent of specific time and frequency positions are Nearest neighbor interpolation, $\mathbf{a}_{l,k}\mathbf{a}_{l,k}^H = 1$, and MMSE interpolation (for flat fading), $\mathbf{a}_{l,k}\mathbf{a}_{l,k}^H = \frac{1}{|P|}$.

A. OFDM

For a compact description, we assume that the interpolation vector $\mathbf{a}_{l,k}$ is chosen so that $\Im\{\mathbf{a}_{l,k}\mathbf{r}_{\mathbf{h}_\mathcal{P},h_{l,k}}\} = 0$ which guarantees that the BEP conditioned a "0' was sent is the same as the BEP conditioned a "1" was sent. Such condition is also required for minimizing the BEP. Using the OFDM signal model in (9) and applying a zero forcing equalizer, allows us to write the BEP for a Gray coded 4-QAM signal constellation as [10]:

$$\mathrm{BEP}_{l,k} = \Pr\left(\Re\left\{\frac{y_{l,k}}{\hat{h}_{l,k}}\right\} < 0 \middle| x_{l,k} = \sqrt{P_\mathcal{D}}\,\frac{1+\mathrm{j}}{\sqrt{2}}\right). \quad (18)$$

In order to employ Lemma 1, we need the corresponding correlation and powers, conditioned on $x_{l,k}$, given by:

$$\mathbb{E}\{y_{l,k}\hat{h}^*_{l,k}|x_{l,k}\} = \mathbf{r}^H_{\mathbf{h}_\mathcal{P},h_{l,k}}\,\mathbf{a}^H_{l,k}\,x_{l,k} \quad (19)$$

$$\mathbb{E}\{|y_{l,k}|^2|x_{l,k}\} = |x_{l,k}|^2 + P_\mathrm{n} \quad (20)$$

$$\mathbb{E}\{|\hat{h}_{l,k}|^2\} = \mathbf{a}_{l,k}\left(\mathbf{R}_{\mathbf{h}_\mathcal{P}} + \frac{P_\mathrm{n}}{P_\mathcal{P}}\mathbf{I}_{|\mathcal{P}|}\right)\mathbf{a}^H_{l,k}. \quad (21)$$

By using Lemma 1 in (18) together with (19) to (21), we immediately find a closed-form solution for the BEP which is valid for doubly-selective channels as long as the noise dominates the interference. Let us now consider the special case of a doubly block-fading channel. The correlation matrix in (21) then transforms to $\mathbf{R}_{\mathbf{h}_\mathcal{P}} = \mathbf{1}_{|\mathcal{P}|}$ and the correlation vector in (19) to $\mathbf{r}_{\mathbf{h}_\mathcal{P},h_{l,k}} = \mathbf{1}_{|\mathcal{P}|\times 1}$. This simplifies the BEP in (18):

$$\mathrm{BEP}_{l,k} = \frac{1}{2} - \frac{1}{2\sqrt{2\left(1+\frac{P_\mathrm{n}}{P_\mathcal{P}}\mathbf{a}_{l,k}\mathbf{a}^H_{l,k}\right)\left(1+\frac{P_\mathrm{n}}{P_\mathcal{D}}\right)-1}}, \quad (22)$$

where we consider a normalized interpolation vector $\mathbf{a}_{l,k}$. Beside the data symbol power $P_\mathcal{D}$, the pilot symbol power $P_\mathcal{P}$, and the noise power P_n, the BEP also depends on the scalar product $\mathbf{a}_{l,k}\mathbf{a}^H_{l,k}$, given for the nearest neighbor interpolation by $\mathbf{a}_{l,k}\mathbf{a}^H_{l,k} = 1$ and for an MMSE interpolation by $\mathbf{a}_{l,k}\mathbf{a}^H_{l,k} = \frac{1}{|P|}$. In general, however, $\mathbf{a}_{l,k}$ depends on a specific time-frequency position, for example, linear interpolation leads to $\mathbf{a}_{l,k}\mathbf{a}^H_{l,k} < 1$, while linear extrapolation to $\mathbf{a}_{l,k}\mathbf{a}^H_{l,k} > 1$.

B. FBMC, Auxiliary Pilots

In contrast to OFDM, the BEP at subcarrier position l and time position k now also depends on all other symbols due to the imaginary interference. Applying a zero forcing equalizer to the FBMC transmission model in (11), allows us to find the BEP for 2-PAM by:

$$\mathrm{BEP}_{l,k} = \frac{1}{2^{|\mathcal{D}||\mathcal{P}|}}\sum_{\mathbf{x}_\mathcal{D}\in|\mathcal{X}|^{|\mathcal{D}|}}\sum_{\mathbf{x}_\mathcal{P}\in|\mathcal{X}|^{|\mathcal{P}|}} \Pr\left(\Re\left\{\frac{y_{l,k}}{\hat{h}_{l,k}}\right\} < 0 \middle| x'_{l,k} = \sqrt{P_\mathcal{D}} + \mathrm{j}\,\Im\{\mathbf{d}_{l,k}\}\mathbf{C}_\mathrm{a}\begin{bmatrix}\mathbf{x}_\mathcal{P}\\ \mathbf{x}_\mathcal{D}\end{bmatrix}\right). \quad (23)$$

A large number of data symbols and pilot symbols makes the evaluation of (23) computationally complex. Because only symbols closest to $x_{l,k}$ have a significant effect on its BEP, we can reduce the computational complexity by including only those closest symbols in our evaluation. Another method to reduce the complexity is Monte Carlo simulation, which we will use in Section VI. Similar as before, we need the second order statistics in order to apply Lemma 1, which we find by:

$$\mathbb{E}\{|y_{l,k}|^2|\mathbf{x}\} = |x'_{l,k}|^2 + P_\mathrm{n} \quad (24)$$

$$\mathbb{E}\{y_{l,k}\hat{h}^*_{l,k}|\mathbf{x}\} = \mathbf{r}^H_{\mathbf{h}_\mathcal{P},h_{l,k}}\mathbf{a}^H_{l,k}\,x'_{l,k} + \mathbf{r}^H_{\mathbf{n}_\mathcal{P},n_{l,k}}\mathrm{diag}(\mathbf{x}_\mathcal{P})^{-1}\mathbf{a}^H_{l,k} \quad (25)$$

$$\mathbb{E}\{|\hat{h}_{l,k}|^2|\mathbf{x}_\mathcal{P}\} = \mathbf{a}_{l,k}\left(\mathbf{R}_{\mathbf{h}_\mathcal{P}} + \mathrm{diag}(\mathbf{x}_\mathcal{P})^{-1}\mathbf{R}_{\mathbf{n}_\mathcal{P}}\,\mathrm{diag}(\mathbf{x}_\mathcal{P})^{-1}\right)\mathbf{a}^H_{l,k}. \quad (26)$$

Opposed to OFDM, the noise term $\mathbf{n}$ is now colored which has to be included in (25) and (26). Applying Lemma 1 in (23), delivers immediately the desired BEP for FBMC. However, the exact solution might not always be necessary so that we propose the following approximation:

$$\mathrm{BEP}_{l,k} \approx \Pr\left(\Re\left\{\frac{y_{l,k}}{\hat{h}_{l,k}}\right\} < 0 \middle| x_{l,k} = \sqrt{P_\mathcal{D}} + \mathrm{j}\sqrt{P_\mathcal{D}}\right), \quad (27)$$

where we replace the "random" imaginary interference term in (23) by a fixed value with the same power. We ignore edge effects and the pilot-to-data power offset, so that the imaginary interference power is the same as the data symbol power. For a doubly block-fading channel and by approximating the colored noise in (25) and (26) by white noise of identical power, (27) in combination with Lemma 1 then leads to:

$$\mathrm{BEP}_{l,k} \approx \frac{1}{2} - \frac{1}{2\sqrt{2\left(1+\frac{P_\mathrm{n}}{P_\mathcal{P}}\mathbf{a}_{l,k}\mathbf{a}^H_{l,k}\right)\left(1+\frac{P_\mathrm{n}}{2P_\mathcal{D}}\right)-1}}. \quad (28)$$

The only difference to (22) is an additional factor of two inside the right bracket. Note, however, that FBMC uses different data and pilot symbol powers when compared to OFDM.

C. FBMC, Coding

Here, the received data symbols have to be decoded according to $\tilde{\mathbf{y}} = \mathbf{C}^T_\mathrm{c}\,\mathbf{y}$. Let us denote the vector $\mathbf{c}_{\mathrm{c_i}}$ as the i-th column vector of $\mathbf{C}_\mathrm{c}$. Due to coding, data symbols close to the pilot symbols are spread over near subcarriers and time positions, so that these data symbols no longer belong to a specific subcarrier position l and time position k. We therefore write the index i instead of (l,k), so that the i-th reassembled received symbol is given by:

$$\tilde{y}_i = h_i\,\mathbf{c}^T_{\mathrm{c_i}}\,\mathbf{D}\,\mathbf{C}_c\begin{bmatrix}\mathbf{x}_\mathcal{P}\\ \mathbf{x}_\mathcal{D}\end{bmatrix} + \mathbf{c}^T_{\mathrm{c_i}}\mathbf{n}. \quad (29)$$

Similar as before, we assume a low delay spread and a low Doppler spread, allowing us to separate the channel h_i as multiplicative factor. Besides the decoding of the received data symbols, we also have to adapt the channel estimates so that they reflect the coding weights. Such channel estimation matrix can be found by:

$$\mathbf{C}_h = \mathbf{C}_c \circ \mathbf{C}_c, \quad (30)$$

where the $\circ$ operator is the Hadamard (point-wise) product. Note that the column sum of $\mathbf{C}_h$ is one. Finally, using the i-th column vector of $\mathbf{C}_\mathrm{h}$, denoted by $\mathbf{c}_{\mathrm{h_i}}$, gives the channel estimate for the reassembled received symbol as:

$$\hat{\tilde{h}}_i = \mathbf{c}^T_{\mathrm{h_i}}\mathbf{A}\,\hat{\mathbf{h}}^{\mathrm{LS}}_\mathcal{P}. \quad (31)$$

TABLE I
OPTIMAL (MINIMIZES THE BEP) PILOT-TO-DATA POWER OFFSET κ_{OPT}.

	OFDM	FBMC, Auxiliary	FBMC, Coding
	$\epsilon_{\mathcal{D}} + \epsilon_{\mathcal{P}} = 1$	$\epsilon_{\mathcal{D}} + \epsilon_{\mathcal{P}} + \epsilon_{\mathcal{A}} = 1$	$\epsilon_{\mathcal{D}} + \epsilon_{\mathcal{P}} = 1 - \epsilon_{\mathcal{P}}$
SNR	$\frac{\lvert\mathcal{D}\rvert P_{\mathcal{D}} + \lvert\mathcal{P}\rvert P_{\mathcal{P}}}{LKP_{\text{n}}}$	$\frac{\lvert\mathcal{D}\rvert P_{\mathcal{D}} + \lvert\mathcal{P}\rvert P_{\mathcal{P}} + \lvert\mathcal{A}\rvert P_{\mathcal{A}}}{LK\frac{P_{\text{n}}}{2}}$	$\frac{\lvert\mathcal{D}\rvert P_{\mathcal{D}} + \lvert\mathcal{P}\rvert P_{\mathcal{P}}}{LK\frac{P_{\text{n}}}{2}}$
κ_{opt}	$\sqrt{\frac{\epsilon_{\mathcal{D}}}{\epsilon_{\mathcal{P}}}\frac{\left(\frac{\epsilon_{\mathcal{D}}}{\text{SNR}}+1\right)}{\left(\frac{\epsilon_{\mathcal{P}}}{\text{SNR}}+\frac{1}{a}\right)}}$	$\sqrt{\frac{\tilde{\epsilon}_{\mathcal{D}\mathcal{A}}}{\epsilon_{\mathcal{P}}}\frac{\left(\frac{\tilde{\epsilon}_{\mathcal{D}\mathcal{A}}}{\text{SNR}}+1\right)}{\left(\frac{\epsilon_{\mathcal{P}}}{\text{SNR}}+\frac{1}{2a}\right)}}$	$\sqrt{\frac{\epsilon_{\mathcal{D}}}{\epsilon_{\mathcal{P}}}\frac{\left(\frac{\epsilon_{\mathcal{D}}}{\text{SNR}}+1\right)}{\left(\frac{\epsilon_{\mathcal{P}}}{\text{SNR}}+\frac{1}{2a}\right)}}$

$\epsilon_{\mathcal{D}} = \frac{\lvert\mathcal{D}\rvert}{KL}$; $\epsilon_{\mathcal{P}} = \frac{\lvert\mathcal{P}\rvert}{KL}$; $\epsilon_{\mathcal{A}} = \frac{\lvert\mathcal{A}\rvert}{KL}$; $\tilde{\epsilon}_{\mathcal{D}\mathcal{A}} = \epsilon_{\mathcal{D}} + \bar{\kappa}_{\mathcal{A}\mathcal{D}}\,\epsilon_{\mathcal{A}}$
$\kappa = \frac{P_{\mathcal{P}}}{P_{\mathcal{D}}}$ Nearest neighbor: $a = 1$; MMSE: $a = \frac{1}{\lvert\mathcal{P}\rvert}$

Such transmission model allows us to find the BEP by:

$$\text{BEP}_{l,k}^{\text{2PAM}} = \frac{1}{2^{|\mathcal{D}||\mathcal{P}|}} \sum_{\mathbf{x}_{\mathcal{D}} \in |\mathcal{X}|^{|\mathcal{D}|}} \sum_{\mathbf{x}_{\mathcal{P}} \in |\mathcal{X}|^{|\mathcal{P}|}} \Pr\left(\Re\left\{ \frac{\tilde{y}_i}{\hat{\tilde{h}}_i} \right\} < 0 \middle| x'_{l,k} = \sqrt{P_{\mathcal{D}}} + \text{j}\, \mathbf{c}_{\text{c}_\text{i}}^T \Im\{\mathbf{D}\}\, \mathbf{C}_c \begin{bmatrix} \mathbf{x}_{\mathcal{P}} \\ \mathbf{x}_{\mathcal{D}} \end{bmatrix} \right), \tag{32}$$

with the corresponding second order statistics for Lemma 1:

$$\mathbb{E}\{|\tilde{y}_i|^2 | \mathbf{x}\} = |x'_{l,k}|^2 + P_{\text{n}} \tag{33}$$

$$\mathbb{E}\{\tilde{y}_i \hat{\tilde{h}}_i^* | \mathbf{x}\} = \mathbf{r}_{\mathbf{h}_{\mathcal{P}},h}^H \mathbf{A}^H \mathbf{c}_{\text{h}_\text{i}}\, x'_{l,k} + \mathbf{c}_{\text{c}_\text{i}}^T \mathbf{R}_{\mathbf{n},\mathbf{n}_{\mathcal{P}}} \text{diag}(\mathbf{x}_{\mathcal{P}})^{-1} \mathbf{A}^H \mathbf{c}_{\text{h}_\text{i}} \tag{34}$$

$$\mathbb{E}\{|\hat{\tilde{h}}_i|^2 | \mathbf{x}_{\mathcal{P}}\} = \mathbf{c}_{\text{h}_\text{i}}^T \mathbf{A} \left(\mathbf{R}_{\mathbf{h}_{\mathcal{P}}} + \text{diag}(\mathbf{x}_{\mathcal{P}})^{-1} \mathbf{R}_{\mathbf{n}_{\mathcal{P}}} \text{diag}(\mathbf{x}_{\mathcal{P}})^{-1} \right) \mathbf{A}^H \mathbf{c}_{\text{h}_\text{i}} \tag{35}$$

We find an approximation of (32) in the same way as for auxiliary pilot symbols. This results in a similar equation as in (28), except that the interpolation factor $\mathbf{a}_{l,k}\mathbf{a}_{l,k}^H$ has to be replaced by $\mathbf{c}_{\text{h}_\text{i}}^T \mathbf{A}\mathbf{A}^H \mathbf{c}_{\text{h}_\text{i}}$. However, for nearest neighbor interpolation and MMSE interpolation this factor stays the same.

V. Optimal Power Allocation

As indicated by our closed-form BEP expressions, (22) and (28), there exists a trade-off between the data symbol power and the pilot symbol power, under the constraint of a constant transmit power. Here, we address the problem of finding the optimal power distribution in the sense of minimizing the BEP. Let us define the pilot-to-data power offset by:

$$\kappa = \frac{P_{\mathcal{P}}}{P_{\mathcal{D}}}. \tag{36}$$

For OFDM, we insert the Signal-to-Noise Ratio (SNR), see Table I, into (22), leading to:

$$\text{BEP}^{\text{OFDM}} = \frac{1}{2} - \frac{1}{2\sqrt{2\left(1 + \frac{\epsilon_{\mathcal{D}} + \epsilon_{\mathcal{P}}\kappa}{\text{SNR}\,\kappa} a\right)\left(1 + \frac{\epsilon_{\mathcal{D}} + \epsilon_{\mathcal{P}}\kappa}{\text{SNR}}\right)} - 1}. \tag{37}$$

The optimal power offset $\kappa_{\text{opt}} = \arg\min_\kappa \text{BEP}$ is then found by setting the first derivative of (37) to zero and solving with respect to κ. For FBMC, we do the same for the approximated BEP in (28). For nearest neighbor and MMSE interpolation, Table I provides the solution of this optimization problem in closed-form. The optimal pilot-to-data power offset depends, in particular, on the data symbol density $\epsilon_{\mathcal{D}}$, the pilot symbol density $\epsilon_{\mathcal{P}}$, and the auxiliary pilot symbol density $\epsilon_{\mathcal{A}}$. In

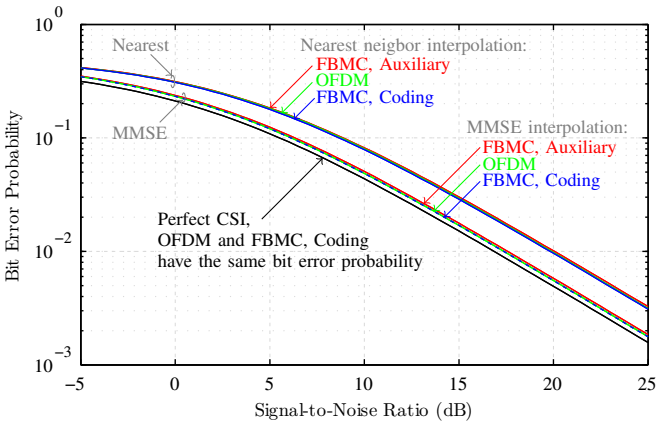

Fig. 1. The BEP strongly depends on the channel estimation method while we observe only marginal differences between FBMC and OFDM (see Figure 2).

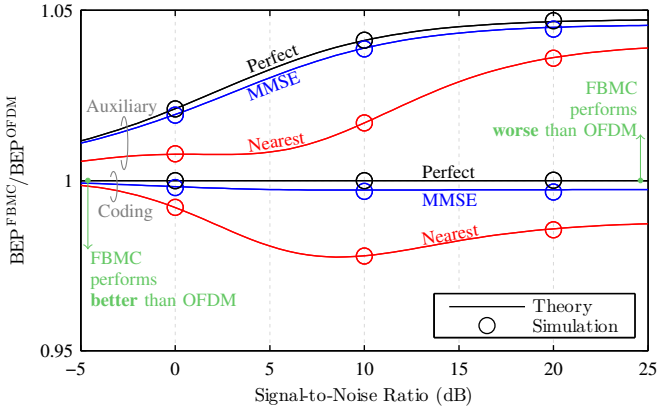

Fig. 2. FBMC which uses coding for the channel estimation, outperforms OFDM in terms of BEP due to the random imaginary interference, while for auxiliary pilots, the performance is worse due to the loss in useful energy. Simulations validate our theoretical derivations, where we see a perfect match.

FBMC, the pilot symbol density is half as much as in OFDM, so that for coding, the optimal power offset is by a factor of two higher than in OFDM. For auxiliary pilot symbols, the power offset is even higher. Note that in FBMC, the noise power is reduced by a factor of two because the imaginary part is canceled. However, for the same transmit power P_{S}, we observe the same SNR as in OFDM because the signal power is also reduced by a factor of two, due to the overlapping structure in time. In FBMC, the channel estimation at the pilot position has to be performed in the complex domain, so that only a pilot-to-data power offset of two guarantees the same data symbol $\text{SNR}_{\mathcal{D}}$ and pilot symbol $\text{SNR}_{\mathcal{P}}$. Thus, we will consider $\kappa = 2$ as our reference power allocation in FBMC. Also, for coding, such offset results in the same data symbol power $P_{\mathcal{D}}$ as for the case of no channel estimation because the coding process saves power, which can be put into pilot symbols. Additionally, a power offset of $\kappa = 2$ for FBMC (coding) and a power offset of $\kappa = 1$ for OFDM, makes both systems equivalent in terms of data symbol $\text{SNR}_{\mathcal{D}}$.

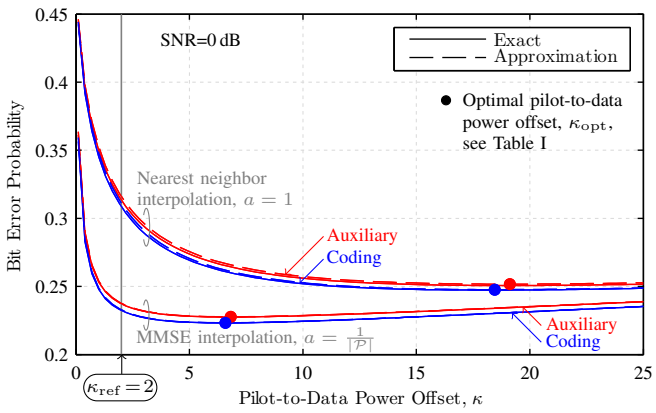

Fig. 3. By increasing the pilot-to-data power offset we can reduce the BEP. Our approximation for FBMC comes close to the exact BEP solution, in particular for MMSE interpolation where they almost perfectly coincide.

VI. Numerical Results

We consider a short transmission block, consisting of $L = 24$ subcarriers and $K = 15$ OFDM symbols, respectively $K = 30$ FBMC symbols. In accordance with our testbed measurements, we will consider a doubly block-fading channel, that is, $h_{l,k} = h$, which, however, varies for different blocks according to Rayleigh fading. For the pilot-symbol pattern, we follow the LTE standard but increase the frequency-pilot-spacing by a factor of two, so that coded symbols do not overlap. Thus, we have a diamond-shaped pilot symbol pattern with $|\mathcal{P}| = 8$ pilot symbols in total. OFDM and FBMC have the same data symbol density, $\epsilon_{\mathcal{D}} = 0.9778$, while the pilot symbol density is in OFDM twice as high as in FBMC, $\epsilon_{\mathcal{P}}^{\text{OFDM}} = 2 \cdot 0.0111$. The number of coded symbols is set to $N = 24$, resulting in a signal-to-interference ratio at the pilot positions of more than 60 dB, so that the condition in (15) is sufficiently fulfilled. The auxiliary-to-data power offset is then given by $\bar{\kappa}_{\mathcal{AD}} = 4.2664$.

Figure 1 shows the BEP, using our closed-form expressions in (22), (23) and (32). OFDM and FBMC achieve almost the same performance, so that we compare FBMC directly with OFDM in Figure 2. If coding is employed for channel estimation in FBMC, it performs better than OFDM, because the random imaginary interference in (32) is sometimes smaller than the fixed imaginary part in (18) for OFDM, reducing the influence of channel estimation errors. Of course, the random interference can also be higher, but as it turns out, the overall effect leads to a lower BEP in FBMC. Note that the colored noise has only a minor influence. In Figure 3, we show how the BEP depends on the pilot-to-data power offset, whereas the optimal power allocation can be found in Table I. The potential improvement of the optimal power offset depends strongly on the interpolation method and the SNR. For example, compared to the optimal power offset, κ_{opt}, the conventional allocation, κ_{ref}, results in a BEP which is approximately 4% higher in case of MMSE interpolation, and, depending on the SNR, 16% to 55% higher in case of nearest neighbor interpolation.

VII. Conclusion

FBMC, which employs coding to cancel the imaginary interference at the pilot positions, shows a slightly better BEP than OFDM, due to the random imaginary interference in FBMC which turns out to be advantageous compared to a fixed imaginary term as in OFDM. However, for perfect channel knowledge we observe the same BEP. On the other hand, FBMC which employs auxiliary pilots performs worse than OFDM because auxiliary pilots are "wasted" energy. Our BEP expressions also allow us to find the optimal pilot-to-data power offset in closed-form. By increasing such power offset, we are able to improve the performance, that is, for the same transmission power, a lower BEP is achieved.

Acknowledgment

The financial support by the Austrian Federal Ministry of Science, Research and Economy, the National Foundation for Research, Technology and Development, and the TU Wien is gratefully acknowledged.

References

[1] G. Wunder, P. Jung, M. Kasparick, T. Wild, F. Schaich, Y. Chen, S. Brink, I. Gaspar, N. Michailow, A. Festag *et al.*, "5GNOW: non-orthogonal, asynchronous waveforms for future mobile applications," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 97–105, 2014.

[2] B. Farhang-Boroujeny, "OFDM versus filter bank multicarrier," *IEEE Signal Processing Magazine*, vol. 28, no. 3, pp. 92–112, 2011.

[3] T. H. Stitz, T. Ihalainen, A. Viholainen, and M. Renfors, "Pilot-based synchronization and equalization in filter bank multicarrier communications," *EURASIP J. Advances in Sig. Proc.*, pp. 1–9, 2010.

[4] C. Lélé, R. Legouable, and P. Siohan, "Channel estimation with scattered pilots in OFDM/OQAM," in *IEEE 9th Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2008, pp. 286–290.

[5] E. Kofidis, D. Katselis, A. Rontogiannis, and S. Theodoridis, "Preamble-based channel estimation in OFDM/OQAM systems: a review," *Signal Processing*, vol. 93, no. 7, pp. 2038–2054, 2013.

[6] R. Nissel, M. Lerch, and M. Rupp, "Experimental validation of the OFDM bit error probability for a moving receive antenna," in *IEEE Vehicular Technology Conference (VTC)*, Vancouver, Canada, Sept 2014.

[7] R. Nissel and M. Rupp, "Doubly-selective MMSE channel estimation and ICI mitigation for OFDM systems," in *IEEE International Conference on Communications (ICC)*, London, UK, June 2015.

[8] M.-X. Chang and Y. T. Su, "Performance analysis of equalized OFDM systems in Rayleigh fading," *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, pp. 721–732, 2002.

[9] P. Tan and N. C. Beaulieu, "Effect of channel estimation error on bit error probability in OFDM systems over Rayleigh and Ricean fading channels," *IEEE Trans. Commun.*, vol. 56, no. 4, pp. 675–685, 2008.

[10] R. Nissel, M. Lerch, M. Šimko, and M. Rupp, "Bit error probability for pilot-symbol-aided OFDM channel estimation in doubly-selective channels," in *18th International ITG Workshop on Smart Antennas (WSA)*, Erlangen, Germany, Mar 2014.

[11] X. Cai and G. B. Giannakis, "Error probability minimizing pilots for OFDM with M-PSK modulation over Rayleigh-fading channels," *IEEE Transactions on Vehicular Technology*, vol. 53, no. 1, pp. 146–155, 2004.

[12] R. Haas and J.-C. Belfiore, "A time-frequency well-localized pulse for multiple carrier transmission," *Wireless Personal Communications*, vol. 5, no. 1, pp. 1–18, 1997.

[13] R. Nissel and M. Rupp, "On pilot-symbol aided channel estimation in FBMC-OQAM," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Shanghai, China, March 2016.

[14] R. Nissel, S. Caban, and M. Rupp, "Closed-form capacity expression for low complexity BICM with uniform inputs," in *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Hong Kong, China, Aug 2015.