ON PILOT-SYMBOL AIDED CHANNEL ESTIMATION IN FBMC-OQAM

Ronald Nissel *Markus Rupp*

Technische Universität Wien, Institute of Telecommunications
Gusshausstraße 25, 1040 Vienna, Austria

ABSTRACT

Filter bank multicarrier modulation is considered as a possible candidate for 5G. In this paper, we consider pilot-symbol aided channel estimation and address the problem of canceling the imaginary interference at the pilot positions. We develop a matrix formulation for the transmission system which allows us to formulate general conditions on the auxiliary pilot symbols, capturing also the interdependency of closely spaced pilots and an arbitrary number of auxiliary pilot symbols. By using two auxiliary symbols per pilot instead of one, we are able to improve the peak-to-average power ratio as well as the achievable capacity for small to medium signal-to-noise ratios. The achievable capacity can further be increased by interference cancellation based on linear precoding for which we propose an algorithm to find the coding matrix required. Finally, we compare auxiliary pilot symbols and linear precoding in terms of complexity and performance.

Index Terms— FBMC-OQAM, Channel estimation, Interference cancellation, Auxiliary pilot, Coding

1. INTRODUCTION

Multicarrier modulation [1] offers many advantages such as simple equalization and adaptive modulation and coding techniques, with Orthogonal Frequency Division Multiplexing (OFDM) being the most prominent scheme, currently employed in many wireless communication standards such as LTE and IEEE 802.11. However, OFDM is based on rectangular pulses which perform poorly in the frequency domain, causing some disadvantages [2]. Recently, Filter Bank MultiCarrier (FBMC) has been identified by many authors [3, 4, 5, 6] as a possible candidate to replace OFDM in the next generation of wireless communications systems (5G). The basic idea of FBMC dates back to the 1960s [7, 8] and was reformulated, for example, in [9, 10, 11]. Note that throughout literature, different names, such as cosine-modulated multitone, staggered multitone, discrete wavelet multitone and OFDM/Offset Quadrature Amplitude Modulation (OQAM), have been used to describe, essentially, the same concept that we will refer to here as FBMC.

Compared to OFDM, channel estimation becomes more challenging in FBMC due to the imaginary interference. Preamble based channel estimation, as for example used in IEEE 802.11, is investigated in [12, 13, 14, 15] for FBMC. However, if the channel varies in time, pilot-symbol aided channel estimation becomes a better choice because it allows to track the channel, which also explains its employment in LTE. To deal with the imaginary interference in pilot-symbol aided channel estimation, [16] uses one symbol per pilot, the so called auxiliary pilot symbol [17], to cancel the imaginary interference. The main drawback of this method is a power offset. To deal with this problem, [18] uses non-linear techniques on a discrete set of possible interference terms, while authors in [19] propose a more elegant method based on linear precoding, in the following simply coding, which completely eliminates the power offset at the expense of increased complexity. Authors in [20] combine the methods of [16] and [19], resulting in intermediate complexity and a (small) power offset.

Novel contribution:

Firstly, the method in [16] does not allow closely spaced pilots or more than one auxiliary symbol per pilot. We therefore propose a more general method to choose the auxiliary pilot symbols, avoiding these drawbacks. By, for example, using two auxiliary symbols per pilot, we are already able to reduce the power offset from 4.3 to 0.8 and to increase the achievable capacity for low Signal-to-Noise Ratio (SNR) values by approximately 5%.

Secondly, authors in [19, 20] consider only coding of up to $N = 8$ symbols. We propose an algorithm to design the coding matrix required for an arbitrary number of coded symbols.

Thirdly, we quantify the complexity difference between auxiliary pilot symbols and coding.

2. SYSTEM MODEL

In FBMC transmissions, the data symbols $x_{l,k}$ at frequency position l and time position k are modulated by the basis pulses $g_{l,k}(t)$, so that the transmit signal $s(t)$ of our system 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

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

The basis pulse $g_{l,k}(t)$ is essentially a time and frequency shifted version of the prototype filter $p(t)$ whereas the time spacing T together with the frequency spacing F determine the spectral efficiency. Our prototype filter $p(t)$ is based on Hermite polynomials $H_n(\cdot)$, as suggested in [21]:

$$p(t) = \frac{1}{\sqrt{T_0}} \mathrm{e}^{-2\pi\left(\frac{t}{T_0}\right)^2} \sum_{\substack{i=\{0,4,8,\\ 12,16,20\}}} a_i H_i\left(2\sqrt{\pi}\frac{t}{T_0}\right), \tag{3}$$

for which we found the coefficients a_i numerically:

$$\begin{aligned} a_0 &= \ \ 1.412692577 & a_{12} &= -2.2611 \cdot 10^{-9} \\ a_4 &= -3.0145 \cdot 10^{-3} & a_{16} &= -4.4570 \cdot 10^{-15}. \\ a_8 &= -8.8041 \cdot 10^{-6} & a_{20} &= \ \ 1.8633 \cdot 10^{-16} \end{aligned} \tag{4}$$

The prototype filter in (3) guarantees orthogonality of the basis pulses for a time spacing of $T = T_0$ and a frequency spacing of $F = \frac{2}{T_0}$. However, maximal spectral efficiency requires $TF = 1$

This work has been funded by the Christian Doppler Laboratory for Wireless Technologies for Sustainable Mobility, the A1 Telekom Austria AG, and the KATHREIN-Werke KG. The financial support by the Federal Ministry of Economy, Family and Youth and the National Foundation for Research, Technology and Development is gratefully acknowledged.

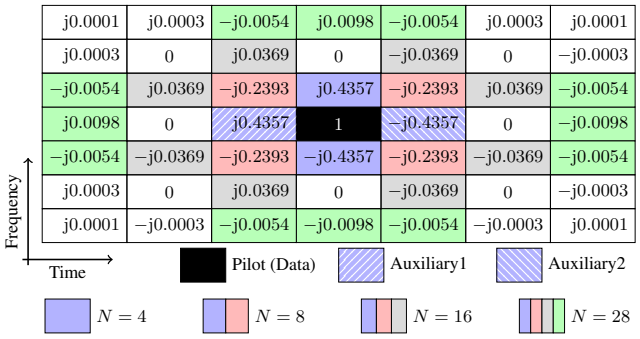

Fig. 1. Imaginary interference weights in FBMC (certain elements of $\mathbf{D}$). At pilot positions, we have to mitigate this interference by using auxiliary pilot symbols or by coding N surrounding symbols.

which is not possible for pulses that are localized in both, time and frequency, according to the Balian-Low theorem [22]. In FBMC, only real valued symbols are transmitted and the orthogonality condition is replaced with the *real* orthogonality condition. The time spacing as well as the frequency spacing can then be reduced by a factor of two, so that $TF = \frac{1}{2}T_0 \, \frac{1}{2} \frac{2}{T0} = \frac{1}{2}$ (real symbols). This results in the same spectral efficiency as OFDM without CP (and ignoring guard bands), that is, $TF = 1$ (complex symbols), but has the additional advantage that the basis pulses are localized both in time and frequency.

In order to keep the analytical investigation simple, we formulate our transmission system in the discrete time domain and employ a matrix description. The basis pulses in (2) are sampled at Δt and placed in a 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:

$$[\mathbf{G}]_{i,l+kL} = \sqrt{\Delta t}\; g_{l,k}(t)\Big|_{t=\Delta t\, i - 3T_0}. \tag{5}$$

The sampled transmit signal in (1) can then be rewritten by:

$$\mathbf{s} = \mathbf{G}\mathbf{x}, \tag{6}$$

whereas all data symbols $x_{l,k}$ are stacked in the vector $\mathbf{x} \in \mathbb{R}^{LK\times 1}$. For a fair comparison of different transmission techniques we always consider the same average transmit power P_{S}, defined as:

$$P_{\mathrm{S}} = \frac{1}{KT}\,\mathrm{tr}\left(\mathbb{E}\left\{\mathbf{s}\mathbf{s}^H\right\}\right)\Delta t. \tag{7}$$

We also assume a low delay spread and a low Doppler spread [23], so that no relevant inter-symbol and inter-carrier interference occurs. Multiplying the received signal with $\mathbf{G}^H$ (matched filter), provides the received data symbol vector $\mathbf{y} \in \mathbb{R}^{LK\times 1}$ as:

$$\mathbf{y} = \mathrm{diag}\{\mathbf{h}\}\mathbf{D}\mathbf{x} + \mathbf{n}, \tag{8}$$

with

$$\mathbf{D} = \mathbf{G}^H\mathbf{G}. \tag{9}$$

The vector $\mathbf{h} \in \mathbb{C}^{LK\times 1}$ represents the channel and $\mathbf{n}$ the complex Gaussian noise vector, $\mathbf{n} \sim \mathcal{CN}(0,\, P_{\mathrm{n}}\,\mathbf{D})$. Note that in OFDM, $\mathbf{D}$ becomes an identity matrix, $\mathbf{D} = \mathbf{I}_{LK}$, whereas in FBMC this matrix has non-diagonal imaginary elements and only $\Re\{\mathbf{D}\} = \mathbf{I}_{LK}$.

3. PILOT-SYMBOL AIDED CHANNEL ESTIMATION

In pilot-symbol aided channel estimation, special "data" symbols, the so called pilot symbols, are known a priori at the receiver. In OFDM, the channel estimation then becomes a trivial task [24, 25, 26]: the received symbols at the pilot positions are divided by the corresponding data symbols which delivers immediately an estimate of the channel coefficients at the pilot positions, see (8) and the fact that $\mathbf{D}$ is an identity matrix. The channel values at the data positions are then obtained through interpolation or extrapolation. Unfortunately, in FBMC such simple approach does not work due to imaginary interference, that is, the matrix $\mathbf{D}$ consists of non-diagonal imaginary elements. FBMC is based on the idea of taking the real part in order to eliminate the imaginary interference. However, this only works after equalization of the phase shift caused by the channel. Because we do not know this phase shift prior to channel estimation, we have to use the complex domain instead of the real domain. Figure 1 shows the imaginary interference weights surrounding one pilot (data) symbol. Since the interference is purely imaginary valued, taking the real part would completely cancel the interference. However, in the complex domain we obtain a Signal-to-Interference Ratio (SIR) of 0 dB which is clearly too low for an accurate channel estimation. In order to employ pilot-symbol aided channel estimation in FBMC, we thus have to mitigate the imaginary interference.

3.1. Auxiliary Pilot Symbols

Sacrificing one additional data symbol, the so called auxiliary pilot symbol, allows to cancel the imaginary interference at one pilot position. In this paper, we use the matrix representation developed in Section 2 in order to express this cancellation condition in a more general way that captures also the interdependency of closely spaced pilot symbols and an arbitrary number of auxiliary pilot symbols. The imaginary interference at the pilot positions can be completely eliminated if the auxiliary pilot symbols are chosen so that:

$$\mathbf{x}_{\mathcal{P}} = \begin{bmatrix}\mathbf{D}_{\mathcal{P},\mathcal{P}} & \mathbf{D}_{\mathcal{P},\mathcal{D}} & \mathbf{D}_{\mathcal{P},\mathcal{A}}\end{bmatrix}\begin{bmatrix}\mathbf{x}_{\mathcal{P}}\\ \mathbf{x}_{\mathcal{D}}\\ \mathbf{x}_{\mathcal{A}}\end{bmatrix}, \tag{10}$$

which follows from our transmission system model in (8). The vector $\mathbf{x}_{\mathcal{P}} \in \mathbb{R}^{|\mathcal{P}|\times 1}$ denotes all those elements of $\mathbf{x}$ at the pilot positions. The same applies to $\mathbf{x}_{\mathcal{D}} \in \mathbb{R}^{|\mathcal{D}|\times 1}$ at the data positions and $\mathbf{x}_{\mathcal{A}} \in \mathbb{R}^{|\mathcal{A}|\times 1}$ at the auxiliary pilot positions. Similar, the matrix $\mathbf{D}_{\mathcal{P},\mathcal{D}} \in \mathbb{C}^{|\mathcal{P}|\times|\mathcal{D}|}$ consists of the row elements and the column elements of $\mathbf{D}$ at the pilot positions respectively data positions. Again, the same is true for $\mathbf{D}_{\mathcal{P},\mathcal{P}}$ and $\mathbf{D}_{\mathcal{P},\mathcal{A}}$. If the number of auxiliary pilot symbols $|\mathcal{A}|$ is larger than the number of pilot symbols $|\mathcal{P}|$, (10) has infinitely many solutions. Because we want to spend as little energy as possible on auxiliary pilot symbols, we can solve (10) using the Moore-Penrose pseudoinverse [27], leading to:

$$\mathbf{x}_{\mathcal{A}} = \mathbf{D}^{\#}_{\mathcal{P},\mathcal{A}}\left(\mathbf{I}_{\mathcal{P}} - \mathbf{D}_{\mathcal{P},\mathcal{P}}\right)\mathbf{x}_{\mathcal{P}} - \mathbf{D}^{\#}_{\mathcal{P},\mathcal{A}}\mathbf{D}_{\mathcal{P},\mathcal{D}}\mathbf{x}_{\mathcal{D}}, \tag{11}$$

with

$$\mathbf{D}^{\#}_{\mathcal{P},\mathcal{A}} = \mathbf{D}^{H}_{\mathcal{P},\mathcal{A}}\left(\mathbf{D}_{\mathcal{P},\mathcal{A}}\mathbf{D}^{H}_{\mathcal{P},\mathcal{A}}\right)^{-1}. \tag{12}$$

If the pilot symbols are spaced sufficiently far away from each other, $\mathbf{D}_{\mathcal{P},\mathcal{P}}$ becomes an identity matrix, simplifying (11). Let us further consider the case of one auxiliary symbol per pilot symbol so that $\mathbf{D}_{\mathcal{P},\mathcal{A}}$ becomes a diagonal matrix whose elements are given by the interference weight at auxiliary pilot position, that is, 0.4357, see Figure 1. Let us also define the auxiliary pilot power offset $\kappa_{\mathcal{A}}$ as follows:

$$\kappa_{\mathcal{A}} = \frac{P_{\mathcal{A}}}{P_{\mathcal{D}}}, \tag{13}$$

with $P_{\mathcal{D}}$ being the power of the data symbols and $P_{\mathcal{A}}$ the power of the auxiliary pilot symbols which is determined by (11). Suppose we want to cancel $N = 8$ closest interferers, see Figure 1. The power offset then becomes: $(3\cdot 0.4357^2 + 4\cdot 0.2393^2)/0.4357^2 = 4.21$. Thus, the auxiliary pilot power is 4.21 times larger than the data power! Let us now consider the case of two auxiliary pilot symbols

	N=4 ⇒ **SIR=9 dB**†			N=8 ⇒ **SIR=22 dB**†		
	Cod.	1Aux.	2Aux.	Cod.	1Aux.	2Aux.
Aux. power offset	-	3	0.5	-	4.21	0.8
TX multiplications*	0	0	0	2	4	4
TX summations	8	2	1	24	6	5
RX multiplications*	0	none	none	8	none	none
RX summations	9	none	none	25	none	none

*multiplications by -1 and 1/2 are considered as no complexity! †at the pilot positions, for $\kappa_{\mathcal{P}} = 2$

	N=16 ⇒ **SIR=35 dB**†			N=28 ⇒ **SIR=61 dB**†		
	Cod.	1Aux.	2Aux.	Cod.	1Aux.	2Aux.
Aux. power offset	-	4.26	0.82	-	4.27	0.82
TX multiplications*	12	12	12	26	24	24
TX summations	88	14	13	184	26	25
RX multiplications*	32	none	none	80	none	none
RX summations	89	none	none	185	none	none

*multiplications by -1 and 1/2 are considered as no complexity! †at the pilot positions, for $\kappa_{\mathcal{P}} = 2$

Table 1. Additional complexity per pilot symbol for canceling N closest interferers: Coding is more complex than using auxiliary pilot symbols. In particular, auxiliary pilot symbols do not require any additional calculations at the RX. Two auxiliary pilot symbols have approximately the same complexity as one auxiliary pilot symbol.

which split the cancellation job equally between themselves, leading to a power offset of $(2 \cdot 0.4357^2 + 4 \cdot 0.2393^2)/(2 \cdot 0.4357)^2 = 0.8$, so that the auxiliary pilot power is actually lower than the data symbol power. The drawback then is a reduction of the number of data symbols. However, as we will show in Section 4, the overall capacity increases for low to medium SNR values because the saved power offsets the loss of data symbols. The transmitted signal in (6) can be rewritten for the case of auxiliary pilot symbols by:

$$\mathbf{s}_{\mathrm{A}} = \mathbf{G}\mathbf{A}\begin{bmatrix}\mathbf{x}_{\mathcal{P}}\\ \mathbf{x}_{\mathcal{D}}\end{bmatrix}, \tag{14}$$

whereas $\mathbf{A} \in \mathbb{R}^{LK\times(LK-|\mathcal{A}|)}$ represents the auxiliary cancellation conditions in (11). Expressing (7) in terms of the data symbol power $P_{\mathcal{D}}$ for a given average transmit power P_{S} leads to:

$$P_{\mathcal{D}} = \frac{T}{\Delta t}\frac{K}{|\mathcal{A}|\bar{\kappa}_{\mathcal{A}} + |\mathcal{P}|\kappa_{\mathcal{P}} + |\mathcal{D}|}P_{\mathrm{S}}, \tag{15}$$

with $\kappa_{\mathcal{P}}$ denoting the pilot power offset and $\bar{\kappa}_{\mathcal{A}}$ the average auxiliary pilot power offset. Equation (15) will become relevant in Section 4 for performance comparison. Table 1 compares the additional complexity per pilot symbol for different imaginary interference mitigation techniques.

3.2. Coding

Instead of using dedicated auxiliary symbols, we code the data in such a way that the imaginary interference is canceled at the pilot positions. Therefore, we rewrite the transmitted signal as:

$$\mathbf{s}_{\mathrm{C}} = \mathbf{G}\mathbf{C}\begin{bmatrix}\mathbf{x}_{\mathcal{P}}\\ \mathbf{x}_{\mathcal{D}}\end{bmatrix}, \tag{16}$$

with $\mathbf{C} \in \mathbb{R}^{LK\times(LK-|\mathcal{P}|)}$ being the coding matrix. After channel estimation, we then decode the equalized received signal by $\mathbf{C}^T$ whereas the condition $\mathbf{C}^T\mathbf{C} = \mathbf{I}_{(LK-|\mathcal{P}|)\times(LK-|\mathcal{P}|)}$ has to hold, that is, the coding vectors $\mathbf{c}_i$ are orthonormal, so that the transmission system becomes similar to the case without coding. In particular, the noise statistics are the same as for the case without coding. We assume that the pilot symbols are spaced sufficiently far away from each other, allowing us to design the coding independently for different pilot symbols. Thus, the problem of finding $\mathbf{C}$ can be simplified into finding a much smaller coding matrix $\tilde{\mathbf{C}} \in \mathbb{R}^{N\times(N-1)}$ which codes N interfering symbols closest to the pilot, see Figure 1. We follow the approach suggested in [19]: all interference weights

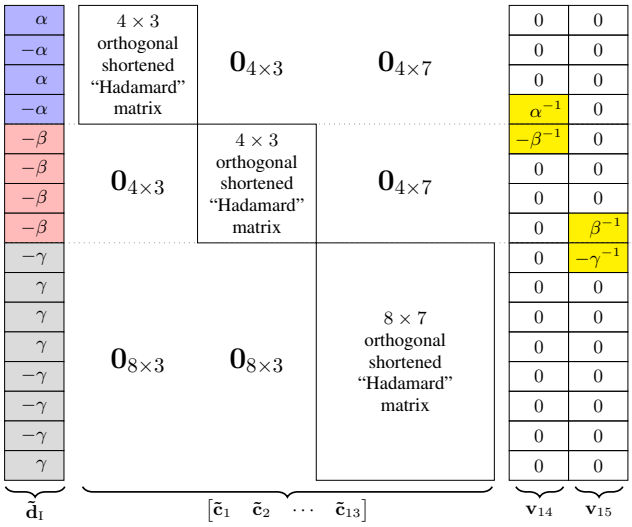

Fig. 2. Illustration of the proposed algorithm for $N = 16$ and $M = 3$ clusters. The key extension to [19, 20] are the vectors $\mathbf{v}_{14}$ and $\mathbf{v}_{15}$ which, in combination with Gram-Schmidt orthogonalization, allow to straightforwardly find the coding matrix $\mathbf{C}$.

with the same magnitude are clustered together and stacked in a vector $\tilde{\mathbf{d}}_{\mathrm{I}} \in \mathbb{R}^{N\times1}$, as shown in Figure 2. We assume there exists M such clusters, each having either 2, 4 or 8 elements. By generating Hadamard matrices, dividing each row by the corresponding interference weight and canceling the column that is not orthogonal to the interference weights, we can find $i = 1 \ldots N - M$ linearly independent coding vectors $\tilde{\mathbf{c}}_i \in \mathbb{R}^{N\times1}$ which are orthogonal to each other and to $\tilde{\mathbf{d}}_{\mathrm{I}}$, see Figure 2. Although the authors in [19] provide general conditions for the remaining $M-1$ coding vectors, they give no detailed instructions on how to construct them. We thus propose the following algorithm:

1. Generate $j = N - M + 1 \ldots N - 1$ vectors $\mathbf{v}_j \in \mathbb{R}^{N\times1}$ which consist of only two nonzero elements, located at the transition between two clusters, and chosen so that $\tilde{\mathbf{d}}_{\mathrm{I}}^T\mathbf{v}_j = 0$, see Figure 2.
2. Use Gram-Schmidt [27] orthogonalization to find the remaining $M - 1$ coding vectors for $j = N - M + 1 \ldots N - 1$:

$$\tilde{\mathbf{c}}_j = \mathbf{v}_j - \sum_{i=1}^{j-1}\frac{\mathbf{v}_j^T\tilde{\mathbf{c}}_i}{\tilde{\mathbf{c}}_i^T\tilde{\mathbf{c}}_i}\tilde{\mathbf{c}}_i. \tag{17}$$

Note that (17) preserves the orthogonality to the interference vector $\tilde{\mathbf{d}}_{\mathrm{I}}$, making such approach feasible. In order to keep the computational complexity low, the vectors $\mathbf{v}_j$ should combine always the clusters with the smallest number of elements for an increasing j. Once two clusters have been combined by the vector $\mathbf{v}_j$, they form a new cluster. Our proposed algorithm describes a general way of finding the coding matrix $\tilde{\mathbf{C}}$ and does not necessarily require orthogonal shortened "Hadamard" matrices, but using them reduces the overall computational complexity. Expressing the data symbol power $P_{\mathcal{D}}$ in terms of the average transmit power P_{S} gives:

$$P_{\mathcal{D}} = \frac{T}{\Delta t}\frac{K}{|\mathcal{P}|\kappa_{\mathcal{P}} + |\mathcal{D}| - |\mathcal{P}|}P_{\mathrm{S}}. \tag{18}$$

In particular, a pilot power offset of $\kappa_{\mathcal{P}} = 2$ results in the same data symbol power $P_{\mathcal{D}}$ as an FBMC system without channel estimation. Additionally, choosing $\kappa_{\mathcal{P}} = 2$ guarantees the same SNR for channel estimation (complex domain) and for data transmission (taking the real part reduces the noise power by a factor of two).

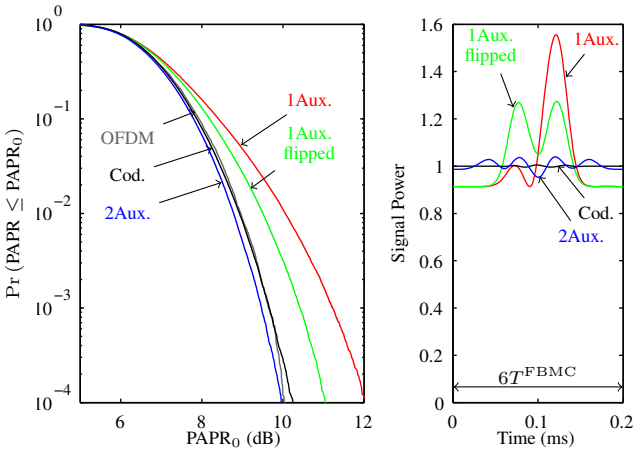

Fig. 3. The large power offset of one auxiliary symbol per pilot results in a high PAPR and generates a high peak in the transmit power. By employing two auxiliary symbols we can mitigate these harmful effects at approximately the same computational complexity.

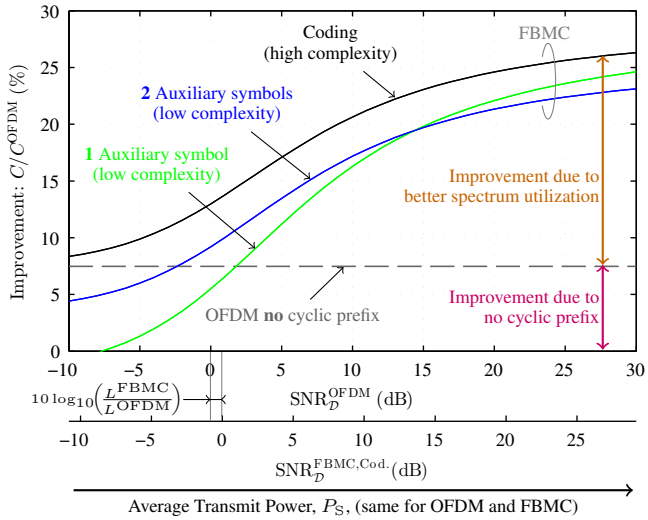

Fig. 4. Capacity improvement of FBMC compared to 1.4 MHz LTE. The maximum improvement of $\frac{15}{14}$(cyclic prefix)$\times\frac{87}{72}$(bandwidth)$\times\frac{0.9556}{0.9524}$(pilot density)$\approx 30\%$ can only be achieved in the limit case of $\mathrm{SNR}\to\infty$. For $\mathrm{SNR}_{\mathcal{D}}^{\mathrm{OFDM}}$ values smaller than 14 dB, two auxiliary symbols per pilot outperform one auxiliary symbol.

4. NUMERICAL RESULTS

FBMC has a higher spectral efficiency than OFDM because it does not employ a Cyclic Prefix (CP) and uses the available bandwidth more efficiently. We will quantify this improvement by comparing FBMC to a 1.4 MHz LTE resembling OFDM signal which uses a subcarrier spacing of $F = 15$ kHz. For this OFDM signal we assume a cycling prefix length of 4.75 µs and $K^{\mathrm{OFDM}} = 14$ OFDM symbols, resulting in a time duration of 1 ms. FBMC on the other hand allows to transmit $K^{\mathrm{FBMC}} = 30$ FBMC symbols within the same 1 ms time interval. However, this symbols are real valued (equivalent to 15 complex symbols). As a reference (dashed line in Figure 4) we also consider an OFDM signal *without* CP that uses $K^{\mathrm{OFDM,noCP}} = 15$ OFDM symbols. Although LTE occupies 1.4 MHz, it only uses $L^{\mathrm{OFDM}} = 72$ subcarriers (1.08 MHz). For FBMC we assume $L^{\mathrm{FBMC}} = 87$ subcarriers so that the power spectral density is below 84 dB of its maximum value for frequencies outside the 1.4 MHz bandwidth. Furthermore, we assume a diamond shaped LTE pilot pattern. OFDM has a data symbol density ($|\mathcal{D}|/LK$) of 0.9524 while OFDM without CP has a density of 0.9556 (they have the same number of pilot symbols). FBMC that uses channel estimation based on coding or one auxiliary pilot symbol has the same density as OFDM without CP, that is, 0.9556, whereas by using two auxiliary symbols the density decreases to 0.9333 because of the decreased number of data symbols.

Figure 3 shows the Peak-to-Average Power Ratio (PAPR), $\max_i |[\mathbf{s}]_i|^2/P_{\mathrm{S}}$, and the transmit power, $\mathrm{diag}\left(\mathbb{E}\left\{\mathbf{s}\mathbf{s}^H\right\}\right)$, for 12 subcarriers, 0.2 ms, and an average signal power of $P_{\mathrm{S}} = 1$. As expected, one auxiliary pilot symbol performs poor due to the power offset. By flipping pilot and auxiliary position at every second pilot subcarrier, we can improve the performance. The performance can further be increased by changing the pilot pattern so that the pilots are evenly distributed in time. However, this requires many subcarriers while two auxiliary symbols and coding still perform better.

Although we could include channel estimation in our capacity considerations [28], we keep the model simple by assuming that only the phase has to be estimated and our channel estimation performs arbitrary close to perfect channel knowledge, delivering a theoretical performance bound of our transmission system. By interpreting each data position as a discrete-time Gaussian channel [29], we find the achievable capacities (which also takes the system overhead such as pilot symbols into account) by:

$$C^{\mathrm{OFDM}} = \frac{|\mathcal{D}|\log_2\left(1+\frac{P_{\mathcal{D}}}{P_{\mathrm{n}}}\right)}{KT} \tag{19}$$

$$C^{\mathrm{FBMC}} = \frac{|\mathcal{D}|\frac{1}{2}\log_2\left(1+\frac{P_{\mathcal{D}}}{P_{\mathrm{n}}/2}\right)}{KT} \tag{20}$$

Because FBMC operates in the real domain, we have to include a factor of $1/2$ in the capacity equation while at the same time the (complex) noise power is reduced by a factor of two, as shown in (20). Note that for the same number of subcarriers L and the same transmit power P_{S}, we have $P_{\mathcal{D}}^{\mathrm{OFDM}} = P_{\mathcal{D}}^{\mathrm{OFDM,noCP}} = 2\,P_{\mathcal{D}}^{\mathrm{FBMC,Cod.}}$, so that $C^{\mathrm{OFDM,noCP}} = C^{\mathrm{FBMC,Cod.}}$ while $C^{\mathrm{OFDM,noCP}} > C^{\mathrm{FBMC,Aux.}}$ because auxiliary pilot symbols are "wasted" energy. Figure 4 shows the improvement in achievable capacity for FBMC relative to 1.4 MHz LTE. For a fair comparison, we keep the average transmit power P_{S} constant.

5. CONCLUSION

If complexity does not constitute an issue, we suggest to use coding for pilot symbol aided channel estimation. Coding mitigates the imaginary interference at pilot positions, offers a low PAPR, no energy is wasted, and well-known channel estimation techniques from OFDM can directly be applied. To reduce the complexity of coding, we suggest to use the algorithm described in Section 3.2. However, if computational complexity becomes relevant, auxiliary pilot symbols might be a better choice. One auxiliary symbol per pilot, as suggested in literature, leads to a severe power offset, increasing the PAPR and wasting too much energy. We thus suggest to use two auxiliary symbols which strongly decreases the PAPR and, in addition, offers a higher capacity for small to medium SNR values because the power that is saved offsets the loss of data symbols.

6. REFERENCES

[1] Alphan Sahin, Ismail Guvenc, and Huseyin Arslan, "A survey on multicarrier communications: Prototype filters, lattice structures, and implementation aspects," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 3, pp. 1312–1338, 2012.

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

[3] Jeffrey G Andrews, Stefano Buzzi, Wan Choi, Stephen V Hanly, Aurelie Lozano, Anthony CK Soong, and Jianzhong Charlie Zhang, "What will 5G be?," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, 2014.

[4] Gerhard Wunder, Peter Jung, Martin Kasparick, Thorsten Wild, Frank Schaich, Yejian Chen, Stephen Brink, Ivan Gaspar, Nicola Michailow, Andreas Festag, et al., "5GNOW: non-orthogonal, asynchronous waveforms for future mobile applications," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 97–105, 2014.

[5] Paolo Banelli, Stefano Buzzi, Giulio Colavolpe, Andrea Modenini, Fredrik Rusek, and Alessandro Ugolini, "Modulation formats and waveforms for 5G networks: Who will be the heir of OFDM?: An overview of alternative modulation schemes for improved spectral efficiency," *IEEE Signal Processing Magazine*, vol. 31, no. 6, pp. 80–93, 2014.

[6] Behrouz Farhang-Boroujeny, "Filter bank multicarrier modulation: A waveform candidate for 5G and beyond," *Advances in Electrical Engineering*, vol. 2014, 2014.

[7] Robert W Chang, "Synthesis of band-limited orthogonal signals for multichannel data transmission," *Bell System Technical Journal*, vol. 45, no. 10, pp. 1775–1796, 1966.

[8] B Saltzberg, "Performance of an efficient parallel data transmission system," *IEEE Transactions on Communication Technology*, vol. 15, no. 6, pp. 805–811, 1967.

[9] Helmut Bölcskei, "Orthogonal frequency division multiplexing based on offset QAM," in *Advances in Gabor analysis*, pp. 321–352. Springer, 2003.

[10] Behrouz Farhang-Boroujeny and Chung Him Yuen, "Cosine modulated and offset QAM filter bank multicarrier techniques: a continuous-time prospect," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, pp. 6, 2010.

[11] M Bellanger, D Le Ruyet, D Roviras, M Terré, J Nossek, L Baltar, Q Bai, D Waldhauser, M Renfors, T Ihalainen, et al., "FBMC physical layer: a primer," *PHYDYAS, January*, 2010.

[12] Chrislin Lélé, J-P Javaudin, Rodolphe Legouable, Alexandre Skrzypczak, and Pierre Siohan, "Channel estimation methods for preamble-based OFDM/OQAM modulations," *European Transactions on Telecommunications*, vol. 19, no. 7, pp. 741–750, 2008.

[13] Dimitrios Katselis, Eleftherios Kofidis, Athanasios Rontogiannis, and Sergios Theodoridis, "Preamble-based channel estimation for CP-OFDM and OFDM/OQAM systems: A comparative study," *IEEE Transactions on Signal Processing*, vol. 58, no. 5, pp. 2911–2916, 2010.

[14] Eleftherios Kofidis, Dimitrios Katselis, Athanasios Rontogiannis, and Sergios Theodoridis, "Preamble-based channel estimation in OFDM/OQAM systems: a review," *Signal Processing*, vol. 93, no. 7, pp. 2038–2054, 2013.

[15] Dejin Kong, Daiming Qu, and Tao Jiang, "Time domain channel estimation for OQAM-OFDM systems: algorithms and performance bounds," *IEEE Transactions on Signal Processing*, vol. 62, no. 2, pp. 322–330, 2014.

[16] Jean-Philippe Javaudin, Dominique Lacroix, and Alexandre Rouxel, "Pilot-aided channel estimation for OFDM/OQAM," in *57th IEEE Vehicular Technology Conference (VTC)*, 2003, vol. 3, pp. 1581–1585.

[17] Tobias Hidalgo Stitz, Tero Ihalainen, Ari Viholainen, and Markku Renfors, "Pilot-based synchronization and equalization in filter bank multicarrier communications," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, pp. 9, 2010.

[18] Jamal Bazzi, Petra Weitkemper, and Katsutoshi Kusume, "Power efficient scattered pilot channel estimation for FBMC/OQAM," in *10th International ITG Conference on Systems, Communications and Coding*. VDE, 2015, pp. 1–6.

[19] C Lélé, R Legouable, and Pierre 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.

[20] W. Cui, D. Qu, T. Jiang, and B. Farhang-Boroujeny, "Coded auxiliary pilots for channel estimation in FBMC-OQAM systems," *IEEE Transactions on Vehicular Technology*, vol. PP, no. 99, pp. 1–1, 2015.

[21] Ralf Haas and Jean-Claude Belfiore, "A time-frequency well-localized pulse for multiple carrier transmission," *Wireless Personal Communications*, vol. 5, no. 1, pp. 1–18, 1997.

[22] Hans G Feichtinger and Thomas Strohmer, *Gabor analysis and algorithms: Theory and applications*, Springer Science & Business Media, 2012.

[23] Franz Hlawatsch and Gerald Matz, *Wireless communications over rapidly time-varying channels*, Academic Press, 2011.

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

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

[26] Ronald Nissel, Martin Lerch, Michal Šimko, and Markus 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.

[27] Todd K Moon and Wynn C Stirling, *Mathematical methods and algorithms for signal processing*, vol. 1, Prentice Hall New York, 2000.

[28] Ronald Nissel, Sebastian Caban, and Markus Rupp, "Closed-Form capacity expression for low complexity BICM with uniform inputs," in *IEEE 26th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Hong Kong, P.R. China, Aug. 2015.

[29] Thomas M Cover and Joy A Thomas, *Elements of information theory*, John Wiley & Sons, 2012.