A Study of Pilot-Aided Channel Estimation in MIMO-GFDM Systems

Shahab Ehsanfar, Maximillian Matthé, Dan Zhang, Gerhard Fettweis, *IEEE Fellow*

Vodafone Chair Mobile Communications Systems
Technische Universität Dresden, Germany
{*firstname.lastname*@ifn.et.tu-dresden.de}

Invited Paper for the Special Session on "Multiple Antenna Concepts for New 5G Air Interface Proposals"

***Abstract*—Generalized frequency division multiplexing (GFDM) is a promising candidate waveform for next generation wireless communications systems. Unlike conventional orthogonal frequency division multiplexing (OFDM) based systems, it is a non-orthogonal waveform subject to inter-carrier and inter-symbol interference. In multiple-input multiple-output (MIMO) systems, the additional inter-antenna interference also takes place. The presence of such three-dimensional interference challenges the receiver design. This paper addresses the MIMO-GFDM channel estimation problem with the aid of known reference signals also referred as pilots. Specifically, the received signal is expressed as the joint effect of the pilot part, unknown data part and noise part. On top of this formulation, least squares (LS) and linear minimum mean square error (LMMSE) estimators are presented, while their performance is evaluated for various pilot arrangements.**

I. Introduction

ULTRA low latency, very high reliability and robustness, low out-of-band (OOB) emission and very high data capacity are among the challenges for the 5th generation (5G) of wireless systems, e.g., [1], [2], [3]. The well-known orthogonal frequency division multiplexing (OFDM) has reached to its boundaries in addressing the above various requirements. Hence, several non-orthogonal waveform candidates have been proposed and rediscovered for the new air interface of 5G, e.g. filter bank multi-carrier (FBMC) [4], universal filtered multi-carrier (UFMC) [5], filtered-OFDM [6] as well as generalized frequency division multiplexing (GFDM) [7].

This paper considers GFDM, since it is equipped with necessary flexibility to address a wide range of requirements envisioned for 5G, e.g., latency, data rates, reliability and OOB emission. Relying on it, a unified air interface can be provided for various service types. The combination of GFDM with multiple antennas, i.e., multiple-input multiple-output (MIMO) GFDM, can further enhance the system performance e.g. [8], [9], [10]. For the MIMO-GFDM receiver design, channel estimation is a critical functional unit. The prior work [11] relied on preamble which is spectrally efficient for continuous transmission over slow fading channels. This paper aims to deliver accurate estimates of channel state information (CSI) for coherent detection by scattered pilot symbols. This type of data-aided channel estimation is more suitable for time and frequency dispersive channels.

In pilot-aided channel estimation, pilot symbols and information-bearing data symbols are multiplexed and transmitted within the same time-frequency resource block, e.g., Fig. 1. At the receiver side, the task of channel estimation is to estimate CSI based on the knowledge of pilot symbols. To this end, different channel estimation techniques have been developed for conventional OFDM systems e.g. [12], [13], [14], [15] and references therein. The extension of OFDM-based channel estimation methods for GFDM are not straightforward, because the orthogonality of OFDM ensures clean pilot observations without interference from unknown data symbols. This property is not valid for GFDM which is a non-orthogonal waveform in general. Moreover, in OFDM many narrow-band subcarriers allow one-tap equalization while on the contrary, in GFDM depending on the transmit signal configuration (e.g. low latency requirement) the subcarriers might have broader bandwidth and consequently, they become frequency selective.

Given the knowledge of data symbols at the transmitter side, it is possible to design pilots such that the interference from data symbols can be properly pre-cancelled. This idea has been applied for channel estimation in a single carrier transmission system over a frequency selective fading channel [16] as well as a GFDM-based system [17]. However, the approach proposed in [17] was developed under the assumption of a nearly flat and slow fading channel, which is unrealistic with respect to broadband communication.

This paper tackles the MIMO-GFDM channel estimation problem for rich multipath fading channels. Two well known estimation techniques, namely least squares (LS) and linear minimum mean square error (LMMSE), are respectively tailored for pilot-aided MIMO-GFDM channel estimation. We evaluate and analyze their performance in accordance with pilot arrangement and correspondingly, we examine their complexity. The LS approach is an unbiased estimator which does not require any probabilistic assumption and therefore, it is being widely used due to its ease of implementation [18]. Nevertheless, the performance loss in LS estimation needs significant attention. On the other hand, LMMSE estimation

is a Bayesian approach which exploits the a-priori knowledge of channel statistics in order to improve the estimation quality at the cost of further implementation complexity.

The rest of this paper is organized as follows: Section II describes the GFDM modulation, pilots insertion and also the assumptions taken into account for the MIMO channel. Section III applies the LS channel estimation method and calculates the closed form expression of the mean squared error (MSE). Based on the computations provided in Sec. III, the LMMSE estimator is then obtained in Section IV. Thereafter, Sec. V compares the numerical results of theoretical calculations with the simulation results; and finally, conclusions are drawn in Sec. VI.

A. Notations

Column-vectors are denoted by vector sign $\vec{X}$ and matrices by boldface $\mathbf{X}$. Time and frequency domain representations are separated by lowercase and uppercase letters respectively. $\mathbb{E}[\cdot]$ is the expectation operator. The trace of a square matrix $\boldsymbol{X}$ is $\mathrm{Tr}(\boldsymbol{X})$. The transpose and Hermitian conjugate of $\boldsymbol{X}$ are $\boldsymbol{X}^T$ and $\boldsymbol{X}^H$ respectively. The Frobenius norm of a matrix $\boldsymbol{X}$ is $\|\boldsymbol{X}\|$ and its square can be written as $\|\boldsymbol{X}\|^2 = \mathrm{Tr}(\boldsymbol{X}\boldsymbol{X}^H)$. The vectorization of a matrix $\boldsymbol{X}$ (i.e. stacking its columns on top of one another from left to right) is denoted by $\mathrm{vec}(\boldsymbol{X})$. The Kronecker and Hadamard products [19] of matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$ are denoted as $\boldsymbol{X}\otimes\boldsymbol{Y}$ and $\boldsymbol{X}\circ\boldsymbol{Y}$ respectively. $\mathrm{diag}(\vec{X})$ is a diagonal matrix whose diagonal entries are the entries of the column vector $\vec{X}$. Furthermore, $\mathrm{diag}(\boldsymbol{X},\cdots,\boldsymbol{Y})$ is a block diagonal matrix according to its matrix entries with $\boldsymbol{X}$ being the top-left and $\boldsymbol{Y}$ being the bottom-right blocks. $\boldsymbol{X} \subseteq \boldsymbol{Y}$ implies that the matrix $\boldsymbol{X}$ is possibly improper submatrix of $\boldsymbol{Y}$. The matrix $\boldsymbol{I}_n$ is the identity matrix of size n. $\vec{0}_n$ is a column vector of size n with all zero entries. $\sqrt{\mathbf{X}}$ is the element-wise square root of matrix $\mathbf{X}$.

II. System Model

A. GFDM Modulation

We assume a GFDM block of length $N = MK$ samples where M complex valued subsymbols are being transmitted on K subcarriers. In GFDM, the entries of vector $\vec{d} \in \mathbb{C}^{N\times 1}$ are filtered through circularly time and frequency shifted versions of a prototype filter $g[n]$. Hence, we define

$$g_{k,m}[n] \triangleq g\left[(n - mK) \bmod N\right] \exp\left[j2\pi \frac{k}{K} n\right], \quad (1)$$

where the circular time shift is acquired via the modulo operation and frequency shift is obtained through the complex exponential term corresponding to subsymbol index m and subcarrier index k respectively.

The superposition of pulse shaped data symbols will then provide the GFDM transmit sample:

$$x[n] = \sum_{k=0}^{K-1} \sum_{m=0}^{M-1} g_{k,m}[n] d_{k,m}, \;\; n = 0, \cdots, N-1 \quad (2)$$

where $d_{k,m}$ is the symbol transmitted on subcarrier k and subsymbol m.

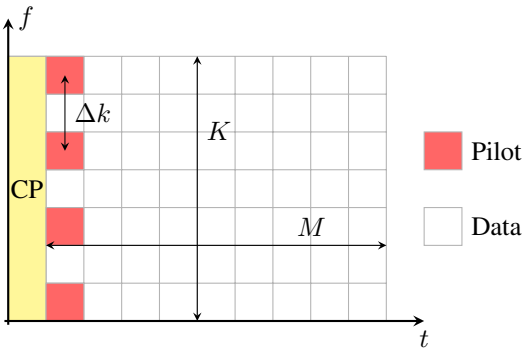


Fig. 1: Pilot positions in time-frequency grid

In terms of matrix and vector notations, the above expression (2) can be rewritten as

$$\vec{x} = \mathbf{A}\vec{d}, \quad (3)$$

where $\vec{d} = (\vec{d}_k[m])^T_{m=0:M-1}$, $\vec{d}_k[m] = (d_{k,m})^T_{k=0:K-1}$ and $\vec{x} = (x[n])^T_{n=0:N-1}$. The GFDM transmit matrix $\mathbf{A}$ follows:

$$\mathbf{A} = \left(\vec{g}_{0,0}, \cdots, \vec{g}_{K-1,0}, \vec{g}_{0,1}, \vec{g}_{1,1}, \cdots, \vec{g}_{K-1,M-1}\right), \quad (4)$$

with column vector $\vec{g}_{k,m} = \left(g_{k,m}[n]\right)^T_{n=0,1,\cdots,N-1}$.

Furthermore, the vector $\vec{d} = \vec{d}_p + \vec{d}_d$ is generated from the summation of pilots sequence $\vec{d}_p \in \mathbb{C}^{N\times 1}$ and data vector $\vec{d}_d \in \mathbb{C}^{N\times 1}$. The pilots sequence $\vec{d}_p$ contains one pilot subsymbol every Δk subcarrier (i.e. $K_p = \lfloor K/\Delta k \rfloor$ pilots) and the rest of subsymbols which are the position of data samples from $\vec{d}_d$ are kept zero. Note that each time-frequency resource element is associated to either pilots or data leading $\vec{d}_p \circ \vec{d}_d = \vec{0}_N$. Fig. 1 shows an example of pilot positions in the time-frequency grid. It is plain that the pilot insertion with a small Δk, reduces the effective rate by the factor

$$\begin{aligned} \eta &\triangleq \frac{\text{no. of data samples}}{\text{no. of total samples}} = (N - K_p)/N \\ &= 1 - \frac{\lfloor K/\Delta k \rfloor}{MK}. \end{aligned} \quad (5)$$

B. MIMO Wireless Channel

Consider a multi-path MIMO block fading channel with $\mathsf{n_T}$ transmit and $\mathsf{n_R}$ receive antennas where the whole GFDM block at each Tx antenna is protected by a single cyclic prefix (CP). Due to the CP, the receive signal $\vec{y}_{i_\mathsf{R}}$ (at Rx antenna i_R) in time is the circular convolution of transmit signal $\vec{x}_{i_\mathsf{T}}$ (from Tx antenna i_T) and the channel impulse response $\vec{h}_{i_\mathsf{T},i_\mathsf{R}}$ (between the antennas i_T and i_R) plus the AWGN process $\vec{w}_{i_\mathsf{R}}$

$$\vec{y}_{i_\mathsf{R}} = \sum_{i_\mathsf{T}=1}^{\mathsf{n_T}} \vec{x}_{i_\mathsf{T}} \circledast \vec{h}_{i_\mathsf{T},i_\mathsf{R}} + \vec{w}_{i_\mathsf{R}}. \quad (6)$$

In the above expression, it is assumed that all the channels have shorter lengths L compared to the CP length. Moreover,

the channel impulse response between the antennas i_{T} and i_{R} is defined as

$$\vec{h}_{i_{\mathsf{T}}i_{\mathsf{R}}} \triangleq \sqrt{\mathrm{diag}(\vec{P}_{i_{\mathsf{T}}i_{\mathsf{R}}})}\vec{\mathfrak{g}}_{i_{\mathsf{T}}i_{\mathsf{R}}}, \tag{7}$$

where $\vec{P}_{i_{\mathsf{T}}i_{\mathsf{R}}} \in \mathbb{R}^{L\times 1}$ is the normalized power delay profile (PDP) between the Tx antenna i_{T} and Rx antenna i_{R}; and $\vec{\mathfrak{g}}_{i_{\mathsf{T}}i_{\mathsf{R}}} \in \mathbb{C}^{L\times 1}$ is a vector of zero mean complex Gaussian random variables with unit variance, representing independent Rayleigh fading for different Tx-Rx antenna pairs.

Due to the circular convolution in (6), the individual channels are diagonal in frequency domain and therefore, the observed signal on Rx antenna i_{R} is characterized by the following linear equation:

$$\vec{Y}'_{i_{\mathsf{R}}} = \sum_{i_{\mathsf{T}}=1}^{\mathsf{n_T}} (\boldsymbol{X}'_{p,i_{\mathsf{T}}} + \mathbf{X}'_{d,i_{\mathsf{T}}})\vec{H}'_{i_{\mathsf{T}},i_{\mathsf{R}}} + \vec{W}'_{i_{\mathsf{R}}}, \tag{8}$$

with $\vec{H}'_{i_{\mathsf{T}},i_{\mathsf{R}}} = \mathbf{F}'_L\vec{h}_{i_{\mathsf{T}},i_{\mathsf{R}}}$. Furthermore, $\boldsymbol{X}'_{s,i_{\mathsf{T}}} = \mathrm{diag}(\vec{X}'_{s,i_{\mathsf{T}}})$ is a diagonal matrix associated either to pilots p or data sequences d (i.e. $s \in \{p,d\}$). $\vec{X}'_{s,i_{\mathsf{T}}}$ is being transmitted on Tx antenna i_{T} and it is defined as $\vec{X}'_{s,i_{\mathsf{T}}} \triangleq (\mathbf{F}'_t\mathbf{A}\vec{d}_s)_{i_{\mathsf{T}}}$. $\mathbf{F}'_t \in \mathbb{C}^{N\times N}$ is the DFT matrix and $\mathbf{F}'_L \in \mathbb{C}^{N\times L}$ contains only the first L columns of $\mathbf{F}'_t$ where L is the channel length. $\vec{W}'_{i_{\mathsf{R}}}$ is the frequency domain counterpart of AWGN process on receive antenna i_{R}.

If the number of pilot subcarriers is smaller than the number of data subcarriers, i.e., the subcarrier spacing $\Delta k > 1$, only a subset of observations in frequency domain with $N_p = \lfloor N/\Delta k \rfloor$ samples that contain the information of pilots will be used for pilot-aided channel estimation. In equations, the received signal at pilot-bearing subcarriers follows:

$$\vec{Y}_{i_{\mathsf{R}}} = \sum_{i_{\mathsf{T}}=1}^{\mathsf{n_T}} (\boldsymbol{X}_{p,i_{\mathsf{T}}} + \mathbf{X}_{d,i_{\mathsf{T}}})\vec{H}_{i_{\mathsf{T}},i_{\mathsf{R}}} + \vec{W}_{i_{\mathsf{R}}}, \tag{9}$$

where $\vec{H}_{i_{\mathsf{T}},i_{\mathsf{R}}} = \mathbf{F}_L\vec{h}_{i_{\mathsf{T}},i_{\mathsf{R}}}$, $\vec{W}_{i_{\mathsf{R}}} = \mathbf{F}_t\vec{w}_{i_{\mathsf{R}}}$, $\boldsymbol{X}_{s,i_{\mathsf{T}}} = \mathrm{diag}(\vec{X}_{s,i_{\mathsf{T}}})$ and $\vec{X}_{s,i_{\mathsf{T}}} = (\mathbf{F}_t\mathbf{A}\vec{d}_s)_{i_{\mathsf{T}}}$. Here, $\mathbf{F}_t \subseteq \mathbf{F}'_t$ and $\mathbf{F}_L \subseteq \mathbf{F}'_L$ are $N_p \times N$ and $N_p \times L$ matrices that take the DFT at pilot subcarriers respectively i.e. every $m + kM$ row of $\mathbf{F}_t, \mathbf{F}_L$ corresponds to $m + kM\Delta k$ row of $\mathbf{F}'_t, \mathbf{F}'_L$ respectively.

We also define the ratio $\xi \triangleq \frac{\mathcal{O}}{\mathcal{U}}$ where $\mathcal{O} = M \times K_p \times \mathsf{n_R}$ is the number of observations (as each pilot generates M samples in frequency domain) while $\mathcal{U} = L \times \mathsf{n_T} \times \mathsf{n_R}$ is the total number of channel taps to be estimated. An LS estimate of the channel exists if and only if the number of observations is greater than or equal to the number of estimation parameters i.e. $\xi \geq 1$. Although, LMMSE technique can still provide a reasonable estimate of the channel due to its prior knowledge of statistical CSI.

We rearrange the expression (9) into matrix form as

$$\mathbf{Y} = (\boldsymbol{X}_p + \boldsymbol{X}_d)\mathbf{F}\mathbf{h} + \mathbf{W}, \text{ with } \begin{cases} \mathbf{Y}, \mathbf{W} \in \mathbb{C}^{N_p\times \mathsf{n_R}} \\ \boldsymbol{X}_p, \boldsymbol{X}_d \in \mathbb{C}^{N_p\times N_p\mathsf{n_T}} \\ \mathbf{F} \in \mathbb{C}^{N_p\mathsf{n_T}\times L\mathsf{n_T}} \\ \mathbf{h} \in \mathbb{C}^{L\mathsf{n_T}\times \mathsf{n_R}} \end{cases} \tag{10}$$

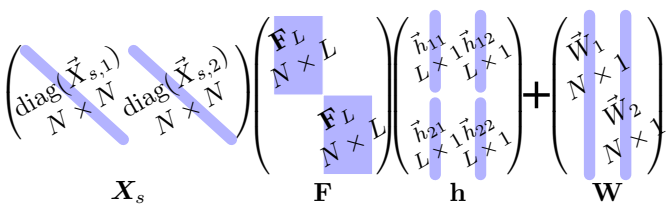


Fig. 2: Overview of the matrix sturctures for a 2×2 MIMO channel

herein, each of the above parameters are defined as

$$\mathbf{Y} \triangleq (\vec{Y}_1, \cdots, \vec{Y}_{i_{\mathsf{R}}}, \cdots, \vec{Y}_{\mathsf{n_R}}), \tag{10a}$$

$$\boldsymbol{X}_s \triangleq (\boldsymbol{X}_{s,1}, \cdots, \boldsymbol{X}_{s,i_{\mathsf{T}}}, \cdots, \boldsymbol{X}_{s,\mathsf{n_T}}), \tag{10b}$$

$$\mathbf{F} \triangleq \mathbf{I}_{\mathsf{n_T}} \otimes \mathbf{F}_L, \tag{10c}$$

$$\mathbf{h} \triangleq \begin{pmatrix} \vec{h}_{11} & \cdots & \vec{h}_{1\mathsf{n_R}} \\ \vdots & \ddots & \vdots \\ \vec{h}_{\mathsf{n_T}1} & \cdots & \vec{h}_{\mathsf{n_T}\mathsf{n_R}} \end{pmatrix}, \tag{10d}$$

$$\mathbf{W} \triangleq (\vec{W}_1, \cdots, \vec{W}_{i_{\mathsf{R}}}, \cdots, \vec{W}_{\mathsf{n_R}}). \tag{10e}$$

Eq. (10) depicts that the observed matrix $\mathbf{Y}$ contains a deterministic term $\boldsymbol{X}_p\mathbf{F}\mathbf{h}$, an interference term due to useful information $\boldsymbol{X}_d\mathbf{F}\mathbf{h}$ and the WGN $\mathbf{W}$. Moreover, Fig. 2 shows an example of matrix structures for $\mathsf{n_T} = 2$ by $\mathsf{n_R} = 2$ antennas. In Fig. 2 it is illustrated that $\boldsymbol{X}_s$ is a wide matrix composed of individual diagonal matrices of transmit signals associated to different Tx antennas. Furthermore, the matrix of channel impulse responses $\mathbf{h}$ is structured as $\mathsf{n_T} \times \mathsf{n_R}$ column vectors. Such matrix structure brings an advantage for mathematical analysis when vectorizing the channel matrix. It is trivial from Fig. 2 that $\vec{h} = \mathrm{vec}(\mathbf{h})$ will consist of $\mathsf{n_T}\mathsf{n_R} = 4$ independent column vectors of channel impulse responses, and thus, considering Rayleigh fading channels with no spatial correlation the covariance matrix of all channel impulse responses $\mathbb{E}\left[\vec{h}\vec{h}^H\right]$ becomes diagonal.

Resorting to the matrix identity $\mathrm{vec}(\mathsf{ABC}) = (\mathsf{C}^T \otimes \mathsf{A})\mathrm{vec}(\mathsf{B})$ [19], the corresponding vectorization of the observed matrix $\mathbf{Y}$ yields the following equation:

$$\vec{Y} = \mathrm{vec}(\mathbf{Y}) = \widetilde{\mathbf{x}}\vec{h} + \vec{W}, \tag{11}$$

where $\widetilde{\mathbf{x}} = (\mathbf{I}_{\mathsf{n_R}} \otimes \boldsymbol{X}\mathbf{F}) \in \mathbb{C}^{N_p\mathsf{n_R}\times L\mathsf{n_T}\mathsf{n_R}}$, $\boldsymbol{X} = \boldsymbol{X}_p + \boldsymbol{X}_d$, $\vec{h} = \mathrm{vec}(\mathbf{h})$ and $\vec{W} = \mathrm{vec}(\mathbf{W})$.

III. Least Squares Estimation

The structure of the transmit signal matrix $\boldsymbol{X}_s$ does not allow to provide a least squares estimate of the channel in frequency domain. As mentioned in Sec. II-B, $\boldsymbol{X}_s$ is a wide matrix of diagonal matrices and therefore, the product of $\boldsymbol{X}_s^H\boldsymbol{X}_s$ or specifically $\boldsymbol{X}_p^H\boldsymbol{X}_p$ is always singular. Although, one can obtain the LS estimate of the channel impulse response by minimizing $\|\mathbf{Y} - \boldsymbol{X}_p\mathbf{F}\mathbf{h}\|^2$ with respect to $\mathbf{h}$. This yields

$$\hat{\mathbf{h}}_{\mathrm{LS}} = \boldsymbol{Q}_{\mathrm{LS}}\mathbf{Y} = \mathbf{h} + \mathbf{E}, \tag{12}$$

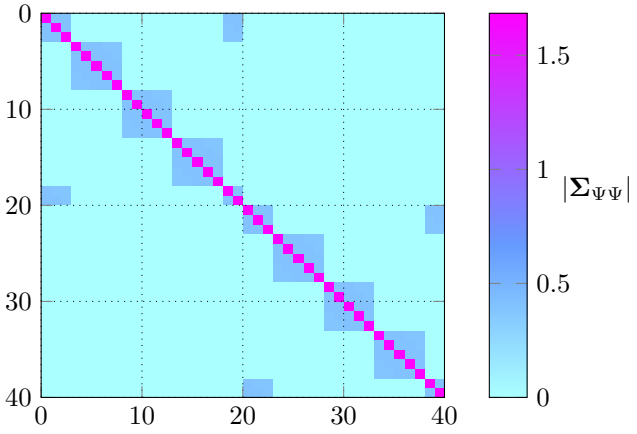


Fig. 3: Example of $\mathbf{\Sigma}_{\Psi\Psi}$ structure for a 2×2 MIMO channel with $K = 16$, $M = 5$ and $\Delta k = 4$

where $\boldsymbol{Q}_{\text{LS}} = \left((\mathbf{X_p F})^H (\mathbf{X_p F})\right)^{-1} (\mathbf{X_p F})^H \in \mathbb{C}^{L\mathsf{n_T} \times N_p}$. The channel frequency response at all data subcarriers is then calculated through Fourier transform of individual estimated channel impulse responses i.e.

$$\widehat{\mathbf{H}}_{\text{LS}} = (\mathbf{I}_{\mathsf{n_T}} \otimes \mathbf{F}'_L)\hat{\mathbf{h}}_{\text{LS}} \tag{13}$$

The above estimation yields the following interference and noise terms:

$$\mathbf{E} = \boldsymbol{Q}_{\text{LS}}\mathbf{\Psi} + \boldsymbol{Q}_{\text{LS}}\mathbf{W}. \tag{14}$$

Here, $\mathbf{\Psi} = \boldsymbol{X}_d\mathbf{F}\mathbf{h}$ leads to an error floor due to the confront of the pilots and useful information and therefore, the norm $\|\boldsymbol{Q}_{\text{LS}}\|$ not only enhances the noise term but also the interference from data.

Accordingly, the result of the MSE calculation follows:

$$\begin{aligned}\text{MSE}_{\text{LS}} &= \mathbb{E}\left[\|\hat{\mathbf{h}}_{\text{LS}} - \mathbf{h}\|^2\right] \\ &= \frac{1}{\Delta k}\text{Tr}\left(\left(\mathbf{I}_{\mathsf{n_R}} \otimes (\boldsymbol{Q}_{\text{LS}}^H \boldsymbol{Q}_{\text{LS}})\right)\mathbf{\Sigma}_{\Psi\Psi}\right) \\ &\quad + \frac{\sigma_w^2}{\Delta k}\text{Tr}\left(\mathbf{I}_{\mathsf{n_R}} \otimes (\boldsymbol{Q}_{\text{LS}}^H \boldsymbol{Q}_{\text{LS}})\right), \end{aligned} \tag{15}$$

where σ_w^2 is the noise variance. Then, we compute the covariance matrix of the interference term as

$$\begin{aligned}\mathbf{\Sigma}_{\Psi\Psi} &= \mathbb{E}\left[\text{vec}(\boldsymbol{X}_d\mathbf{F}\mathbf{h})\text{vec}(\boldsymbol{X}_d\mathbf{F}\mathbf{h})^H\right] \\ &= \mathbb{E}_{\boldsymbol{X}_d}\left[(\mathbf{I}_{\mathsf{n_R}} \otimes \boldsymbol{X}_d\mathbf{F})\,\mathbb{E}_{\mathbf{h}}\left[\vec{h}\vec{h}^H|\boldsymbol{X}_d\right](\mathbf{I}_{\mathsf{n_R}} \otimes \boldsymbol{X}_d\mathbf{F})^H\right] \\ &= \mathbb{E}_{\boldsymbol{X}_d}\left[(\mathbf{I}_{\mathsf{n_R}} \otimes \boldsymbol{X}_d\mathbf{F})\,\mathbf{\Sigma_{hh}}\,(\mathbf{I}_{\mathsf{n_R}} \otimes \boldsymbol{X}_d\mathbf{F})^H\right]. \end{aligned} \tag{16}$$

Here, an important fact arises that both of the above matrices $(\mathbf{I}_{\mathsf{n_R}} \otimes \boldsymbol{X}_d\mathbf{F})$ and $\mathbf{\Sigma_{hh}}$ have block diagonal structures as

$$\begin{aligned}\mathbf{I}_{\mathsf{n_R}} \otimes \boldsymbol{X}_d\mathbf{F} = \text{diag}([\boldsymbol{X}_{d,1}\mathbf{F}_L, \cdots, \boldsymbol{X}_{d,\mathsf{n_T}}\mathbf{F}_L], \cdots, \\ [\boldsymbol{X}_{d,1}\mathbf{F}_L, \cdots, \boldsymbol{X}_{d,\mathsf{n_T}}\mathbf{F}_L]), \end{aligned} \tag{17}$$

$$\begin{aligned}\mathbf{\Sigma_{hh}} = \text{diag}(\mathbf{\Sigma}_{h_{11}}, \cdots, \mathbf{\Sigma}_{h_{\mathsf{n_T}1}}, \\ \cdots, \mathbf{\Sigma}_{h_{(\mathsf{n_T}-1)\mathsf{n_R}}}, \mathbf{\Sigma}_{h_{\mathsf{n_T}\mathsf{n_R}}}), \end{aligned} \tag{18}$$

where $\mathbf{\Sigma}_{h_{i_\mathsf{T} i_\mathsf{R}}} \in \mathbb{R}^{L\times L}$ is the diagonal covariance matrix of channel impulse response, computed as

$$\begin{aligned}\mathbf{\Sigma}_{h_{i_\mathsf{T} i_\mathsf{R}}} &= \mathbb{E}\left[\vec{h}_{i_\mathsf{T} i_\mathsf{R}}\vec{h}^H_{i_\mathsf{T} i_\mathsf{R}}\right] \\ &= \text{diag}(\vec{P}_{i_\mathsf{T} i_\mathsf{R}}). \end{aligned} \tag{19}$$

The product of (17), (18) and the hermitian conjugate of (17) will then provide a block diagonal structure for the interference covariance matrix $\mathbf{\Sigma}_{\Psi\Psi}$ as expressed in (16). This is due to the fact that independent Rayleigh fading has been considered for the individual channels (see Sec. II-B). As a result, it is possible to perform the computations separately for the individual blocks. Hence, for the Tx antenna i_T and Rx antenna i_R we have [19]:

$$\begin{aligned}\mathbf{\Sigma}_{\Psi\Psi_{i_\mathsf{T} i_\mathsf{R}}} &= \mathbb{E}_{\boldsymbol{X}_{d,i_\mathsf{T}}}\left[\boldsymbol{X}_{d,i_\mathsf{T}}\mathbf{F}_L\mathbb{E}_h\left[\vec{h}_{i_\mathsf{T} i_\mathsf{R}}\vec{h}^H_{i_\mathsf{T} i_\mathsf{R}}|\boldsymbol{X}_{d,i_\mathsf{T}}\right]\mathbf{F}_L^H\boldsymbol{X}_{d,i_\mathsf{T}}^H\right] \\ &= \mathbf{\Upsilon}_{i_\mathsf{T} i_\mathsf{R}} \circ \mathbf{\Sigma}_{X_dX_d,i_\mathsf{T}}, \end{aligned} \tag{20}$$

where $\mathbf{\Upsilon}_{i_\mathsf{T} i_\mathsf{R}} = \mathbf{F}_L\text{diag}(\vec{P}_{i_\mathsf{T} i_\mathsf{R}})\mathbf{F}_L^H$. Furthermore, the covariance matrix of data is being calculated as

$$\begin{aligned}\mathbf{\Sigma}_{X_dX_d,i_\mathsf{T}} &= \mathbb{E}[(\mathbf{F}_t\mathbf{A}\vec{d}_d)_{i_\mathsf{T}}(\mathbf{F}_t\mathbf{A}\vec{d}_d)^H_{i_\mathsf{T}}] \\ &= (\mathbf{F}_t\mathbf{A}\text{diag}(\vec{\sigma}_d^2)\mathbf{A}^H\mathbf{F}_t^H)_{i_\mathsf{T}}, \end{aligned} \tag{21}$$

where, $\vec{\sigma}_d^2$ is the vector of data variances with zero entries at pilot positions.

Consequently, for each Rx antenna i_R we calculate the individual diagonal blocks of $\mathbf{\Sigma}_{\Psi\Psi}$ as

$$\mathbf{\Sigma}_{\Psi\Psi(i_\mathsf{R})} = \sum_{i_\mathsf{T}=1}^{\mathsf{n_T}} \mathbf{\Upsilon}_{i_\mathsf{T} i_\mathsf{R}} \circ \mathbf{\Sigma}_{X_dX_d,i_\mathsf{T}}. \tag{22}$$

Hence, the full interference covariance matrix follows:

$$\mathbf{\Sigma}_{\Psi\Psi} = \text{diag}(\mathbf{\Sigma}_{\Psi\Psi(i_\mathsf{R}=1)}, \mathbf{\Sigma}_{\Psi\Psi(i_\mathsf{R}=2)}, \cdots, \mathbf{\Sigma}_{\Psi\Psi(i_\mathsf{R}=\mathsf{n_R})}). \tag{23}$$

Fig. 3 shows an example of $\mathbf{\Sigma}_{\Psi\Psi}$ for a 2×2 MIMO channel. Notice that the individual blocks of $\mathbf{\Sigma}_{\Psi\Psi}$ are sparse matrices with most elements equal to zero.

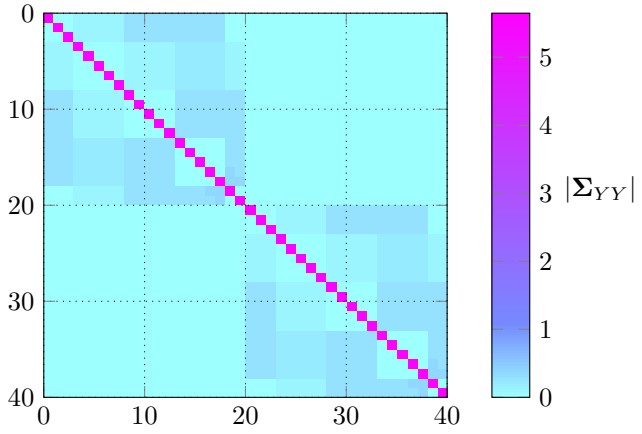


Fig. 4: Example of $\mathbf{\Sigma}_{YY}$ structure for a 2×2 MIMO channel with $K = 16$, $M = 5$ and $\Delta k = 4$

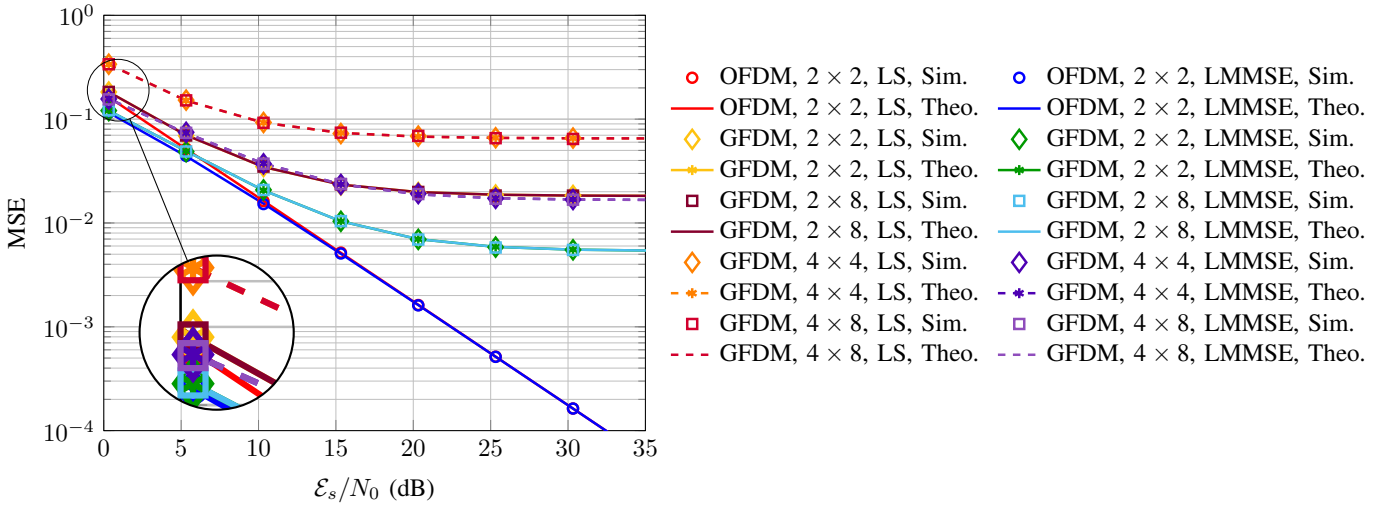


Fig. 5: MSE results of channel estimation vs. SNR for simulation and theoretical calculations in Rayleigh fading MIMO channel with pilot spacing of $\Delta k = 2$ and $K = 128$ subcarriers.

IV. LMMSE

The LMMSE estimation calculates the coefficients of a linear filter aiming at minimum mean square error. In accordance with (10) and the corresponding vectorization in (11), we formally have:

$$\hat{\vec{h}}_{\text{LMMSE}} = \mathbf{\Sigma}_{hY}\mathbf{\Sigma}_{YY}^{-1}\vec{Y}, \tag{24}$$

with the matrices defined as

$$\mathbf{\Sigma}_{YY} = \widetilde{\mathbf{x}}_p\mathbf{\Sigma}_{\mathbf{hh}}\widetilde{\mathbf{x}}_p^H + \mathbf{\Sigma}_{\Psi\Psi} + \sigma_w^2\mathbf{I}_{N\mathsf{n}_\mathsf{R}}, \in \mathbb{C}^{N_p\mathsf{n}_\mathsf{R}\times N_p\mathsf{n}_\mathsf{R}} \tag{25}$$

$$\mathbf{\Sigma}_{hY} = \mathbf{\Sigma}_{\mathbf{hh}}\widetilde{\mathbf{x}}_p^H, \in \mathbb{C}^{L\mathsf{n}_\mathsf{T}\mathsf{n}_\mathsf{R}\times N_p\mathsf{n}_\mathsf{R}} \tag{26}$$

where $\widetilde{\mathbf{x}}_p = (\mathbf{I}_{\mathsf{n}_\mathsf{R}} \otimes \boldsymbol{X}_p\mathbf{F})$. Note that $\hat{\vec{h}}_{\text{LMMSE}}$ is a column vector containing $\mathsf{n}_\mathsf{T}\mathsf{n}_\mathsf{R}$ individual column vectors of size L, associated to the LMMSE estimates of the individual channel impulse responses.

The resulting MSE performance of the LMMSE estimation follows:

$$\text{MSE}_{\text{LMMSE}} = \text{Tr}\left(\mathbf{\Sigma}_{\mathbf{HH}} - \widehat{\mathbf{\Sigma}}_{\mathbf{HH}}\right) \tag{27}$$

with

$$\mathbf{\Sigma}_{\mathbf{HH}} = (\mathbf{I}_{\mathsf{n}_\mathsf{R}} \otimes \mathbf{F})\mathbf{\Sigma}_{\mathbf{hh}}(\mathbf{I}_{\mathsf{n}_\mathsf{R}} \otimes \mathbf{F})^H \tag{28}$$

$$\widehat{\mathbf{\Sigma}}_{\mathbf{HH}} = (\mathbf{I}_{\mathsf{n}_\mathsf{R}} \otimes \mathbf{F})\mathbf{\Sigma}_{hY}\mathbf{\Sigma}_{YY}^{-1}\mathbf{\Sigma}_{hY}^H(\mathbf{I}_{\mathsf{n}_\mathsf{R}} \otimes \mathbf{F})^H \tag{29}$$

where $\mathbf{F}$ is a block diagonal Fourier matrix defined in (10c).

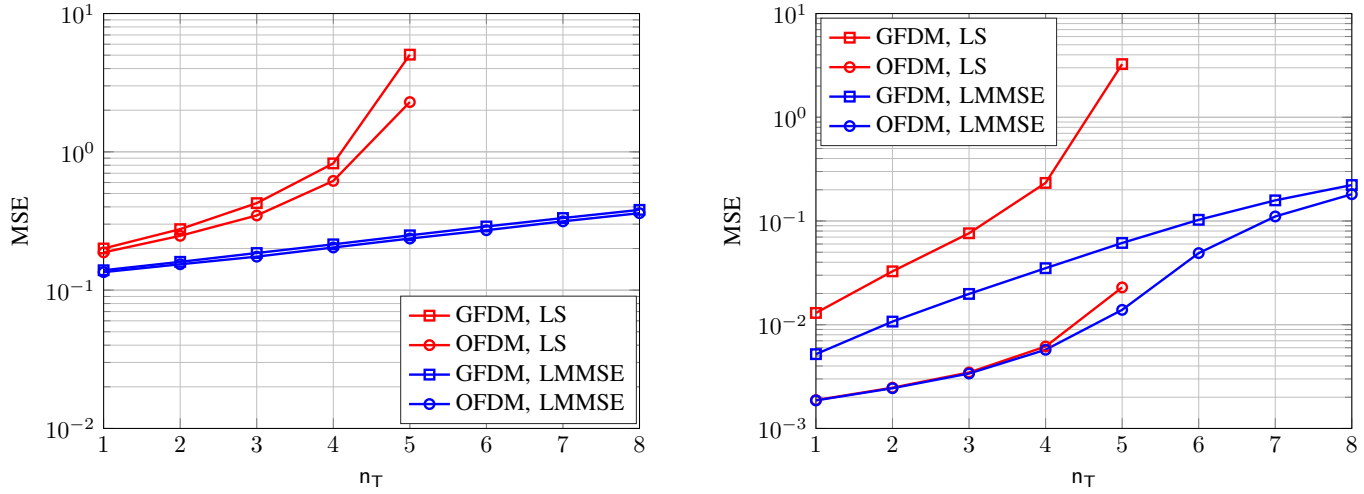
A. Complexity

The complexity of the LMMSE implementation in GFDM is increased with respect to the LMMSE estimation in OFDM due to the further computations of the interference covariance matrix (23). However, if the PDP, the configuration of transmitter matrix $\mathbf{A}$ and the pilot pattern (i.e. the position of pilots in time-frequency grid) remain unchanged over consecutive transmissions, the computation of (23) is required only once. In addition, note that although the complexity of the matrix inversion $\mathbf{\Sigma}_{YY}^{-1}$ is $O\left((N_p\mathsf{n}_\mathsf{R})^3\right)$, the matrix $\mathbf{\Sigma}_{YY}$ is constructed in form of block diagonal with n_R blocks. Hence, the complexity reduces to $O(\mathsf{n}_\mathsf{R}N_p^3)$ due to n_R separate matrix inversions. Fig. 4 shows an example of $\mathbf{\Sigma}_{YY}$ for a 2×2 MIMO channel and thus 2 individual diagonal blocks which can be inverted separately.

V. Numerical Results

In this section we present the simulation and numerical results in order to validate the closed-form expressions of mean squared error of MIMO GFDM channel estimation proposed in the previous sections. Moreover, we compare the performance of the GFDM channel estimation with OFDM. In the end, we show the simulation results of symbol error rate (SER) performance of GFDM under the impact of channel estimation where we compare it with OFDM along with GFDM Genie-aided receiver.

Here, we consider a sequence of 16-QAM symbols with energy per symbol $\mathcal{E}_s$ being transmitted through a multipath MIMO channel with noise energy N_0 and with $\mathsf{n}_\mathsf{T} = \{2, 3, 4\}$ and $\mathsf{n}_\mathsf{R} = \{2, 3, \cdots, 8\}$ antennas. The channel gains are considered to have Rayleigh distribution, hence, the PDP is exponentially distributed with $L = 9$ independent fading gains. A single block of GFDM signal contains $M = 7$ subsymbols and it is filtered by a Raised Cosine pulse with roll-off factor $\alpha = 0.3$. For comparison purpose, we configure OFDM to have $K' = MK$ subcarriers. Considering the same block length of GFDM and OFDM, their bandwidth becomes equivalent. Note that the subcarrier spacing of GFDM is M times wider than the case of OFDM. Therefore, the number of subcarriers in GFDM is M times less than that of OFDM, while each subcarrier carries M data symbols.



(a) $\mathcal{E}_s/N_0 = 0$ dB

(b) $\mathcal{E}_s/N_0 = 20$ dB

Fig. 6: MSE vs. no. of Tx-Rx antennas with $K = 96, \Delta k = 2, L = 9, \mathsf{n_R} = \mathsf{n_T}$

Fig. 5 illustrates the MSE evaluations for theoretical analysis as well as simulation results via Monte Carlo method. It is clearly observable that the channel estimation for GFDM contains an error floor due to the interference from data symbols while for OFDM, MSE decreases linearly with SNR. Moreover, comparing the GFDM channel estimation results for various number of Tx and Rx antennas, we notice that the error does not directly depend to the number of receive antennas e.g. The MSE curves for 2×2 vs. 2×8 antennas are overlapped (as well as 4×4 vs. 4×8). This is due to the fact that, by linearly increasing the number of Rx antennas we increase the number of observations while the number of estimation parameters (i.e. channel taps) also increases linearly e.g. Doubling the number of Rx antennas, we also double the number of channels while the ratio $\xi = \frac{\mathcal{O}}{U}$ remains identical. As a consequence, no analytical difference should be expected for this case.

On the other hand, as we increase the number of Tx-Rx antennas, the estimation performance for both LS and LMMSE estimators degrades (see Fig. 6), because, by linearly increasing the number of Tx-Rx antennas, the number of channel taps is increased quadratically and thus, the parameter ξ decreases leading to estimation performance degradation. Note that for the specified combination of $K, \Delta k$ & L in Fig. 6, the MSE curves for LS estimation with $\mathsf{n_T} > 5$ do not exist for OFDM, since the parameter ξ becomes smaller than one. Furthermore, the error gets too large also for GFDM as $\xi > 1/M$ because the norm of the LS estimator becomes larger and consequently enhances the interference term significantly. In addition, Fig. 6a and 6b also show that in low SNR case, the performance of GFDM and OFDM is almost the same, though, in high SNR the gap between GFDM and OFDM increases due to the interference from data and correspondingly the error floor in GFDM channel estimation.

In order to assess the influence of pilot spacing Δk in MIMO GFDM channel estimation, Fig. 7 compares the mean squared error for different Δk values. The pilot spacing of $\Delta k = 2$ in GFDM corresponds to $\Delta k' = 2M$ in OFDM as one GFDM block contains $MK = 896$ samples with $K_p = 64$ pilots while also OFDM includes the same number of samples and pilots. An immediate observation from Fig. 7 is the overlapping of channel estimation MSE in OFDM system with the MSE in GFDM due to noise only. However, because of the interference in MIMO GFDM channel estimation, the summation of MSE due to noise and due to interference in GFDM becomes larger than the MSE of OFDM. Furthermore, it can be seen that a smaller value of Δk leads better estimation results for both LS and LMMSE methods, because, more number of pilots are transmitted and more observations are acquired at the receiver side. Note that the error floor due to the interference that is independent of SNR, also increases vertically with the increase of Δk. Nevertheless, the effective rate for $\Delta k = \{1, 2, 4\}$ is reduced by $\eta = \{86\%,\ 93\%,\ 96\%\}$ respectively.

The simulation results for uncoded SER performance of GFDM under the influence of channel estimation is provided in Fig. 8. Here, we compare the MIMO GFDM channel estimation with OFDM as well as GFDM Genie-aided receiver. The GFDM Genie-aided receiver transmits no pilot symbols (only data transmission) but it has the perfect knowledge of the CSI due to a genie at no cost. Here again, we notice that the GFDM SER performance has an error floor at high SNRs due to the error floor in channel estimation, while for OFDM SER decreases proportionally with the SNR. Moreover, notice that by increasing the pilot spacing Δk (Fig. 8 left to right) not only the SER due to the noise term increases, but also the error floor which is due to the interference term from data in channel estimation. Fewer pilot subcarriers makes the Fourier matrix wider, resulting in a larger value of $\|\boldsymbol{Q}_{\mathrm{LS}}\|$ and therefore, further noise and interference enhancement.

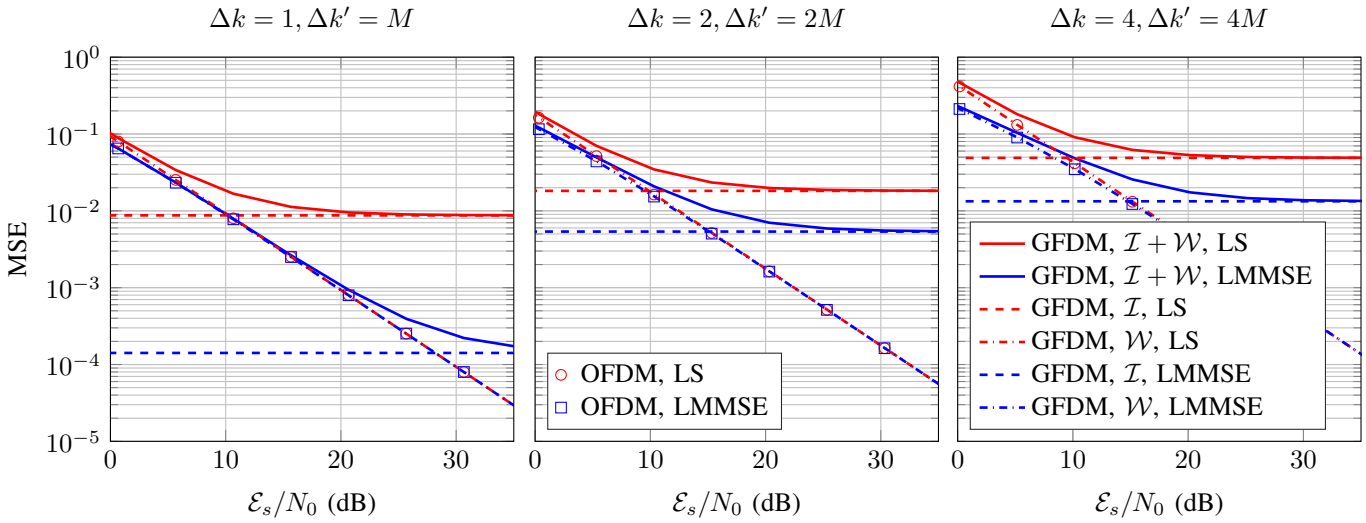


Fig. 7: MSE of channel estimation vs. SNR in Rayleigh fading 2×2 MIMO channel with $K = 128$ subcarriers and $M = 7$ subsymbols. $\mathcal{I}$ represents the MSE due to the interference from data while $\mathcal{W}$ stands for the MSE due to noise only.

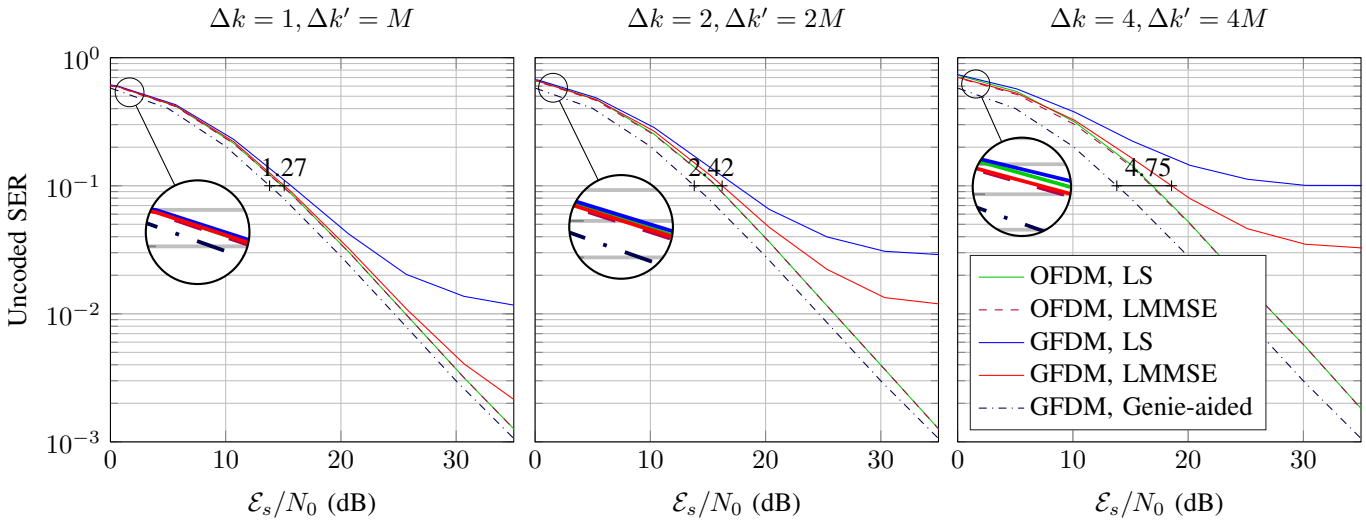


Fig. 8: Symbol Error Rate performance of GFDM under the influence of channel estimation in Rayleigh fading 2×2 MIMO channel with $K = 128$ subcarriers and $M = 7$ subsymbols

VI. Conclusions and Future Work

This paper presents a system model (10), serving as a general framework for deriving various pilot-aided channel estimators for MIMO-GFDM systems. Both LS and LMMSE criterion-based estimators are derived and their resulting MSE performance is analyzed. The MSE simulations confirm the closed form expressions of theoretical MSE calculations. Moreover, the MSE and SER performances of the MIMO-GFDM channel estimation show that both performances are similar to OFDM at low SNR whereas at high SNRs an error floor exists due to interference from data symbols.

The future work will rely on pre-canceling the interference term at the transmitter side due to its knowledge on data-bearing subsymbols. By properly pre-canceling the interference term, the GFDM channel estimation would be able to provide similar performance as in OFDM, while the GFDM waveform provides more flexibility and advantages such as extreme low latency, high throughput and low OOB emission.

Acknowledgment

The computations were performed on a computing cluster at the Center for Information Services and High Performance Computing (ZIH) at TU Dresden. The work presented in this paper has been performed in the framework of the SATURN project with contract no. 100235995 which is funded by the

Europäischer Fonds für regionale Entwicklung (EFRE). Part of this work is also sponsored by the Federal Ministry of Education and Research within the programme "Twenty20 - Partnership for Innovation" under contract 03ZZ0505B - "Fast Wireless".

REFERENCES

[1] G. Fettweis, "The Tactile Internet: Applications and Challenges," *IEEE Veh. Technol. Mag.*, vol. 9, no. 1, pp. 64–70, 2014.

[2] 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 Commun. Mag.*, vol. 52, no. 2, pp. 97–105, 2014.

[3] G. Fettweis and S. Alamouti, "5G: Personal mobile internet beyond what cellular did to telephony," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 140–145, 2014.

[4] B. Farhang-Boroujeny, "OFDM Versus Filter Bank Multicarrier," *IEEE Signal Process. Mag.*, vol. 28, no. 3, pp. 92–112, 2011.

[5] V. Vakilian, T. Wild, F. Schaich, S. ten Brink, and J.-F. Frigon, "Universal-filtered multi-carrier technique for wireless systems beyond LTE," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, 2013, pp. 223–228.

[6] X. Zhang, M. Jia, L. Chen, J. Ma, and J. Qiu, "Filtered-OFDM-Enabler for Flexible Waveform in The 5th Generation Cellular Networks," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, 2015, to be published.

[7] N. Michailow et. al., "Generalized Frequency Division Multiplexing for 5th Generation Cellular Networks," *IEEE Trans. Commun.*, vol. 62, no. 9, pp. 3045–3061, 2014.

[8] M. Matthé, I. Gaspar, D. Zhang, and G. Fettweis, "Short Paper: Near-ML Detection for MIMO-GFDM," in *Proc. IEEE 82nd Vehicular Technology Conference*, Boston, 2015.

[9] M. Matthé, L. Mendes, N. Michailow, D. Zhang, and G. Fettweis, "Widely Linear Estimation for Space-Time-Coded GFDM in Low-Latency Applications," *IEEE Trans. Commun.*, 2015.

[10] D. Zhang, L. Mendes, M. Matthé, I. Gaspar, N. Michailow, and G. Fettweis, "Expectation Propagation for Near-Optimum Detection of MIMO-GFDM Signals," *IEEE Trans. Wireless Commun.*, 2015.

[11] M. Danneberg, N. Michailow, I. Gaspar, M. Matthé, D. Zhang, L. L. Mendes, and G. Fettweis, "Implementation of a 2 by 2 MIMO-GFDM Transceiver for Robust 5G Networks."

[12] S. Coleri, M. Ergen, A. Puri, and A. Bahai, "Channel estimation techniques based on pilot arrangement in OFDM systems," *IEEE Trans. Broadcast.*, vol. 48, no. 3, pp. 223–229, 2002.

[13] M. Biguesh and A. Gershman, "MIMO channel estimation: optimal training and tradeoffs between estimation techniques," in *Proc. IEEE Int. Conf. on Commun.*, vol. 5, 2004, pp. 2658–2662 Vol.5.

[14] G. Taricco and E. Biglieri, "Space-time decoding with imperfect channel estimation," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1874–1888, 2005.

[15] T.-D. Chiueh, P.-Y. Tsai, and I.-W. Lai, *Baseband receiver design for wireless MIMO-OFDM communications*. John Wiley & Sons, 2012.

[16] Ghogho, M. and McLernon, D. and Alameda-Hernandez, E. and Swami, A., "Channel estimation and symbol detection for block transmission using data-dependent superimposed training," *IEEE Signal Process. Lett.*, vol. 12, no. 3, pp. 226–229, 2005.

[17] U. Vilaipornsawai and M. Jia, "Scattered-pilot channel estimation for GFDM," in *Proc. IEEE Wireless Commun. and Netw. Conf. (WCNC)*, 2014, pp. 1053–1058.

[18] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ: Prentice Hall, 1993, vol. 1.

[19] A. H. Roger and R. J. Charles, "Topics in matrix analysis," 1991.