TECHNISCHE UNIVERSITÄT DRESDEN

Advanced Channel Estimation Techniques for Multiple-Input Multiple-Output Multi-Carrier Systems in Doubly-Dispersive Channels

Shahab Ehsan Far

der Fakultät Elektrotechnik und Informationstechnik
der Technischen Universität Dresden

zur Erlangung des akademischen Grades eines

Doktoringenieurs

(Dr.-Ing.)

genehmigte Dissertation

Vorsitzender:	Prof. Dr. techn. Klaus Janschek
Gutachter:	Prof. Dr.-Ing. Dr. h.c. Gerhard Fettweis
	Prof. Dr. Mario Huemer
Tag der Einreichung:	30.10.2019
Tag der Verteidigung:	21.01.2020

Abstract

Flexible numerology of the physical layer has been introduced in the latest release of 5G new radio (NR) and the baseline waveform generation is chosen to be cyclic-prefix based orthogonal frequency division multiplexing (CP-OFDM). Thanks to the narrow subcarrier spacing and low complexity one tap equalization (EQ) of OFDM, it suits well to time-dispersive channels. For the upcoming 5G and beyond use-case scenarios, it is foreseen that the users might experience high mobility conditions. While the frame structure of the 5G NR is designed for long coherence times, the synchronization and channel estimation (CE) procedures are not fully and reliably covered for diverse applications. The research on alternative multi-carrier waveforms has brought up valuable results in terms of spectral efficiency, applications coexistence and flexibility. Nevertheless, the receiver design becomes more challenging for multiple-input multiple-output (MIMO) non-orthogonal multi-carriers because the receiver must deal with multiple dimensions of interference. This thesis aims to deliver accurate pilot-aided estimations of the wireless channel for coherent detection. Considering a MIMO non-orthogonal multi-carrier, e.g. generalized frequency division multiplexing (GFDM), we initially derive the classical and Bayesian estimators for rich multi-path fading channels, where we theoretically assess the choice of pilot design. Moreover, the well time- and frequency-localization of the pilots in non-orthogonal multi-carriers allows to reuse their energy from cyclic-prefix (CP). Taking advantage of this feature, we derive an iterative approach for joint CE and EQ of MIMO systems. Furthermore, exploiting the block-circularity of GFDM, we comprehensively analyze the complexity aspects, and propose a solution for low complexity implementation. Assuming very high mobility use-cases where the channel varies within the symbol duration, further considerations, particularly the channel coherence time must be taken into account. A promising candidate that is fully independent of the multi-carrier choice is unique word (UW) transmission, where the CP of random nature is replaced by a deterministic sequence. This feature, allows per-block synchronization and channel estimation for robust transmission over extremely doubly-dispersive channels. In this thesis, we propose a novel approach to extend the UW-based physical layer design to MIMO systems and we provide an in-depth study of their out-of-band emission, synchronization, CE and EQ procedures. Via theoretical derivations and simulation results, and comparisons with respect to the state-of-the-art CP-OFDM systems, we show that the proposed UW-based frame design facilitates robust transmission over extremely doubly-dispersive channels.

Acknowledgement

This dissertation is the result of four years diligent work as a Research Associate at Vodafone Chair Mobile Communications Systems at Technische Universität Dresden. I am cordially grateful to Prof. Gerhard Fettweis for giving me an opportunity to be a part of his team. His professional inspiration, continuous support and valuable discussions cleared the path for me towards this achievement.

I would like to thank Prof. Mario Huemer for accepting to be the second reviewer of my thesis and giving valuable feedback in our meeting. I am also grateful to the chair of examination committee Prof. Klaus Janschek and the board member of the committee Prof. Frank Fitzek.

Additionally, I would like to thank my group leaders Dr. Dan Zhang and Prof. Marwa Chafii as well as my senior colleague Dr. Maximilian Matthé for the fruitful discussions in our meetings and teaching me how to efficiently write and revise scientific papers. Prof. Chafii and Dr. Matthé have also reviewed draft versions of this thesis and I sincerely appreciate the time and the effort they have spent on it. I would also thank my co-authors and colleagues Martin, Ahmad, Roberto, Song, Mostafa, Behnam, Zhitao, Zhonjgu, Ana, Ivo and many other colleagues at the chair for the interesting talks, meetings, lunch breaks and social dinners. My special thanks also go to Kathrin, Sylvia, Raffael and Rüdiger for supporting my administrative and organizational matters.

Finally, I am profoundly grateful to my family and friends for giving me the encouragement throughout the years of study and research. This work could not have been possible without my mother's support on my studies from Iran to abroad in Italy and now in Germany.

Dresden, January 2020 Shahab Ehsanfar

Contents

List of Figures

List of Tables

Acronyms

5G fifth generation

AC auto-correlation

AWGN additive white Gaussian noise

BER bit error rate

CC cross-correlation

CCM channel covariance matrix

CE channel estimation

CFO carrier frequency offset

CIR channel impulse response

CP cyclic prefix

CP-OFDM cyclic-prefix based orthogonal frequency division multiplexing

CRLB Cramér-Rao lower bound

CSI channel state information

CWCU component-wise conditionally unbiased

DFT discrete Fourier transform

DZT discrete ZAK transform

eMBB enhanced mobile broadband

EQ equalization

ETU extended typical urban

EVA extended vehicular A

FER frame error rate

FFT fast Fourier transform

FIR finite impulse response

FLOPs floating point operations

GFDM generalized frequency division multiplexing

GS guard symbol

HM harmonic mean

Hz Hertz

i.i.d. independent and identically distributed

IAI inter-antenna-interference

IBI inter-block-interference

ICI inter-carrier-interference

IFPI interference-free pilot insertion

ISI inter-symbol-interference

LMMSE linear minimum mean square error

LS least squares

LTE Long Term Evolution

MCS modulation and coding scheme

MF matched filter

MI mutual information

MIMO multiple-input multiple-output

ML maximum likelihood

mMTC massive machine type communication

MSE mean squared error

NR New Radio

OFDM orthogonal frequency division multiplexing

OOB out-of-band

PAPR peak to average power ratio

PCCC parallel concatenated convolutional codes

pdf probability density function

PDP power-delay-profile

PIC parallel interference cancellation

PSD power spectral density

RB resource block

RC Raised-Cosine

Rx receive

SC single-carrier

SC-FDE single carrier frequency domain equalization

SC-FDMA single carrier frequency domain multiple access

SER symbol error rate

SISO single-input single-output

SLMMSE sequential LMMSE

SNR signal-to-noise ratio

SoA state-of-the-art

STO symbol time offset

Tx transmit

URLLC ultra reliable low latency communication

UW unique-word

ZF zero forcing

Chapter 1

Introduction

HOLISTIC fifth generation (5G) cellular networks enable new applications and use-cases for the vertical markets, and also new challenges for more efficient service provisioning. Thanks to the flexible numerology of the physical layer in latest release of 5G New Radio (NR) [3GP18], diverse application services such as enhanced mobile broadband (eMBB), ultra reliable low latency communication (URLLC), and massive machine type communication (mMTC) are envisioned to coexist within the same time interval and also the same frequency band. Examples of these use-cases are augmented reality, virtual reality and 3d video streaming, which massively increase the data-rate requirements; Further examples are industrial automation, remote healthcare and autonomous driving, which require stringent latency of sub-millisecond; and massive number of low-cost connected devices for sensing, tracking, monitoring, etc. On top of these use cases, the Tactile Internet [Fet14] further challenges the state-of-the-art (SoA) wireless networks for applications such as free-viewpoint-video with latency requirements of below 10ms, 1-10ms synchronization and 100Gb/s data-rate. A further immersive application that has been defined by the Tactile Internet is the remote control of humanoid robots and robotic assisted surgery that require an end-to-end round-trip latency of 1ms. While in Rel. 15 of 5G NR, eMBB service has been the main focus, frame designs and in particular synchronization and channel estimation procedures are not fully and reliably covered for diverse applications. Putting high mobility on top, a robust data transmission becomes even more challenging, specially, if the spectrum is also being shared for multiple services.

1.1 Multi-Carrier Waveforms

While orthogonal frequency division multiplexing (OFDM) is being the key modulation technique for current wireless communication systems, the research on its alternatives is still ongoing and it has brought up valuable results in terms of spectral efficiency, applications coexistence and flexibility of the radio networks. For instance, [MMG+14] proposed a unified air-interface namely generalized frequency division multiplexing (GFDM)

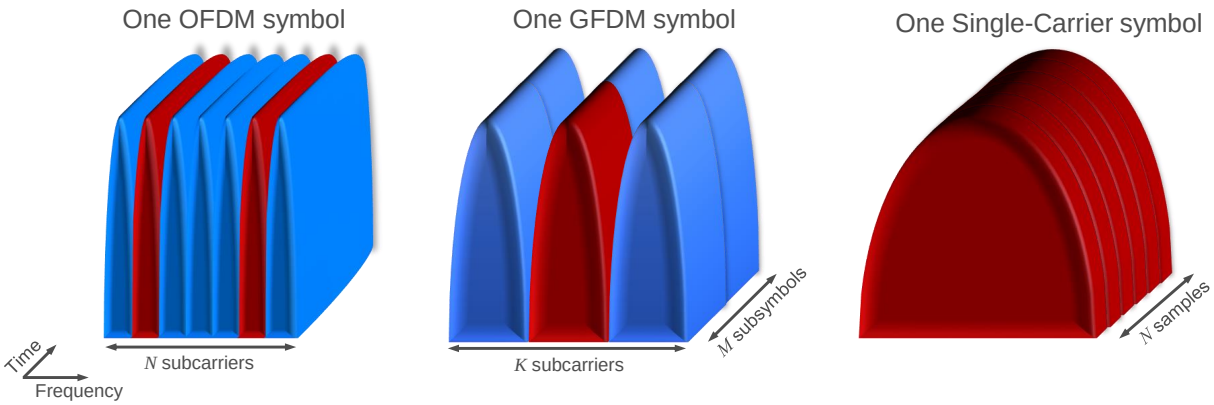


Fig. 1.1: Pilot insertion in different transmission techniques (red: pilot(s), blue: data).

that emulates any single/multi-carrier system [DNM+18]. Via simple parameter configurations, GFDM can adjust the physical layer to serve the above mentioned diverse services [FRG+17], although, its design also represents a trade-off between complexity, flexibility and performance. Thanks to the possibility of circular-filtering in GFDM, peak to average power ratio (PAPR)—that is a key characteristic for battery-life of wireless device—and/or out-of-band (OOB) emission—that is a necessary consideration for spectrum sharing—can be reduced with respect to OFDM[1]. Nevertheless, GFDM is being positioned in the context of non-orthogonal waveforms, because if multiple symbols are packed to use a single cyclic prefix (CP), and also because of the subcarrier-wise filtering, the receiver sees a signal that is subject to inter-symbol-interference (ISI) and inter-carrier-interference (ICI). In multiple-input multiple-output (MIMO) systems, also inter-antenna-interference (IAI) unites with the self-interference generation of non-orthogonal waveforms.

Pilot Contamination in Orthogonal vs. Non-Orthogonal Waveforms. Fig. 1.1 illustrates the pilot insertion in multiple systems. An OFDM symbol comprises N narrow subcarriers. Since in OFDM every symbol is protected by a dedicated CP, the N subcarriers remain orthogonal to each other in a perfectly time and frequency synchronized receive (Rx) signal. Therefore, the pilots can be easily separated from the data subcarriers via discrete Fourier transform (DFT) operation. An example of a practical system that exploits this orthogonality principle is the downlink resource block (RB) of Long Term Evolution (LTE) [Acc09], which inserts scattered pilots at different time and frequency resource elements. On the other hand, if the channel becomes time-varying, frequency dispersion causes the OFDM data-subcarriers to leak over the pilot-subcarriers. In this case, ICI occurs and the pilots become contaminated by the data subcarriers.

[1] We should point out that in [ZMMF17], two major configurations of GFDM have been studied in terms of PAPR and OOB emission. According to [ZMMF17], GFDM type-I achieves a smaller OOB radiation compared to OFDM, but worse PAPR. GFDM type-II configuration achieves a smaller PAPR with respect to OFDM, but gains no advantage in terms of OOB emission.

In a non-orthogonal system e.g. single-carrier (SC), N time samples (also protected by a CP) are overlapped in frequency. Therefore, an SC symbol must be either pilot or data to avoid interference[2]. In a practically implemented system e.g. LTE uplink RB, the 4-th and the 11-th symbols are entirely configured as pilots. Similarly in GFDM, M subsymbols are superimposed in frequency domain. In time domain, only a single CP is used for the M subsymbols. Therefore, the pilots are subject to both ISI and ICI and thus, clear extraction of the pilots (whether in time nor frequency) from the Rx signal is not straightforward. In addition, employing a MIMO technique—that will be discussed in Sec. 2.3—for the non-orthogonal systems, adds a further dimension of pilot contamination, i.e. inter-antenna-interference. In Sec. 2.1, we will formally define the non-orthogonality in multi-carrier systems.

1.2 MIMO Systems

Nowadays, a crucial part of any modern transceiver is MIMO, in which, the spatial diversity is either used to improve throughput or robustness. The MIMO systems can be considered as centralized or distributed. In distributed MIMO systems, the transmit (Tx) antennas are placed farther apart (likewise, the Rx antennas). Each antenna might use a different oscillator clock source and also the power-delay-profile (PDP) for each Tx-Rx antenna pair might differ. In centralized MIMO systems however, the Tx antennas are collocated, and thus, they use a single oscillator clock source (likewise, the Rx antennas). Therefore, it can be assumed that the transmit frames are perfectly aligned by the time they arrive at the Rx antennas and there is no timing misalignment between them (see [MS01],[CT08] and references therein). Moreover, due to the collocation of antennas in centralized MIMO systems, the PDP is commonly assumed to remain identical between each Tx-Rx antenna pair. Throughout this thesis, we mostly consider centralized MIMO systems, but the algorithm developments can analogously be extended to distributed MIMO systems as well.

Pilot Design in MIMO Systems. Several pilot schemes and patterns have been proposed in literature to accommodate different MIMO scenarios. For instance, [CTL12] outlines three major MIMO OFDM pilot allocation schemes namely, preamble-based, space-time pilots and space-frequency pilots.

- The MIMO preamble design is spectrally efficient for continuous transmission over slow fading channels.

[2] To put it another way, if pilots are time multiplexed with the data within few samples of an SC symbol, or they are frequency multiplexed over few subcarriers, pilots become contaminated by the data samples. Thus, clear pilot observation is achieved if *all* active subcarriers of an SC symbol are pilots.

- The space-time pilot insertion is an extension of comb-type pilot pattern[3] to MIMO. But in order to reduce the overhead of comb-type pilots, it considers distinct pilot sets over multiple OFDM symbols. Consequently, the space-time pilot design is not suitable for fast-fading channels.
- The space-frequency pilot pattern increases the pilot overhead by inserting distinct pilot sets on different subcarriers of every OFDM symbol. This way, the channel can be estimated on a per-symbol basis which is suitable for doubly-dispersive channels, i.e. both time- and frequency-selective. Nonetheless, when employing space-frequency pilots, it is still assumed that the channels remain stationary during a symbol period.

This thesis aims to deliver accurate estimates of the channel impulse response (CIR) for coherent detection via pilot-aided schemes of MIMO systems, and regardless of the single/multi-carrier choice. Thus, employing a non-orthogonal waveform in doubly-dispersive scenarios, sophisticated pilot designs need to be analyzed. For instance, a question that comes to the mind is, how can one relate the pilot insertion to the self-interference statistics? In other words, which kind of pilot pattern allows a MIMO receiver to achieve sufficient estimation quality, despite the fact that self-interference terms still exist? In chapter 3, we will address these questions by means of mathematical derivations as well as extensive simulations.

Moving forward towards extreme fast-fading—as well as frequency-selective—conditions where the channel starts to vary within the symbol duration, further considerations, particularly the channel coherence time must be taken into account. A promising candidate that is fully independent of the waveform choice is unique-word (UW) transmission [DGE00], [WMS+02], where the CP of random nature is replaced by a deterministic sequence. This way, per-block synchronization and channel estimation allows a more robust transmission over extreme doubly-dispersive channels. In chapter 2, we review the SoA of UW designs in single-input single-output (SISO) systems, and in chapter 4, we show an efficient approach to extend the UW design concept to MIMO applications.

1.3 Contributions and Thesis Structure

Integrating the non-orthogonal flexible GFDM into a MIMO framework, a sophisticated receiver design which estimates the MIMO wireless channels with high precision is crucial. Although, numerous algorithms for joint channel estimation and equalization of MIMO-OFDM exist in the literature, e.g. [TC07], [KY02], [CT08], and references therein, they are not directly applicable to non-orthogonal multi-carriers, because they do not consider multiple dimensions of interference, i.e. ISI, ICI, IAI. Considering MIMO interference-limited non-orthogonal waveforms as a baseline, this thesis tailors the following aspects:

[3] In comb-type pilot pattern, pilots with some pilot-subcarrier spacing are inserted into every OFDM symbol.

Chapter 2, mathematically formulates the channel estimation (CE) problem by means of basic linear algebra. Afterwards, a non-orthogonal modulation approach, in particular GFDM is outlined. Then, we review the fundamentals of wireless MIMO transmission along with CIR estimation in MIMO OFDM systems. Thereafter, we review the SoA UW based transmission schemes, in which, the conventional CP insertion is replaced by employing a deterministic sequence instead of random CP.

Chapter 3, provides the contributions of this thesis in terms of non-orthogonal MIMO channel estimation. We begin our work with standard derivations of classical and Bayesian channel estimation for MIMO GFDM systems. Although, GFDM is considered as the main non-orthogonal waveform in this work (due to its flexibility in emulating various single/multi-carriers [DNM+18]), the derivations are valid for other orthogonal/non-orthogonal systems as well. In Chapter 3, we also study the pilot design for GFDM where we theoretically show that rectangular-grid pilot pattern of GFDM better suppresses the interference. Immediately thereafter, we adjust the frequency domain GFDM modulator for orthogonal pilot insertion with the purpose that interference-free pilots can be observed at the MIMO GFDM receiver. Since the initial non-iterative approaches attain only a sub-optimal performance of CE, we extend our MIMO CE techniques to iterative linear minimum mean square error (LMMSE) parallel interference cancellation (PIC) for close-to-optimal joint *channel-estimation-and-equalization* of MIMO non-orthogonal systems. In our derivations, we also take into account a case in which the pilots' information from the whole transmission block, including its CP, is considered.

In Chapter 4, we extend the SoA SISO UW-based transmission to MIMO for synchronization and channel estimation of any single/multi-carrier system, where the underlying wireless channel scenario is considered to be extremely doubly-dispersive. There, we first propose a UW based frame design and we compare its overhead requirements to an equivalent CP-based system. Since the signal characteristics (particularly OOB emission) of a non-orthogonal waveform is one of its key advantages compared to OFDM, we analyze the spectral properties of MIMO UW sequences. Then, assuming a centralized MIMO system, we elaborate the timing and frequency synchronization metrics of UW-based systems, where per-block synchronization is its key aspect for providing a robust transmission in doubly-dispersive channels. Thereafter, we tailor the UW-based channel estimation for highly frequency-selective and time-variant channels by deriving the *Wiener-Hopf* adaptive filters. For the sake of more robust channel estimation in frequency-selective channels, we also propose a frame structure that circularly rotates the UW sequences transmitted from different antennas. There, we derive and simulate the Circular (Circ.) UW-based LMMSE channel estimation and we compare its performance with a CP-based system as well as the basic UW-based system. In Chapter 4, we propose and derive a *UW-free* LMMSE-based joint *equalization-and-demodulation* technique that also considers imperfect channel knowledge for close-to-optimal detection of transmitted signal constellations.

We conclude the main outcomes of this thesis in Chapter 5 and subsequently, our discus-

sion follows by the remaining challenges and open topics for MIMO channel estimation and equalization in doubly-dispersive channels.

1.4 Notations

- Column-vectors are denoted by vector sign $\vec{X}$ and matrices by boldface $\mathbf{X}$.
- Time domain signals are represented by lowercase letters while normal font uppercase letters are used for DFT and discrete ZAK transform (DZT) domain signals.
- $\mathbb{E}[\cdot]$ is the expectation operator.
- $\lfloor \cdot \rfloor$ denotes the floor operation.
- X modulo Y is denoted by $X \mod Y$.
- The trace of a square matrix $\boldsymbol{X}$ is $\mathrm{Tr}(\boldsymbol{X})$.
- The real and imaginary components of a complex value X is denoted by $\Re(X)$ and $\Im(X)$, respectively.
- The conjugate, transpose and Hermitian transpose of $\boldsymbol{X}$ are denoted by $\boldsymbol{X}^*$, $\boldsymbol{X}^T$ and $\boldsymbol{X}^H$, respectively.
- The Frobenius norm of a matrix $\boldsymbol{X}$ is $\|\boldsymbol{X}\|$ and its square can be written as $\|\boldsymbol{X}\|^2 = \mathrm{Tr}(\boldsymbol{X}\boldsymbol{X}^H)$.
- $\left|\vec{X}\right|^2$ denotes the element-wise absolute square of vector $\vec{X}$.
- $\boldsymbol{X} \otimes \boldsymbol{Y}$ and $\boldsymbol{X} \circ \boldsymbol{Y}$ are the Kronecker and Hadamard products [RC91] of matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$, respectively.
- $\frac{\vec{X}}{\vec{Y}}$ is the element-wise division of vector $\vec{X}$ over $\vec{Y}$.
- $\sqrt{\mathbf{X}}$ is the element-wise square root of matrix $\mathbf{X}$.
- $\boldsymbol{X} \oplus \boldsymbol{Y}$ is the direct sum of two matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$.
- $\mathrm{diag}(\vec{X})$ is a diagonal matrix whose diagonal entries are the entries of the column vector $\vec{X}$.
- $\mathrm{blkdiag}(\boldsymbol{X}, \cdots, \boldsymbol{Y})$ 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.
- $\mathrm{circ}(\cdot)$ returns a circulant matrix associated to its input row vector.
- $\mathtt{vec}(\boldsymbol{X})$ is the operation of stacking the columns of $\boldsymbol{X}$ on top of one another.

- The matrices $\boldsymbol{F}_n$ and $\boldsymbol{I}_n$ are the unitary DFT and the identity matrices of size $n \times n$, respectively.
- The matrix $\mathbf{0}_{m \times n}$ is an all *zero* matrix of size $m \times n$.
- $\vec{0}_n$ and $\vec{1}_n$ are column vectors of size n with all *zero* and *one* entries, respectively.
- The distribution of a complex multivariate Gaussian random vector $\vec{X}$ with mean $\vec{\mu} = \mathbb{E}[\vec{X}]$ and covariance matrix $\mathbf{\Sigma} = \mathbb{E}[(\vec{X} - \vec{\mu})(\vec{X} - \mu)^H]$ is denoted by $\mathcal{N}_C(\vec{\mu}, \mathbf{\Sigma})$.

Chapter 2

State-of-the-art and Fundamentals

Since 4G LTE up to the latest releases of 5G NR, the baseline waveform generation of cellular networks physical layers has been chosen to be cyclic-prefix based orthogonal frequency division multiplexing (CP-OFDM). Thanks to the narrow subcarrier spacing and low complexity one tap equalization of OFDM, it suits well to time-dispersive channels. In the following, before diving into the CE techniques of the SoA CP-OFDM, we formally formulate the problem of orthogonality vs. non-orthogonality in wireless transmission from a channel estimation perspective. Afterwards, having the MIMO wireless channel defined in linear algebra, we recap the most popular CE techniques for CP-OFDM, e.g. as in [CT08], [CTL12], [BLM03]. Thereafter, reviewing the SoA UW-based transmission—which is an alternative to the CP-based systems [WMS+02] for emulating a cyclic signal at the receiver—and comparing it with respect to a CP-based system, we provide initial recommendations of frame design based on the outcomes of the SoA approaches. However, since in related works, the UW-based systems are separately designed for synchronization [DGE00], channel estimation [CSBM06] and equalization [HOH11] modules, further analysis need to be taken into account, particularly, if the frame designs are going to be extended to multiple-antenna systems. In our problem formulation in this chapter, we consider that a single approach of frame design must jointly deal with all the above mentioned modules of the receiver, and later in Chapter 4, we propose our solutions in order to address each of the challenges.

2.1 Linear Systems and Problem Statement

Assume a linear system, in which the measurement vector $\vec{y}$ is characterized by the following expression

$$\vec{y} = \boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{M}}(\underbrace{\vec{d}_p + \vec{d}_d}_{\vec{d}}) + \vec{w}, \tag{2.1}$$

where, $\vec{d}_p \in \mathbb{C}^M$ denotes a known vector (i.e. pilots), $\vec{d}_d \in \mathbb{C}^M$ denotes the vector of unknown data symbols with $\vec{d}_d \sim \mathcal{N}_C(\vec{0}, \boldsymbol{\Sigma}_{dd})$, $\vec{w} \in \mathbb{C}^N$ denotes the additive white Gaus-

sian noise (AWGN) vector with $\vec{w} \sim \mathcal{N}_C(\vec{0}, \sigma_w^2 \boldsymbol{I}_N)$. The matrix $\boldsymbol{\mathcal{M}} \in \mathbb{C}^{N \times M}$ is a complex modulation matrix that maps the M-dimensional vector $\vec{d}$ into an N-dimensional transmit signal. Further, the matrix $\boldsymbol{\mathcal{H}} \in \mathbb{C}^{N \times N}$ denotes a complex square matrix of the channel which depends on the parameters vector $\vec{h} \in \mathbb{C}^L$ with $L << N$ unknown parameters.

Under the above definitions, the observation vector $\vec{y}$ depends on two types of unknown parameters i.e. $\vec{h}$ and $\vec{d}_d$, while in order to extract the data vector $\vec{d}_d$ from the measurement vector $\vec{y}$, knowledge of $\vec{h}$ is essential. Generally, it is assumed that $\vec{h}$ and $\vec{d}_d$ are independent. Therefore, in a sequential manner, one exploits the knowledge of the vector $\vec{d}_p$ to estimate $\vec{h}$ and subsequently, to decode the data vector $\vec{d}_d$.

Given the above linear model, we define the CE problem as

a) Find a method to estimate the parameters vector $\vec{h}$ out of $\boldsymbol{\mathcal{H}}$ from the measurements $\vec{y}$.

b) Optimize the known signal $\vec{x}_p = \boldsymbol{\mathcal{M}}\vec{d}_p$ and the distribution of $\vec{d}_d$ for a suitable estimation performance with constraints on the signal PAPR and OOB emission.

The following assumptions alleviate the above CE problem:

- *Block-fading:* If $\vec{h}$ remains static during a transmission, $\boldsymbol{\mathcal{H}} \triangleq \boldsymbol{\mathcal{T}}_N$ becomes a lower triangular Toeplitz matrix $\boldsymbol{\mathcal{T}}_N \in \mathbb{C}^{N \times N}$ with $\vec{h}$ on its first column.
- If in a *block-fading* situation, the system employs a CP that is equal or longer than L, $\boldsymbol{\mathcal{H}} \triangleq \boldsymbol{\mathcal{C}}_N$ becomes circulant $\boldsymbol{\mathcal{C}}_N \in \mathbb{C}^{N \times N}$.
- The measurements $\vec{y}$ and the transmit signal $\vec{x} = \boldsymbol{\mathcal{M}}\vec{d}$ are perfectly synchronized in time and in frequency.

Thereupon, in a *CP-OFDM* system under the assumption of block-fading as well as perfect synchronization, the matrix $\boldsymbol{\mathfrak{D}} = \boldsymbol{F}_N \boldsymbol{\mathcal{H}} \boldsymbol{\mathcal{M}}_{\text{OFDM}}$ with $\boldsymbol{\mathcal{M}}_{\text{OFDM}} = \boldsymbol{F}_N^H$ is a diagonal matrix. If $\vec{d}_p \circ \vec{d}_d = \vec{0}_N$ holds, the elements of $\boldsymbol{F}_N \boldsymbol{\mathcal{M}} \vec{d}_p$ and $\boldsymbol{F}_N \boldsymbol{\mathcal{M}} \vec{d}_d$ become orthogonal which yields clear pilot observation at the receiver side. Moreover, given that the noise vector $\vec{w}$ is white and it is inversely proportional to the signal power—and thus signal-to-noise ratio (SNR)—and also $\boldsymbol{F}_N$ being unitary, the DFT domain $\boldsymbol{F}_N \vec{w}$ is white too. In a time-varying scenario where the channel transfer function varies rapidly in time and in frequency, the CIRs on the columns of $\boldsymbol{\mathcal{H}}$ differ from one another, and consequently, $\boldsymbol{\mathfrak{D}}' = \boldsymbol{F}_N \boldsymbol{\mathcal{H}} \boldsymbol{\mathcal{M}}_{\text{OFDM}}$ is no longer a diagonal matrix. In this case, due to the frequency dispersion, the orthogonality principles of OFDM is being lost and thus, in estimation of $\vec{h}$, an error floor independent of SNR rises.

In a *non-orthogonal* multi-carrier system, $\boldsymbol{\mathcal{M}}$ is not restricted to $\boldsymbol{F}_N^H$ and it can apply any linear process on the vector $\vec{d}$. For instance, via linear or circular filtering the orthogonality principles can be partially abandoned in order to better control the Tx signal impairments,

e.g. lower PAPR and/or lower OOB emission. In this case, there is no linear operation $\boldsymbol{\mathcal{Z}}$ independent of $\vec{h}$ that makes the matrix $\boldsymbol{\mathfrak{D}}'' = \boldsymbol{\mathcal{Z}}\boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{M}}$ diagonal. Thus, despite having $\vec{d}_p \circ \vec{d}_d = \vec{0}_N$, the elements of $\vec{x}_p$ and $\vec{x}_d = \boldsymbol{\mathcal{M}}\vec{d}_d$ might be superimposed (both in time and in frequency domain), which results into pilot contamination. Under such circumstances, the noise-plus-interference vector $\vec{E} = \boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{M}}\vec{d}_d + \vec{w}$ is correlated with covariance matrix $\boldsymbol{\Sigma}_{EE}$. Due to $\boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{M}}\vec{d}_d$ being independent of SNR, estimation of $\vec{h}$ always subjects to error floor.

2.2 GFDM Modulation

We adopt a GFDM data matrix $\boldsymbol{D} \in \mathbb{C}^{K\times M}$ with K subcarriers and M complex valued subsymbols. The subsymbol at time index m and subcarrier index k is denoted by $d_k[m]$. In GFDM, each data symbol $d_k[m]$ is pulse-shaped using circularly time and frequency shifted version of a prototype filter $g[n]$ whose energy during one period $N = MK$ is normalized to unity. Formally, the GFDM transmit sample [MMG+14] is given by

$$x[n] = \sum_{k=0}^{K-1}\sum_{m=0}^{M-1} d_k[m]g\Big[(n-mK) \bmod N\Big]e^{j2\pi\frac{k}{K}n}, \tag{2.2}$$

with $n = 0, \cdots, N-1$ and $d_k[m] = 0$ if it belongs to an inactive subcarrier. In the above expression, circular filtering is obtained via the modulo operation. Collecting the samples in matrix-vector form, the GFDM transmit signal becomes

$$\vec{x} = \boldsymbol{A}\vec{d}, \tag{2.3}$$

where $\vec{d} = [\vec{d}_0^T, \cdots, \vec{d}_{K-1}^T]^T \in \mathbb{C}^N$ is vectorized version of $\boldsymbol{D}$ while for each subcarrier k its corresponding $\vec{d}_k$ vector is given by $\vec{d}_k = \Big(d_k[m]\Big)^T_{(m=0,\cdots,M-1)}$. Moreover, $\vec{d}_k \in \{\vec{d}_{k,p}, \vec{d}_{k,d}\}$, where $\vec{d}_{k,p}$ carries information (consisting both pilots and data) if k belongs to a pilot subcarrier ($k \in \mathcal{K}_p$). Otherwise, if k refers to a data subcarrier ($k \in \mathcal{K}_d$), then $\vec{d}_{k,p} = \vec{0}_M$. In a similar fashion, $\vec{d}_{k,d}$ carries the symbols at data subcarriers, with $\vec{d}_{k,d} = \vec{0}_M$ if $k \in \mathcal{K}_p$. Hence, $\vec{d}_{k,p} \circ \vec{d}_{k,d} = \vec{0}_M$ meaning no information between pilot and data subcarriers is superimposed. For future use, we also denote the total number of pilot subcarriers by $K_p = \lfloor K/\Delta k \rfloor$ with Δk being the pilot subcarrier spacing. Likewise, $K_d = K - K_p$ is the total number of data subcarriers. The modulation matrix is denoted by $\boldsymbol{A}$ and it follows [MMG+14]:

$$\boldsymbol{A} \triangleq \Big(\vec{g}_{0,0}, \cdots, \vec{g}_{K-1,0}, \vec{g}_{0,1}, \vec{g}_{1,1}, \cdots, \vec{g}_{K-1,M-1}\Big), \tag{2.4}$$

where $\vec{g}_{k,m} = \Big(g_{k,m}[n]\Big)^T_{n=0,1,\cdots,N-1}$ denotes the prototype filter at subcarrier index k and subsymbol index m, i.e. $g_{k,m}[n] \triangleq g\Big[(n-mK) \bmod N\Big]e^{j2\pi\frac{k}{K}n}$. Next, we write the equivalent version of (2.3) for the signal associated to the data subcarriers as [GMN+13]

$$\vec{\tilde{x}} = \boldsymbol{F}_N^H \sum_{k=0}^{K-1} \boldsymbol{P}^{(k)}\boldsymbol{G}^{(\delta)}\boldsymbol{R}^{(\delta)}\boldsymbol{F}_M\vec{d}_{k,d}, \tag{2.5}$$

where $\boldsymbol{R}^{(\delta)}$ is δ-fold repetition matrix which concatenates δ identity matrices $\boldsymbol{I}_M$ of size M, i.e. $\boldsymbol{R}^{(\delta)} = (\boldsymbol{I}_M \ \boldsymbol{I}_M \ \cdots)^T$. The value of δ is based on the number of non-zero values in filter frequency response, e.g. if a filter spans over two subcarriers, δ is then typically selected as $\delta = 2$. In (2.5) due to the circular filtering, the subcarrier filter $\boldsymbol{G}^{(\delta)} = \text{diag}(\vec{G}^{(\delta)})$ is diagonal in frequency domain with $\vec{G}^{(\delta)} = \boldsymbol{F}_{M\delta}\vec{g}^{(\delta)}$. The circulant filter $\vec{g}^{(\delta)}$ is the down-sampled version of $\vec{g} = \big(g[n]\big)_{n=0,\cdots,N-1}$ by factor K/δ. The permutation matrix $\boldsymbol{P}^{(k)}$ shifts the DC signals to their corresponding subcarriers (i.e. k) and is given by

$$\boldsymbol{P}^{(k)} = \boldsymbol{C}_\ell \begin{bmatrix} \boldsymbol{0}_{M\delta/2} & \boldsymbol{I}_{M\delta/2} & \boldsymbol{0}_{M\delta/2\times(N-\delta M)} \\ \boldsymbol{I}_{M\delta/2} & \boldsymbol{0}_{M\delta/2} & \boldsymbol{0}_{M\delta/2\times(N-\delta M)} \end{bmatrix}^T, \tag{2.6}$$

where $\ell = kM - M\delta/2$. The circulant matrix $\boldsymbol{C}_\ell$ follows:

$$\boldsymbol{C}_\ell = \text{circ}([\vec{0}^T_{\ell \bmod N},\ 1,\ \vec{0}^T_{(N-\ell-1) \bmod N}]). \tag{2.7}$$

2.3 MIMO Wireless Channel

We consider a multipath MIMO block fading channel with $\mathfrak{I}$ Tx and $\mathfrak{Q}$ Rx antennas. At the Tx antenna i, the signal is modulated using a linear multi-carrier approach and then protected by a CP. We assume that all the channels have shorter lengths L than the CP length. Thus, the receive signal after CP removal at Rx antenna q is characterized by the following equation:

$$\vec{y}_q = \sum_{i=1}^{\mathfrak{I}} \vec{x}_i \circledast \vec{h}_{i,q} + \vec{w}_q, \tag{2.8}$$

where $\circledast$ denotes the circular convolution of the CIR $\vec{h}_{i,q}$ and the transmit signal $\vec{x}_i$. Further, $\vec{x}_i = \vec{x}_{p,i} + \vec{x}_{d,i}$ can be decomposed into a known signal of pilots $\vec{x}_{p,i}$ and random data signal $\vec{x}_{d,i}$. The vector $\vec{w}_q$ denotes the AWGN process with variance σ_w^2 received at antenna q. A general overview of the system has been depicted in Fig. 2.1.

Although, the CIR may vary in time according to the maximum Doppler shift f_d, typically, it is assumed that it remains stationary during one or multiple transmission blocks. Hence, for a CP-less Tx signal, we model the L-length CIR between antennas i and q as a linear finite impulse response (FIR) filter given by

$$h_{iq}[n] = \sum_{\ell=0}^{L-1} h_{\ell,i,q}\delta\left[n - \tilde{n}_\ell\right], \tag{2.9}$$

where $\tilde{n}_\ell$ is the delay of the ℓ-th tap. $h_{\ell,i,q}$ is the envelope of the CIR at tap ℓ and it can be written as a complex value with its *real* and *imaginary* components independent and identically distributed (i.i.d.) zero mean Gaussian processes parameterized by the profile $\mathbb{E}[|h_{\ell,i,q}|^2] = P_{i,q}[\ell]$. If the system employs a CP with length $N_{\text{CP}} > L$, the above linear FIR filter turns into a circular filter. In vector notations, we write the CIR as

$$\vec{h}_{iq} = \sqrt{\text{diag}(\vec{P}_{iq})}\vec{\mathfrak{g}}_{iq}, \tag{2.10}$$

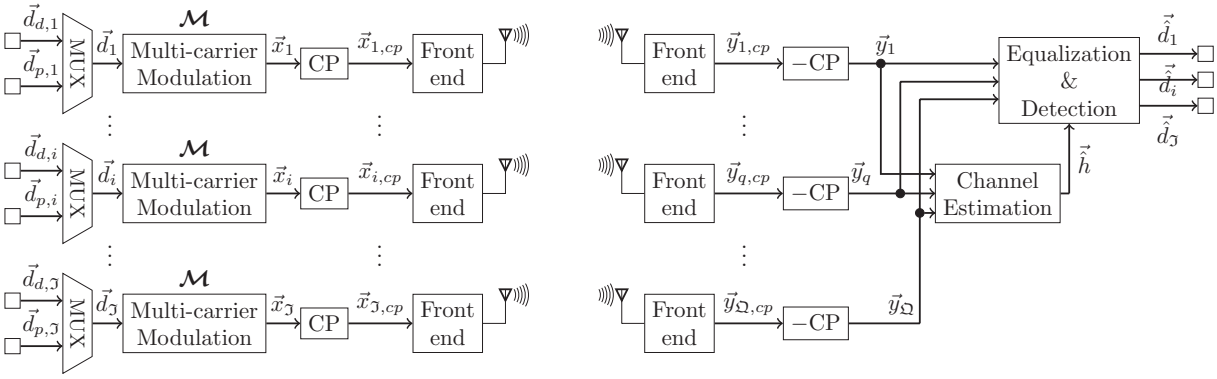


Fig. 2.1: Overall block diagram of the generic MIMO system.

where $\vec{\mathfrak{g}}_{iq} \sim \mathcal{N}_C(\vec{0}, \boldsymbol{I}_L)$, and $\vec{P}_{iq} \in \mathbb{R}^L$ denotes the normalized PDP between antennas i and q, i.e. $\|\vec{P}_{iq} = 1\|$.

In DFT domain, circular convolution turns into multiplication and therefore, the observed signal on Rx antenna q is given by

$$\vec{Y}'_q = \sum_{i=1}^{\mathfrak{I}} (\boldsymbol{X}'_{p,i} + \mathbf{X}'_{d,i}) \vec{H}'_{i,q} + \vec{W}'_q, \tag{2.11}$$

with $\vec{H}'_{i,q} = \sqrt{N}\mathbf{F}_{N,L}\vec{h}_{i,q}$. Furthermore, $\boldsymbol{X}'_{s,i} = \mathrm{diag}(\vec{X}'_{s,i})$ is a diagonal matrix associated either to pilots p or data sequences d (i.e. $s \in \{p, d\}$). $\vec{X}'_{s,i}$ is being transmitted on Tx antenna i and it is defined as $\vec{X}'_{s,i} \triangleq (\boldsymbol{F}_N \boldsymbol{\mathcal{M}} \vec{d}_s)_i$. The matrix $\mathbf{F}_{N,L} \in \mathbb{C}^{N \times L}$ contains only the first L columns of the unitary DFT matrix $\mathbf{F}_N \in \mathbb{C}^{N \times N}$. $\vec{W}'_q$ is the frequency domain counterpart of AWGN process on receive antenna q.

If the number of pilot subcarriers is smaller than the number of data subcarriers, i.e., the pilot-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:

$$\vec{Y}_q = \sum_{i=1}^{\mathfrak{I}} (\boldsymbol{X}_{p,i} + \mathbf{X}_{d,i}) \vec{H}_{i,q} + \vec{W}_q, \tag{2.12}$$

where $\vec{H}_{i,q} = \sqrt{N}\mathbf{F}_L \vec{h}_{i,q}$, $\vec{W}_q = \mathbf{F}_{N_p,N} \vec{w}_q$, $\boldsymbol{X}_{s,i} = \mathrm{diag}(\vec{X}_{s,i})$ and $\vec{X}_{s,i} = (\mathbf{F}_{N_p,N} \boldsymbol{\mathcal{M}} \vec{d}_s)_i$. Here, $\mathbf{F}_{N_p,N} \subseteq \mathbf{F}_N$ and $\mathbf{F}_L \subseteq \mathbf{F}_{N,L}$ 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 $\mathbf{F}_{N_p,N}, \mathbf{F}_L$ corresponds to $m + kM\Delta k$ row of $\mathbf{F}_N, \mathbf{F}_{N,L}$ respectively.

We also define the ratio

$$\Omega \triangleq \frac{MK_p\mathfrak{Q}}{L\mathfrak{I}\mathfrak{Q}}, \tag{2.13}$$

where $MK_p\mathfrak{Q}$ number of observations (as each pilot generates M samples in frequency domain) shall be used to estimate $L\mathfrak{I}\mathfrak{Q}$ channel parameters. A least squares (LS) estimate of the channel exists if and only if the number of observations is greater than or equal

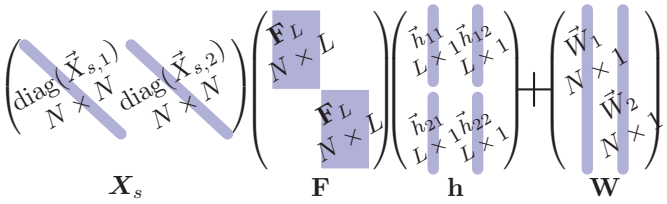


Fig. 2.2: Overview of the matrix structures for a 2×2 MIMO channel.

to the number of estimation parameters i.e. $\Omega \geq 1$. Although, LMMSE technique can still provide a reasonable estimate of the channel due to its prior knowledge of statistical channel state information (CSI).

We rearrange the expression (2.12) into matrix form as

$$\mathbf{Y} = (\boldsymbol{X}_p + \boldsymbol{X}_d)\mathbf{F}\mathbf{h} + \mathbf{W}, \text{ with } \begin{cases} \mathbf{Y}, \mathbf{W} \in \mathbb{C}^{N_p \times \mathfrak{Q}} \\ \boldsymbol{X}_p, \boldsymbol{X}_d \in \mathbb{C}^{N_p \times N_p \mathfrak{I}} \\ \mathbf{F} \in \mathbb{C}^{N_p \mathfrak{I} \times L \mathfrak{I}} \\ \mathbf{h} \in \mathbb{C}^{L\mathfrak{I} \times \mathfrak{Q}} \end{cases} \tag{2.14}$$

herein, each of the above parameters are defined as

$$\mathbf{Y} \triangleq (\vec{Y}_1, \ \cdots, \vec{Y}_q, \ \cdots, \vec{Y}_{\mathfrak{Q}}), \tag{2.14a}$$

$$\boldsymbol{X}_s \triangleq (\boldsymbol{X}_{s,1}, \ \cdots, \boldsymbol{X}_{s,i}, \ \cdots, \boldsymbol{X}_{s,\mathfrak{I}}), \tag{2.14b}$$

$$\mathbf{F} \triangleq \mathbf{I}_{\mathfrak{I}} \otimes \mathbf{F}_L, \tag{2.14c}$$

$$\mathbf{h} \triangleq \begin{pmatrix} \vec{h}_{11} & \cdots & \vec{h}_{1\mathfrak{Q}} \\ \vdots & \ddots & \vdots \\ \vec{h}_{\mathfrak{I}1} & \cdots & \vec{h}_{\mathfrak{I}\mathfrak{Q}} \end{pmatrix}, \tag{2.14d}$$

$$\mathbf{W} \triangleq (\vec{W}_1, \ \cdots, \vec{W}_q, \ \cdots, \vec{W}_{\mathfrak{Q}}). \tag{2.14e}$$

Expression (2.14) 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. 2.2 shows an example of matrix structures for $\mathfrak{I} = 2$ by $\mathfrak{Q} = 2$ antennas. In Fig. 2.2, 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 $\mathfrak{I} \times \mathfrak{Q}$ column vectors. Such matrix structure brings an advantage for mathematical analysis when vectorizing the channel matrix. It is trivial from Fig. 2.2 that $\vec{h} = \text{vec}(\mathbf{h})$ will consist of $\mathfrak{I}\mathfrak{Q} = 4$ 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 (A.4) from Appendix A.1, the corresponding vectorization of the observed matrix $\mathbf{Y}$ yields the following equation:

$$\vec{Y} = \text{vec}(\mathbf{Y}) = \tilde{\mathbf{x}}\vec{h} + \vec{W}, \tag{2.15}$$

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

For further mathematical treatments of the equalization approaches that we use later in Sec. 3.4, we also define the time domain channel covariance matrix as

$$\begin{aligned}\overline{\boldsymbol{\Sigma}}_{hh} &\triangleq \mathbb{E}[\mathtt{vec}(\boldsymbol{h})\mathtt{vec}(\boldsymbol{h})^H] \\ &= \boldsymbol{\Sigma}^1_{hh} \oplus \cdots \oplus \boldsymbol{\Sigma}^q_{hh} \oplus \cdots \oplus \boldsymbol{\Sigma}^{\mathfrak{Q}}_{hh},\end{aligned} \tag{2.16}$$

where $\boldsymbol{\Sigma}^q_{hh} = \boldsymbol{\Sigma}_{hh,1q} \oplus \cdots \oplus \boldsymbol{\Sigma}_{hh,\mathfrak{I}q}$ and $\boldsymbol{\Sigma}_{hh,iq} = \mathrm{diag}(\vec{P}_{iq})$. The MIMO matrix of CIRs $\boldsymbol{h}$ follows

$$\boldsymbol{h} = [\vec{h}_1 \ \cdots \ \vec{h}_q \ \cdots \ \vec{h}_{\mathfrak{Q}}] \in \mathbb{C}^{\mathfrak{I}L\times\mathfrak{Q}}, \tag{2.17}$$

where $\vec{h}_q = [\vec{h}^T_{1q} \ \cdots \ \vec{h}^T_{iq} \ \cdots \ \vec{h}^T_{\mathfrak{I}q}]^T$.

Since each transmit signal is separated by a CP, we assume the individual channels between the antenna pairs i-q are circulant. Therefore, under the assumption of perfect time and frequency synchronization, the receive signal is given by

$$\begin{aligned}\vec{y}_{\mathfrak{Q}N} &= [\vec{y}^T_{1,N} \ \cdots \ \vec{y}^T_{q,N} \ \cdots \ \vec{y}^T_{\mathfrak{Q},N}]^T \\ &= \boldsymbol{\mathcal{C}}_N\overline{\boldsymbol{\mathcal{M}}}\vec{d} + \vec{w}_{\mathfrak{Q}N},\end{aligned} \tag{2.18}$$

where $\vec{y}^T_{q,N}$ is the time domain Rx signal on antenna q, $\vec{d} = [\vec{d}^T_1 \ \cdots \ \vec{d}^T_{\mathfrak{I}}]^T$, $\overline{\boldsymbol{\mathcal{M}}} = \boldsymbol{I}_{\mathfrak{I}} \otimes \boldsymbol{\mathcal{M}}$, $\vec{w}_{\mathfrak{Q}N} \in \mathbb{C}^{\mathfrak{Q}N}$ is the vector of AWGN samples, $\boldsymbol{\mathcal{C}}_N \in \mathbb{C}^{\mathfrak{Q}N\times\mathfrak{I}N}$ is the MIMO version of $\mathfrak{Q}\times\mathfrak{I}$ circulant channels $\boldsymbol{\mathcal{C}}_{N,qi} \in \mathbb{C}^{N\times N}$ with the impulse responses $\vec{h}_{iq}$ on their first column, i.e.

$$\boldsymbol{\mathcal{C}}_N = \begin{pmatrix} \boldsymbol{\mathcal{C}}_{N,11} & \cdots & \boldsymbol{\mathcal{C}}_{N,1\mathfrak{I}} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\mathcal{C}}_{N,\mathfrak{Q}1} & \cdots & \boldsymbol{\mathcal{C}}_{N,\mathfrak{Q}\mathfrak{I}} \end{pmatrix}. \tag{2.19}$$

For future purpose, we also define the matrix

$$\boldsymbol{y}_N \triangleq [\vec{y}_{1,N} \ \cdots \ \vec{y}_{\mathfrak{Q},N}] = \mathtt{vec}^{-1}(\vec{y}_{\mathfrak{Q}N}). \tag{2.20}$$

In frequency domain, the circulant channels become diagonal. Therefore, the received signal becomes:

$$\vec{Y}_{\mathfrak{Q}N} = \overline{\boldsymbol{H}}\bar{\boldsymbol{F}}\overline{\boldsymbol{\mathcal{M}}}\vec{d} + \vec{W}_{\mathfrak{Q}N}, \tag{2.21}$$

where $\vec{W}_{\mathfrak{Q}N} = (\boldsymbol{I}_{\mathfrak{Q}} \otimes \boldsymbol{F}_N)\vec{w}_{\mathfrak{Q}N}$, $\bar{\boldsymbol{F}} = \boldsymbol{I}_{\mathfrak{I}} \otimes \boldsymbol{F}_N$, $\overline{\boldsymbol{H}}$ is the $\mathfrak{Q}\times\mathfrak{I}$ matrix of frequency domain MIMO channels $\boldsymbol{H}_{qi} = \mathrm{diag}(\sqrt{N}\boldsymbol{F}_{N,L}\vec{h}_{i,q})$. The matrix $\boldsymbol{F}_{N,L} \subseteq \boldsymbol{F}_N$ comprises the L columns of the DFT matrix $\boldsymbol{F}_N$ associated to the delays $\tilde{n}_\ell$ of the L taps.

2.4 Classical and Bayesian Channel Estimation in MIMO OFDM Systems

In CP-OFDM with $\boldsymbol{\mathcal{M}}_{\mathrm{OFDM}} = \boldsymbol{F}^H_N$, given that the channel is bock-fading, the pilot and data signals become orthogonal to each other, i.e. $\boldsymbol{X}_p \circ \boldsymbol{X}_d = \mathbf{0}_{N_p\times N_p\mathfrak{I}}$. Thus, under

the assumption of perfect time and frequency synchronization and ideal front-ends, the measurements matrix $\mathbf{Y}$ simplifies to

$$\mathbf{Y}_{\text{cp-ofdm}} = \boldsymbol{X}_p\mathbf{F}\mathbf{h} + \mathbf{W}, \tag{2.22}$$

and accordingly, $\mathbf{Y}_{\text{cp-ofdm}}$ solely depends on the known observation matrix $\boldsymbol{X}_p\mathbf{F}$, without any influence from $\boldsymbol{X}_d$.

Classical Estimation. The LS estimate of the channel follows the form [BLM03],[Kay93]:

$$\hat{\mathbf{h}}_{\text{LS,ofdm}} = \boldsymbol{Q}_{\text{LS}}\mathbf{Y} = \mathbf{h} + \mathbf{E}_{\text{ofdm}}, \tag{2.23}$$

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

$$\widehat{\mathbf{H}}_{\text{LS}} = (\mathbf{I}_{\mathfrak{I}} \otimes \mathbf{F}_{N,L})\hat{\mathbf{h}}_{\text{LS}} \tag{2.24}$$

Because of the orthogonality of pilots and data symbols, the *noise-enhancement-factor* becomes $\mathbf{E}_{\text{ofdm}} = \boldsymbol{Q}_{\text{LS}}\mathbf{W}$, which yields the mean squared error (MSE)

$$\begin{aligned} \text{MSE}_{\text{LS,ofdm}} &= \mathbb{E}\left[\|\hat{\mathbf{h}}_{\text{LS}} - \mathbf{h}\|^2\right] \\ &= \frac{\sigma_w^2}{\Delta k}\text{Tr}\left(\mathbf{I}_{\mathfrak{Q}} \otimes (\boldsymbol{Q}_{\text{LS}}^H\boldsymbol{Q}_{\text{LS}})\right), \end{aligned} \tag{2.25}$$

where, Δk and σ_w^2 denote the pilot-subcarrier-spacing and white-noise variance, respectively. Since (2.23) satisfies the condition $\mathbb{E}[\mathbf{E}_{\text{ofdm}}] = \mathbf{0}_{N\times\mathfrak{Q}}$, $\hat{\mathbf{h}}_{\text{LS,ofdm}}$ is an unbiased estimator.

Bayesian Estimation. The LMMSE estimation calculates the coefficients of a linear filter aiming at minimum mean square error. In accordance with (2.14) and the corresponding vectorization in (2.15), we formally have:

$$\vec{\hat{h}}_{\text{LMMSE,ofdm}} = \boldsymbol{\Sigma}_{hY}\boldsymbol{\Sigma}_{YY,\text{ofdm}}^{-1}\vec{Y}_{\text{cp-ofdm}}, \tag{2.26}$$

with the matrices defined as

$$\boldsymbol{\Sigma}_{YY,\text{ofdm}} = \widetilde{\mathbf{x}}_p\boldsymbol{\Sigma}_{\mathbf{hh}}\widetilde{\mathbf{x}}_p^H + \sigma_w^2\mathbf{I}_{N\mathfrak{Q}}, \in \mathbb{C}^{N_p\mathfrak{Q}\times N_p\mathfrak{Q}} \tag{2.27}$$

$$\boldsymbol{\Sigma}_{hY} = \boldsymbol{\Sigma}_{\mathbf{hh}}\widetilde{\mathbf{x}}_p^H, \quad \in \mathbb{C}^{L\mathfrak{I}\mathfrak{Q}\times N_p\mathfrak{Q}} \tag{2.28}$$

where $\widetilde{\mathbf{x}}_p = (\mathbf{I}_{\mathfrak{Q}}\otimes\boldsymbol{X}_p\mathbf{F})$, and $\vec{Y}_{\text{cp-ofdm}} = \text{vec}(\mathbf{Y}_{\text{cp-ofdm}})$. Note that $\vec{\hat{h}}_{\text{LMMSE}}$ is a column vector containing $\mathfrak{I}\mathfrak{Q}$ individual column vectors of size L, associated to the LMMSE estimates of the individual CIRs. In addition, the matrices $\boldsymbol{\Sigma}_{YY,\text{ofdm}}$ and $\boldsymbol{\Sigma}_{hY}$ are block-diagonal with $\mathfrak{Q}$ matrices on their diagonal, each with $N_p \times N_p$ and $L\mathfrak{I} \times N_p$ dimensions, respectively.

Hence, inversion of $\mathbf{\Sigma}_{YY,\text{ofdm}}$ comprises of an $N_p \times N_p$ matrix inversion, which is then $\mathfrak{Q}$ times copied onto the diagonal blocks of $\mathbf{\Sigma}_{YY,\text{ofdm}}$.

The resulting MSE performance of the LMMSE estimation follows:

$$\text{MSE}_{\text{LMMSE,ofdm}} = \text{Tr}\left(\mathbf{\Sigma_{HH}} - \widehat{\mathbf{\Sigma}}_{\mathbf{HH}}\right) \tag{2.29}$$

with

$$\mathbf{\Sigma_{HH}} = (\mathbf{I}_{\mathfrak{Q}} \otimes \mathbf{F})\mathbf{\Sigma_{hh}}(\mathbf{I}_{\mathfrak{Q}} \otimes \mathbf{F})^H \tag{2.30}$$

$$\widehat{\mathbf{\Sigma}}_{\mathbf{HH}} = (\mathbf{I}_{\mathfrak{Q}} \otimes \mathbf{F})\mathbf{\Sigma}_{hY}\mathbf{\Sigma}^{-1}_{YY,\text{ofdm}}\mathbf{\Sigma}^H_{hY}(\mathbf{I}_{\mathfrak{Q}} \otimes \mathbf{F})^H, \tag{2.31}$$

which is *clear* from any interference.

Resorting to the matrix identity (A.5), the LMMSE estimation (2.26) can alternatively be written as

$$\vec{\hat{h}}_{\text{LMMSE,ofdm}} = (\sigma_w^2\mathbf{\Sigma_{hh}} + \widetilde{\mathbf{x}}_p^H\widetilde{\mathbf{x}}_p)^{-1}\widetilde{\mathbf{x}}_p^H\vec{Y}_{\text{cp-ofdm}}, \tag{2.32}$$

which requires a matrix inversion with dimension $\mathfrak{I}L \times \mathfrak{I}L$, instead of $N_p \times N_p$ as in (2.26). Also note that in (2.32), if $\sigma_w^2 \neq 0$, we have $\mathbb{E}[\vec{\hat{h}}_{\text{LMMSE,ofdm}}] \neq \vec{h}$, which means the estimator $\vec{\hat{h}}_{\text{LMMSE,ofdm}}$ is biased.

Summarizing the above two estimation techniques, one must consider a trade-off between the two choices. The LS approach is an unbiased estimation which does not require any probabilistic assumptions on the channel statistics, and therefore, it is being widely used due to its ease of implementation. Nonetheless, the LMMSE estimation exploits the a-priori knowledge of channel statistics, in order to improve the estimation quality by minimizing the MSE. The LMMSE estimation has a higher complexity and it is more costly to be implemented, because e.g. (2.32) needs to calculate the coefficients of the estimator for any σ_w^2 value, whereas in LS approach, $\boldsymbol{Q}_{\text{LS}}$ can be calculated offline. One should also note that at very high SNR values, one shall expect the LMMSE and LS estimations to have identical quality, i.e. $\lim_{\sigma_w^2 \to 0} \vec{\hat{h}}_{\text{LMMSE,ofdm}} = \text{vec}(\hat{\mathbf{h}}_{\text{LS,ofdm}})$.

2.5 UW-Based Transmission in SISO Systems

In a unique word (UW) based system [WMS+02], [DGE00], instead of a CP of random nature, a deterministic sequence is added as a prefix and also as a suffix around the transmission block (e.g. see Fig. 2.3). Therefore, the symbol duration reduces from $N + N_{\text{cp}}$ to N samples[1]. Here, N_{cp} is the number of time-domain samples of the CP, and N is the data fast Fourier transform (FFT) size. Similar to CP-based systems, if the UW length N_u is longer than the channel delay spread with length L, and also the channel remains constant during the block duration of N samples, the receiver can interpret a circular channel

[1] In the SoA approaches, pilot insertion for the CP-based system and its corresponding overhead versus UW has not been taken into account. We will discuss such overhead analysis in Chapter 4.

transfer function and perform FFT based channel estimation and equalization approaches. Moreover, since UW sequences are deterministic and known, per-block synchronization can be achieved.

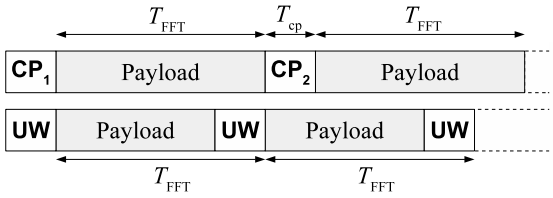


Fig. 2.3: UW vs. CP design according to [DGE00].

The state-of-the-art data-aided synchronization approaches are mostly preamble based, and they do not specifically address MIMO UW based systems, and moreover, they mostly assume either low mobility scenarios or totally time-invariant systems in their simulations [ZZZW05],[MS01]. On the other hand, the synchronization techniques that have been proposed for UW-based systems are mostly for single-input-single-output systems, and they also consider time-invariant channel conditions [DGE00], [HWH03], [Abo12]. In [ZFF17], embedding a Zadoff-Chu sequence within the GFDM signal, a UW-based frame structure has been designed for vehicular communications. Nevertheless, all the above related works were focusing on SISO systems and their approaches are not straightforwardly applicable to MIMO multi-carrier systems, because in MIMO, the UW sequences being transmitted from each antenna must accommodate a high quality MIMO synchronization [MS01] as well as MIMO channel estimation. We will address these challenges in Chapter 4 by considering a single frame design that will be jointly used for MIMO synchronization and MIMO channel estimation of UW-based systems.

From a channel estimation perspective, if the UW length N_{u} is set to L samples of the channel delay spread length, the UW observations would always be interference-limited, because taking the FFT of the frame over N samples of T_{FFT} duration, payload becomes overlapped with UW sequences in frequency domain. In [CSBM06], authors suggested that the UW length must be at least twice of the channel length (i.e. [CSBM06] $N_u = 2L$) to achieve an interference-free observation at the second half of the UW. Although, in order to use the energy of the UW from the first half, they suggested an iterative interference cancellation approach. Unfortunately, the performance of the approach in [CSBM06] was not compared to an equivalent CP-based system and moreover, it is also only applicable to low mobility channel conditions with SISO systems. In Chapter 4, after proposing a MIMO UW-based frame structure and their corresponding UW sequences, we develop further CE techniques for MIMO UW-based systems in frequency-selective and fast-fading channels.

The SoA equalization techniques [HOH11], [DdL17] use the circularity of the channel over *Payload-UW* slots. In this case, given the payload size N_d and under the assumption of perfect channel knowledge, finding an optimal Bayesian data estimator requires $N_d \times N_d$

matrix inversion [HOH11]. Using the special matrix structures of the precoder proposed in [HOH11], the complexity could be reduced to $N_u \times N_u$ matrix inversion for SISO OFDM systems. However, to the best of our knowledge, there is no low complexity UW-based equalization approach for MIMO non-orthogonal multi-carriers.

Now, for the sake of formulating the problem of employing UW sequences for non-orthogonal modulation matrix $\boldsymbol{A}_{N_d} \in \mathbb{C}^{N_d \times N_d}$, consider the following linear model in a SISO system

$$\begin{aligned} \vec{Y}_N &= \boldsymbol{H}\boldsymbol{F}_N[\vec{x}_d^T \ \vec{x}_u^T]^T + \vec{W} \\ &= \boldsymbol{H}\boldsymbol{F}_N \begin{bmatrix} \boldsymbol{A}_{N_d} & \\ & \boldsymbol{I}_{N_u} \end{bmatrix} \begin{bmatrix} \vec{d}_d \\ \vec{x}_u \end{bmatrix} + \vec{W} \end{aligned} \tag{2.33}$$

where $\boldsymbol{H}$ is a diagonal channel, $\vec{x}_u \in \mathbb{C}^{N_u}$ denotes a UW sequence, and $\vec{x}_d = \boldsymbol{A}_{N_d}\vec{d}_d \in \mathbb{C}^{N_d}$ denotes the time-domain payload signal. Focusing on data part we have

$$\vec{Y}_{N,d} = \boldsymbol{H}\boldsymbol{F}_{N,N_d}\boldsymbol{A}_{N_d}\vec{d}_d + \vec{W} \tag{2.34}$$

where $\boldsymbol{F}_{N,N_d}$ denotes the first N_d columns of $\boldsymbol{F}_N$. Here, since an N-size DFT is being applied to an $N_d \times N_d$ dimensional matrix $\boldsymbol{A}_{N_d}$, an optimum—from error probability perspective—low complexity receiver which uses the particular structures of $\boldsymbol{A}_{N_d}$ whether in time, DFT or DZT domains is not straightforwardly achievable. Therefore, finding an optimal $\boldsymbol{\mathcal{Z}}_{\text{opt}}$ for MMSE based estimation of $\vec{d}_d$, i.e. joint *channel-equalization-and-demodulation*, requires $N \times N$ matrix inversion. Extending such a system to MIMO will further increase the complexity to $\mathfrak{I}N \times \mathfrak{I}N$ matrix inversion, which becomes entirely impracticable for real-time implementation, when $\mathfrak{I}$ is large. Therefore, in Chapter 4, we propose a UW-free joint channel-equalization-and-demodulation approach that is a general close-to-optimal form which can be applied to any MIMO single/multi-carrier system.

2.6 Summary

In this chapter, we presented the fundamentals of MIMO channel estimation for wireless transmission in linear systems context. The fundamentals presented in this chapter would be considered as benchmark for the algorithm developments in Chapters 3 and 4. In addition to the MIMO CP-OFDM channel estimation techniques, we have also reviewed the pros and cons of the SoA UW-based systems. In Chapter 4, we will extend the SoA UW-based SISO systems to MIMO, and we will also show that our proposed algorithm developments outperform the SoA CP- and UW-based systems.

Chapter 3

Channel Estimation for MIMO Non-Orthogonal Waveforms

An essential part of future wireless communications is MIMO transmission. Compared to straightforward and low complexity implementation of OFDM, the receiver design becomes more challenging for MIMO non-orthogonal multi-carriers because the receiver has to deal with multiple dimensions of interference, i.e. ISI, ICI and IAI. The multi-dimensional interference influences the received pilots, in which, it poses challenges on CE and subsequent equalization (EQ) of the Rx signal, particularly, in a broadband type of communication where the channel is doubly dispersive.

The main outcomes of our contributions in this chapter have been published in [2], [3], [9], [8], [7], and [6], where the following aspects concerning MIMO non-orthogonal channel estimation have been addressed.

- Considering rich multi-path fading channels, two well-known estimation techniques, namely LS and LMMSE are here tailored for pilot-aided MIMO-GFDM channel estimation [3], [9].

- We theoretically assess the choice of pilot design in GFDM and in this regard, we propose a new low complexity CE technique as an alternative to the classical LS estimation. We will also analyze the impact of pilots pattern on self-interference of GFDM system [8].

- Modifying the low complexity GFDM transmitter based on sparse frequency domain processing, we propose a general approach in order to insert orthogonal pilots [7],[3] and subsequently, to achieve an interference-free channel estimation for MIMO-GFDM. Even though, the approach is applied to GFDM, it is also applicable to different non-orthogonal multi-carriers which use frequency domain channel estimation.

- The well time- and frequency-localization of the pilots in non-orthogonal waveforms allows to reuse their energy from CP. Taking advantage of this feature in non-orthogonal multi-carriers, we derive an iterative approach for joint channel estimation and equalization of MIMO systems [2]. Next, considering the block-circularity structure of GFDM, we derive a low complexity CE technique via sequential LMMSE estimation.

- We extend the pilot- and CP-aided channel estimation of non-orthogonal systems to time-varying scenarios where we use *Wiener-Hopf* filters for joint *smoothing-interpolation-prediction* of the channel responses [6].

3.1 Classical and Bayesian Channel Estimation in MIMO GFDM Systems

Pilot Allocation. Consider the GFDM model (2.3) where the vector $\vec{d} = \vec{d}_p + \vec{d}_d$ is generated from the summation of pilots sequence $\vec{d}_p \in \mathbb{C}^N$ and data vector $\vec{d}_d \in \mathbb{C}^N$. 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 the elements in $\vec{d}_p$ which are the positions of data samples from $\vec{d}_d$ are kept zero. Thus, each time-frequency resource element is associated to either pilots or data, i.e. $\vec{d}_p \circ \vec{d}_d = \vec{0}_N$. The data vector $\vec{d}_d$ comprises complex valued data symbols from a modulation alphabet, e.g. 2^μ-QAM. Fig. 3.1 shows an example of pilot positions in the time-frequency resource grid, where a single CP protects multiple subsymbols. It is plain that the pilot insertion with pilot-subcarrier spacing Δk, reduces the effective rate by the factor

$$\begin{aligned} \eta_{\text{CE}} &\triangleq \frac{\text{no. of data samples}}{\text{no. of total samples}} = (N - K_p)/N \\ &= 1 - \frac{\lfloor K/\Delta k \rfloor}{MK}. \end{aligned} \tag{3.1}$$

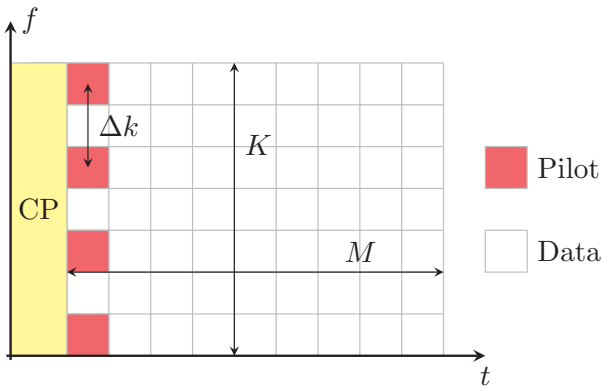


Fig. 3.1: Block-type pilots in time-frequency grid of GFDM for $M_p = 2, \Delta k = 2$.

3.1.1 MIMO LS Channel Estimation

Similar to (2.23), the LS estimator of $\mathbf{h}$ is found by minimizing $\|\mathbf{Y} - \boldsymbol{X}_p\mathbf{F}\mathbf{h}\|^2$. This yields

$$\hat{\mathbf{h}}_{\mathrm{LS}} = \boldsymbol{Q}_{\mathrm{LS}}\mathbf{Y} = \mathbf{h} + \mathbf{E}. \tag{3.2}$$

In contrast to (2.25), the above estimation yields the following interference and noise enhancement terms:

$$\mathbf{E} = \boldsymbol{Q}_{\mathrm{LS}}\boldsymbol{\Psi} + \boldsymbol{Q}_{\mathrm{LS}}\mathbf{W}. \tag{3.3}$$

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

Accordingly, the result of the MSE calculation follows:

$$\begin{aligned}
\mathrm{MSE}_{\mathrm{LS}} &= \mathbb{E}\left[\|\hat{\mathbf{h}}_{\mathrm{LS}} - \mathbf{h}\|^2\right] \\
&= \frac{1}{\Delta k}\mathrm{Tr}\left(\left(\mathbf{I}_{\mathfrak{Q}} \otimes (\boldsymbol{Q}_{\mathrm{LS}}^H\boldsymbol{Q}_{\mathrm{LS}})\right)\boldsymbol{\Sigma}_{\Psi\Psi}\right) \\
&\qquad + \frac{\sigma_w^2}{\Delta k}\mathrm{Tr}\left(\mathbf{I}_{\mathfrak{Q}} \otimes (\boldsymbol{Q}_{\mathrm{LS}}^H\boldsymbol{Q}_{\mathrm{LS}})\right),
\end{aligned} \tag{3.4}$$

where σ_w^2 is the noise variance. In the above expression, the first term is due to the interference from data subsymbols on the pilot subcarriers and second term is the noise enhancement. Then, we compute the covariance matrix of the interference term as

$$\begin{aligned}
\boldsymbol{\Sigma}_{\Psi\Psi} &= \mathbb{E}\left[\mathrm{vec}(\boldsymbol{X}_d\mathbf{F}\mathbf{h})\mathrm{vec}(\boldsymbol{X}_d\mathbf{F}\mathbf{h})^H\right] \\
&= \mathbb{E}_{\boldsymbol{X}_d}\left[(\mathbf{I}_{\mathfrak{Q}} \otimes \boldsymbol{X}_d\mathbf{F})\,\mathbb{E}_{\mathbf{h}}\left[\vec{h}\vec{h}^H|\boldsymbol{X}_d\right](\mathbf{I}_{\mathfrak{Q}} \otimes \boldsymbol{X}_d\mathbf{F})^H\right] \\
&= \mathbb{E}_{\boldsymbol{X}_d}\left[(\mathbf{I}_{\mathfrak{Q}} \otimes \boldsymbol{X}_d\mathbf{F})\,\boldsymbol{\Sigma}_{\mathbf{hh}}\,(\mathbf{I}_{\mathfrak{Q}} \otimes \boldsymbol{X}_d\mathbf{F})^H\right].
\end{aligned} \tag{3.5}$$

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

$$\mathbf{I}_{\mathfrak{Q}} \otimes \boldsymbol{X}_d\mathbf{F} = \mathrm{blkdiag}([\boldsymbol{X}_{d,1}\mathbf{F}_L, \cdots, \boldsymbol{X}_{d,\mathfrak{I}}\mathbf{F}_L], \cdots, [\boldsymbol{X}_{d,1}\mathbf{F}_L, \cdots, \boldsymbol{X}_{d,\mathfrak{I}}\mathbf{F}_L]), \tag{3.6}$$

$$\boldsymbol{\Sigma}_{\mathbf{hh}} = \mathrm{blkdiag}(\boldsymbol{\Sigma}_{h_{11}}, \cdots, \boldsymbol{\Sigma}_{h_{\mathfrak{I}1}}, \cdots, \boldsymbol{\Sigma}_{h_{(\mathfrak{I}-1)\mathfrak{Q}}}, \boldsymbol{\Sigma}_{h_{\mathfrak{I}\mathfrak{Q}}}), \tag{3.7}$$

where $\boldsymbol{\Sigma}_{h_{iq}} \in \mathbb{R}^{L\times L}$ is the diagonal covariance matrix of CIR, computed as

$$\begin{aligned}
\boldsymbol{\Sigma}_{h_{iq}} &= \mathbb{E}\left[\vec{h}_{iq}\vec{h}_{iq}^H\right] \\
&= \mathrm{diag}(\vec{P}_{iq}).
\end{aligned} \tag{3.8}$$

The product of (3.6), (3.7) and the hermitian conjugate of (3.6) will then provide a block diagonal structure for the interference covariance matrix $\boldsymbol{\Sigma}_{\Psi\Psi}$ as expressed in (3.5). This is due to the fact that independent Rayleigh fading has been considered for the individual channels (see Sec. 2.3). As a result, it is possible to perform the computations separately

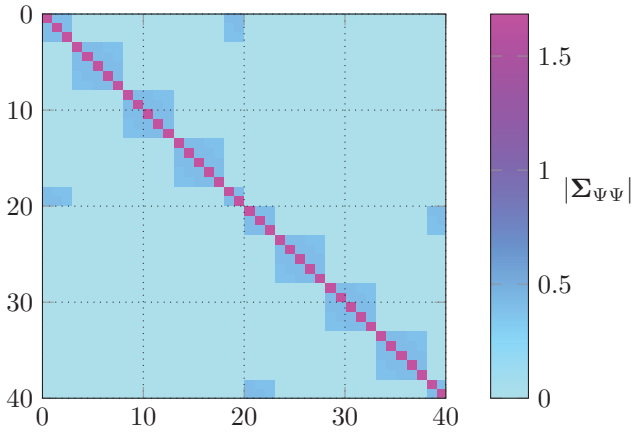


Fig. 3.2: Example of $\mathbf{\Sigma}_{\Psi\Psi}$ structure for a 2×2 MIMO channel with $K = 16$, $M = 5$ and $\Delta k = 4$. The matrix is block diagonal according to $\mathfrak{Q} = 2$ Rx antennas. Furthermore, due to the block-circulant structure of $\mathbf{A}$ matrix, the individual blocks are banded matrices with periodic boundary conditions. Note that the non-zero terms at the off-diagonal edges would vanish if guard subcarriers are inserted.

for the individual blocks. Hence, via the matrix identity A.6, for the Tx antenna i and Rx antenna q we have [RC91]:

$$\begin{aligned} \mathbf{\Sigma}_{\Psi\Psi_{iq}} &= \mathbb{E}_{\boldsymbol{X}_{d,i}}\left[\boldsymbol{X}_{d,i}\mathbf{F}_L\mathbb{E}_h\left[\vec{h}_{iq}\vec{h}_{iq}^{H}|\boldsymbol{X}_{d,i}\right]\mathbf{F}_L^H\boldsymbol{X}_{d,i}^H\right] \\ &= \mathbf{F}_L\mathrm{diag}(\vec{P}_{iq})\mathbf{F}_L^H \circ \mathbf{\Sigma}_{X_dX_d,i}. \end{aligned} \tag{3.9}$$

Furthermore, the covariance matrix of data is being calculated as

$$\begin{aligned} \mathbf{\Sigma}_{X_dX_d,i} &= \mathbb{E}[(\mathbf{F}_{N_p,N}\mathbf{A}\vec{d}_d)_i(\mathbf{F}_{N_p,N}\mathbf{A}\vec{d}_d)_i^H] \\ &= (\mathbf{F}_{N_p,N}\mathbf{A}\mathrm{diag}(\vec{\sigma}_d^2)\mathbf{A}^H\mathbf{F}_{N_p,N}^H)_i, \end{aligned} \tag{3.10}$$

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

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

$$\mathbf{\Sigma}_{\Psi\Psi(q)} = \sum_{i=1}^{\mathfrak{I}}(\mathbf{F}_L\mathrm{diag}(\vec{P}_{iq})\mathbf{F}_L^H) \circ \mathbf{\Sigma}_{X_dX_d,i}. \tag{3.11}$$

The full interference covariance matrix follows:

$$\mathbf{\Sigma}_{\Psi\Psi} = \mathrm{blkdiag}(\mathbf{\Sigma}_{\Psi\Psi(q=1)}, \mathbf{\Sigma}_{\Psi\Psi(q=2)}, \cdots, \mathbf{\Sigma}_{\Psi\Psi(q=\mathfrak{Q})}). \tag{3.12}$$

Fig. 3.2 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.

3.1.2 MIMO LMMSE Channel Estimation

Analogously to (2.26), we calculate the coefficients of the LMMSE filter by

$$\hat{\vec{h}}_{\mathrm{LMMSE}} = \mathbf{\Sigma}_{hY}\mathbf{\Sigma}_{YY}^{-1}\vec{Y}, \tag{3.13}$$

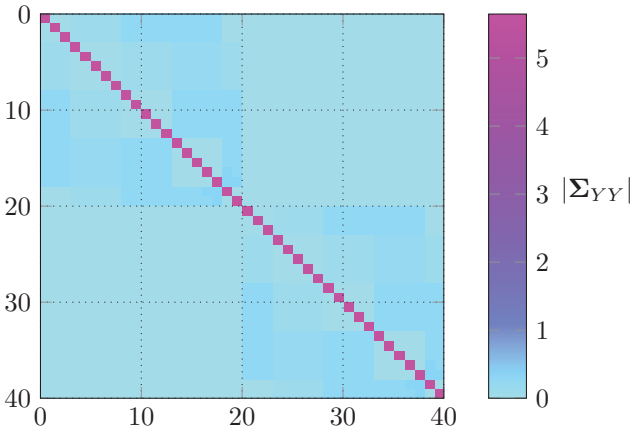


Fig. 3.3: Example of $\mathbf{\Sigma}_{YY}$ structure for a 2×2 MIMO channel with $K = 16$, $M = 5$ and $\Delta k = 4$. The matrix is block-diagonal according to $\mathfrak{Q} = 2$ Rx antennas. Here, although in individual matrix blocks, the diagonal elements are significantly larger than the off-diagonal elements, they are not diagonally dominant matrices yet.

in which, $\mathbf{\Sigma}_{YY}$ follows

$$\mathbf{\Sigma}_{YY} = \tilde{\mathbf{x}}_p \mathbf{\Sigma}_{\mathbf{hh}} \tilde{\mathbf{x}}_p^H + \mathbf{\Sigma}_{\Psi\Psi} + \sigma_w^2 \mathbf{I}_{N\mathfrak{Q}}. \tag{3.14}$$

Note that here, the MSE performance of LMMSE estimation

$$\mathrm{MSE}_{\mathrm{LMMSE}} = \mathrm{Tr}\left(\mathbf{\Sigma}_{\mathbf{HH}} - (\mathbf{I}_{\mathfrak{Q}} \otimes \mathbf{F})\mathbf{\Sigma}_{hY}\mathbf{\Sigma}_{YY}^{-1}\mathbf{\Sigma}_{hY}^H(\mathbf{I}_{\mathfrak{Q}} \otimes \mathbf{F})^H\right) \tag{3.15}$$

also includes the interference covariance matrix $\mathbf{\Sigma}_{\Psi\Psi}$.

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 (3.12). 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 (3.12) is required only once. In addition, note that although the complexity of the matrix inversion $\mathbf{\Sigma}_{YY}^{-1}$ is $O\left((N_p\mathfrak{Q})^3\right)$, the matrix $\mathbf{\Sigma}_{YY}$ is constructed in form of block diagonal with $\mathfrak{Q}$ blocks. Hence, the complexity reduces to $O(\mathfrak{Q}N_p^3)$ due to $\mathfrak{Q}$ separate matrix inversions. Fig. 3.3 shows an example of $\mathbf{\Sigma}_{YY}$ for a 2×2 MIMO channel and thus 2 individual diagonal blocks which can be inverted separately.

3.1.3 Simulation 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 derived in Sec. 3.1.1 and Sec. 3.1.2. Here, we consider the OFDM CE techniques (i.e. Sec. 2.4)

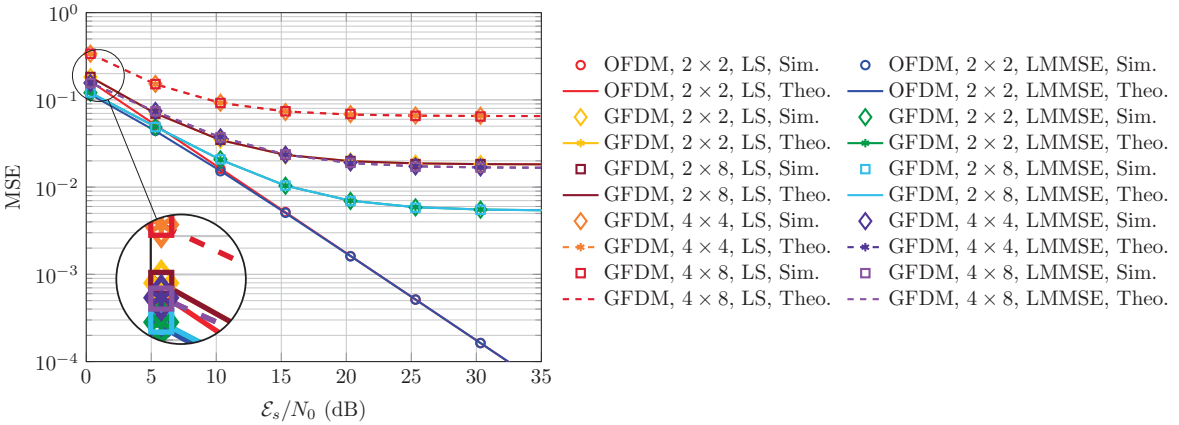


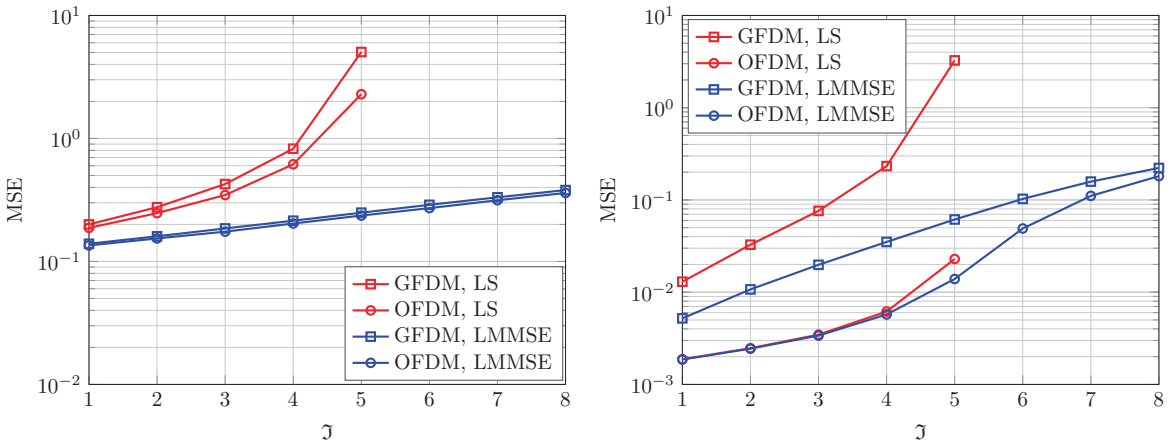
Fig. 3.4: MSE results of channel estimation vs. SNR for simulation and theoretical calculations in Rayleigh fading MIMO channel with pilot spacing of $\Delta k = 2$ and $K = 128$ subcarriers.

as benchmark. 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 $\mathfrak{I} \in \{2, 3, 4\}$ and $\mathfrak{Q} \in \{2, 3, \cdots, 8\}$ antennas. The channel gains are considered to have Rayleigh distribution, hence, the PDP is exponentially distributed with $L = 9$ independent fading gains. Furthermore, the channel realizations are assumed to remain static during the transmission of a GFDM block. A single block of GFDM signal contains $M = 7$ subsymbols and it is filtered by a Raised-Cosine (RC)[1] pulse with roll-off factor $\alpha = 0.3$. For comparison purpose, we configure OFDM to have $K_{\text{OFDM}} = MK$ 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.

Fig. 3.4 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 on 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.

[1] Thorough analysis of the choice of pulse-shaping filter has been studied in [MMGF14]. Indeed, for a given non-zero roll-off factor and thus, similar amount of out-of-band suppression, the RC filter generates the least amount of self-interference variance in comparison with root-RC, Xia and Gaussian pulses.



(a) $\mathcal{E}_s/N_0 = 0$ dB **(b)** $\mathcal{E}_s/N_0 = 20$ dB

Fig. 3.5: MSE vs. no. of Tx-Rx antennas with $K = 96, \Delta k = 2, L = 9, \mathfrak{Q} = \mathfrak{I}$. Here, LS estimators for $\mathfrak{I} > 5$ do not exist, because number of estimation parameters become larger than the number of observations.

Doubling the number of Rx antennas, we also double the number of channels, while the ratio Ω—defined in (2.13)—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. 3.5), because, by linearly increasing the number of Tx-Rx antennas, the number of channel taps increases quadratically and thus, the parameter Ω—defined in (2.13)—decreases, leading to estimation performance degradation. Note that for the specified combination of $K, \Delta k$ & L in Fig. 3.5, the MSE curves for LS estimation with $\mathfrak{I} > 5$ do not exist for OFDM, since the parameter Ω becomes smaller than one. Furthermore, the error gets too large also for GFDM when $\Omega < M$, because *a*) The energy of the pilots in GFDM spread over M frequency bins (i.e. $\frac{\mathcal{E}_s}{M}$ energy per bin), and *b*) The MK_p observations (associated to K_p pilots) are not equispaced in frequency. Formally, such condition yields the norm of the LS estimator to become too large, and thus, it significantly enhances the noise and interference terms. In addition, Fig. 3.5a and 3.5b 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. 3.6 compares the MSE for different Δk values. The pilot spacing of $\Delta k = 2$ in GFDM corresponds to $\Delta k_{\text{OFDM}} = 2M$ 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. 3.6 is the overlapping of channel estimation MSE in OFDM system with the MSE in GFDM due to noise only. However, because of

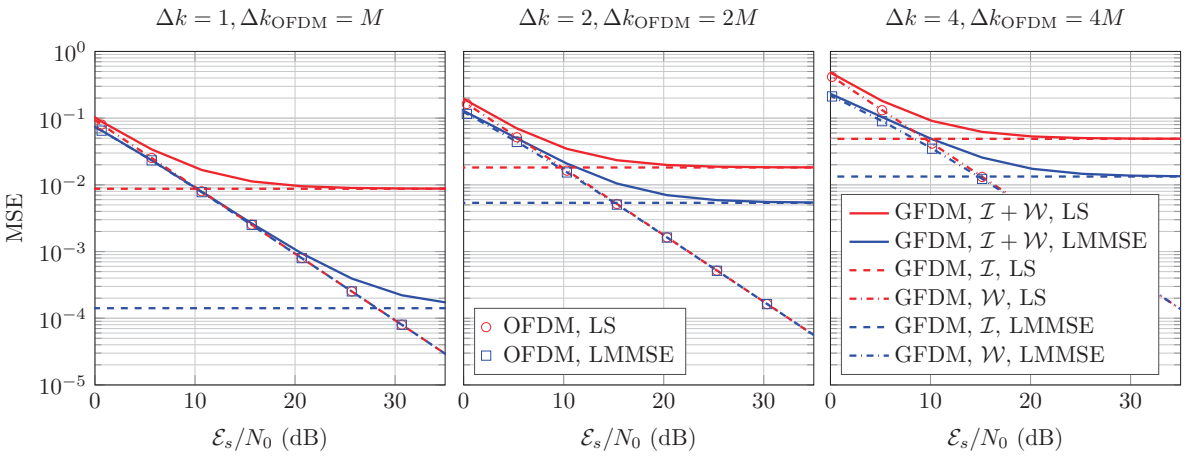


Fig. 3.6: MSE of channel estimation vs. SNR in Rayleigh fading 2×2 MIMO channel with $K = 128$ subcarriers and $M = 7$ subsymbols. $\mathcal{I}$ represents the MSE due to the interference from data, while $\mathcal{W}$ stands for the MSE due to noise only.

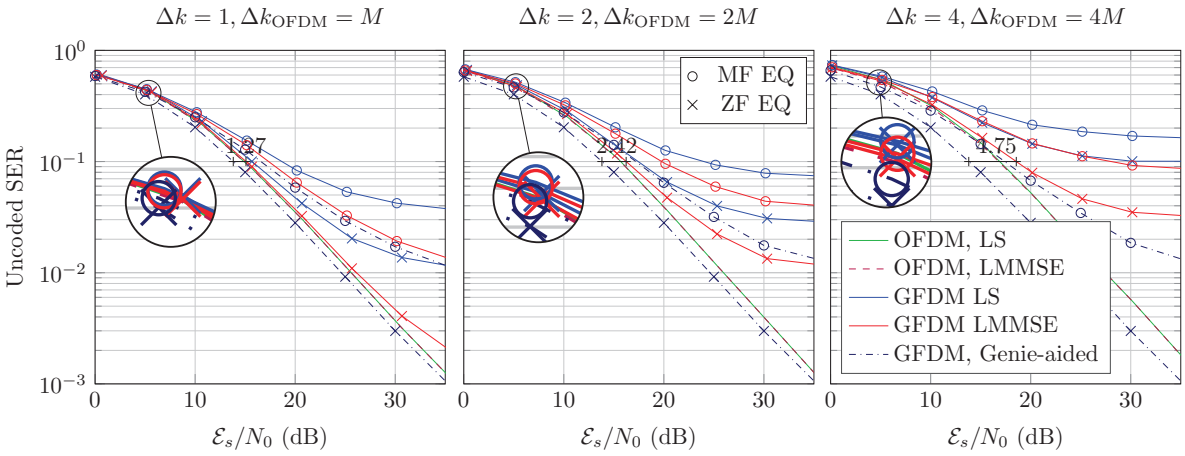


Fig. 3.7: Symbol Error Rate performance of GFDM under the influence of channel estimation in Rayleigh fading 2×2 MIMO channel with $K = 128$ subcarriers, $M = 7$ subsymbols and 16-QAM symbol mapping.

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—defined in (3.1)—for $\Delta k = \{1, 2, 4\}$ is reduced by $\eta_{\text{CE}} = \{86\%,\ 93\%,\ 96\%\}$, respectively.

The simulation results for uncoded SER performance of GFDM under the influence of channel estimation is provided in Fig. 3.7. 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 knowl-

edge of the CSI due to a genie at no cost. Moreover, the GFDM receiver employs either a matched filter (MF)—shown by o marks—approach or zero forcing (ZF)—shown by x marks—to equalize the $\boldsymbol{A}$ matrix, i.e. $\boldsymbol{A}^H$ and $\boldsymbol{A}^{-1}$, respectively. Here again, we notice that the GFDM SER performance has an error floor at high SNRs due to the error floor in channel estimation, whereas, for OFDM, SER decreases proportionally with the SNR. We also note that since the GFDM MF receiver subjects to additional ISI and ICI—due to non-zero roll-off factor of the prototype filter—it suffers from a larger error floor, when in it compared to the ZF receiver. Although, the vertical gap of the error floors decreases, as the error floor due to the channel estimation increases for different pilot spacing Δk values (Fig. 3.7 left to right). In addition, notice that by increasing Δk, 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 $\mathbf{F}_L$ wider, resulting in a larger value of $\|\boldsymbol{Q}_{\text{LS}}\|$ and therefore, further noise and interference enhancement.

To Summarize we derived the standard classical and Bayesian estimators for CE of interference-limited MIMO GFDM systems. Here we observed that at low SNR regions, the non-orthogonal MIMO GFDM channel estimation achieves a similar performance as in OFDM CE. Although, due to the interference terms, the non-orthogonal channel estimation is subject to error floor at high SNR values. In OFDM CE, the error floor does not occur, because as long as the channel is stationary and the subcarriers are orthogonal, clear pilot observation allows interference-free channel estimation. In the next section, we analyze the choice of pilot patterns for GFDM channel estimation and considering a SISO system model, we derive low complexity channel estimators based on frequency domain processing.

3.2 Basic Pilot Designs for GFDM Channel Estimation

In this section, we theoretically study the performance of GFDM channel estimation for different pilot patterns via LS, harmonic mean (HM), and LMMSE estimation techniques. While in the previous section, only block-type pilot pattern was being utilized, we further study its impact on interference statistics versus rectangular grid pilots. If the channel frequency-selectivity is not significant, the number of interfering data subsymbols within the GFDM pilot subcarriers would be smaller than that of block-type pilots, because taking M-point DFT—as in (2.5)—yields $M_d = M - M_p$ interfering data subsymbols in frequency domain. Here, M_p denotes the number of pilot subsymbols that are multiplexed with M_d data subsymbols at the pilot subcarriers.

We consider two cases of *one pilot* (OP) and *multiple pilots* (MP) per pilot subcarrier with

the pilot-subcarrier-spacing Δk. Fig. 3.8 shows an example of the time-frequency resource grid. In the OP case (i.e. $M_p^{\mathrm{OP}} = 1$), one pilot subsymbol is inserted into every pilot subcarrier (i.e. yellow in Fig. 3.8), whereas in the MP case, the same value of the first pilot is transmitted at different time slots (e.g. for $M_p^{\mathrm{MP}} = 2$, green in Fig. 3.8). Correspondingly, the effective rate in each scenario is reduced by a factor $(M_d K_p + M K_d)/M(K_p + K_d)$, where $K_d = K - K_p$ is the number of subcarriers that transmit only data. Similar to Sec. 3.1, we consider $\vec{d} = \vec{d}_d + \vec{d}_p$, where $\vec{d}_d \circ \vec{d}_p = \vec{0}_N$ holds, i.e. pilots and data symbols are not superimposed, but rather multiplexed. Additionally, in the MP scenario we have: $\vec{d}_p^{\mathrm{MP}} = \vec{d}_{p,1} + \vec{d}_{p,2} + \cdots + \vec{d}_{p,M_p}$, where $\vec{d}_{p,1}$ and $\vec{d}_{p,M_p}$ contain the first and the M_p-th pilot respectively, with the property $\vec{d}_{p,1} \circ \vec{d}_{p,2} \circ \cdots \circ \vec{d}_{p,M_p} = \vec{0}_N$.

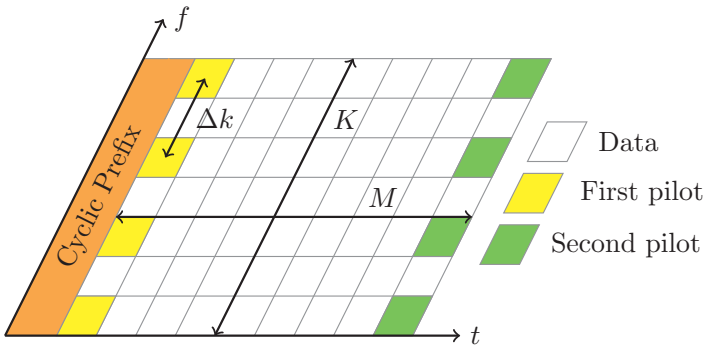


Fig. 3.8: Overview of the pilots structure for $M_p = 2, \Delta k = 2$.

Review of the system model. In the following, for the purpose of deriving low complexity channel estimators in frequency domain, we consider a SISO channel model. The CE approaches that are being proposed in this section will also be used in Sec. 3.3 for MIMO applications, where the pilots can be treated separately. Consider the measurement signal $\vec{Y}_{N_p}$ in frequency domain as

$$\vec{Y}_{N_p} = \boldsymbol{F}_N \boldsymbol{\mathcal{C}}_N \vec{x} + \boldsymbol{F}_N \vec{w}, \tag{3.16}$$

where, $\boldsymbol{\mathcal{C}}_N$ is the circulant channel matrix with impulse response $\vec{h}_N = [\vec{h}^T \ \vec{0}_{N-L}^T]^T$ being its first column. Diagonalizing $\boldsymbol{\mathcal{C}}_N$ yields

$$\vec{Y}_N = \mathrm{diag}(\vec{H}_N)\vec{X} + \vec{W}_N, \tag{3.17}$$

with $\vec{H}_N = \sqrt{N}\boldsymbol{F}_N \vec{h}_N$ and $\vec{X} = \boldsymbol{F}_N \vec{x}$. If pilot-subcarrier-spacing $\Delta k > 1$, only a subset of the frequencies containing $N_p = \lfloor \frac{K}{\Delta k} \rfloor$ pilot samples shall be utilized. Therefore, we define the frequency domain of the receive signal at pilot subcarriers as

$$\vec{Y}_{N_p} = (\boldsymbol{X}_p + \boldsymbol{X}_d)\vec{H}_{N_p} + \vec{W}_{N_p}, \tag{3.18}$$

where $\boldsymbol{X}_p = \mathrm{diag}(\boldsymbol{F}_{N_p,N}\boldsymbol{A}\vec{d}_p)$, $\boldsymbol{X}_d = \mathrm{diag}(\boldsymbol{F}_{N_p,N}\boldsymbol{A}\vec{d}_d)$, $\vec{H}_{N_p} = \sqrt{N}\boldsymbol{F}_{N_p,N}\vec{h}_N$ and $\vec{W}_{N_p} = \boldsymbol{F}_{N_p,N}\vec{w}_N$. Here, $\boldsymbol{F}_{N_p,N} \subseteq \boldsymbol{F}_N$ is the $N_p \times N$ matrix that takes the DFT at pilot subcarriers, i.e. every $m + kM$ row of $\boldsymbol{F}_{N_p,N}$ corresponds to $m + kM\Delta k$ row of $\boldsymbol{F}_N$.

3.2.1 LS/HM Channel Estimation

LS channel estimation for One-Pilot scheme. Assuming only one pilot subsymbol is inserted into each subcarrier, the LS channel estimation which minimizes $\|\vec{Y} - \boldsymbol{X}_p\vec{H}\|^2$ with respect to $\vec{H}$ simply follows:

$$\vec{\tilde{H}}_{\mathrm{OP}} = \boldsymbol{X}_p^{-1}\vec{Y}_{N_p} = \vec{H} + \vec{\tilde{I}}_{\mathrm{OP}} + \vec{\tilde{W}}_{\mathrm{OP}}, \tag{3.19}$$

here, $\vec{\tilde{I}}_{\mathrm{OP}} = \boldsymbol{X}_p^{-1}\boldsymbol{X}_d\vec{H}$ and $\vec{\tilde{W}}_{\mathrm{OP}} = \boldsymbol{X}_p^{-1}\vec{W}$ are the interference and the enhanced noise terms, respectively.

HM channel estimation for Multiple Pilots scheme. The HM channel estimation is equivalent to the mean of multiple OP estimations. Assuming M_p pilots are transmitted on pilot subcarriers (as in Fig. 3.8), the HM channel estimation follows:

$$\begin{aligned}\vec{\tilde{H}}_{\mathrm{HM}} &= \left(M_p[\boldsymbol{X}_{\boldsymbol{p_1}}^{-1} + \boldsymbol{X}_{\boldsymbol{p_2}}^{-1} + \cdots + \boldsymbol{X}_{\boldsymbol{p_{M_p}}}^{-1}]^{-1}\right)^{-1}\vec{Y}_{N_p} \\ &= \frac{1}{M_p}(\boldsymbol{X}_{\boldsymbol{p_1}}^{-1} + \boldsymbol{X}_{\boldsymbol{p_2}}^{-1} + \cdots + \boldsymbol{X}_{\boldsymbol{p_{M_p}}}^{-1})\vec{Y}_{N_p},\end{aligned} \tag{3.20}$$

where $\boldsymbol{X}_{\boldsymbol{p_{M_p}}} = \mathrm{diag}(\boldsymbol{F}_{N_p,N}\boldsymbol{A}\vec{d}_{p,M_p})$. The HM channel estimation is no longer unbiased and the multiple pilots might possibly interfere one another. Hence, the choice of pilots being transmitted at different time slots plays an important role in the quality of channel estimation.

Similar to the OP case, an interference and enhanced noise term arise in HM estimation as well. In the following, for the sake of simplicity, we set $M_p = 2$ for the MP case. Thus, the bias, the interference and enhanced noise terms follow respectively:

$$\vec{B}_{\mathrm{HM}} = \frac{1}{2}\left(\boldsymbol{X}_{\boldsymbol{p_1}}^{-1}\boldsymbol{X}_{\boldsymbol{p_2}} + \boldsymbol{X}_{\boldsymbol{p_2}}^{-1}\boldsymbol{X}_{\boldsymbol{p_1}}\right)\vec{H}_{N_p}, \tag{3.21}$$

$$\vec{\tilde{I}}_{\mathrm{HM}} = \frac{1}{2}\left(\boldsymbol{X}_{\boldsymbol{p_1}}^{-1}\boldsymbol{X}_d + \boldsymbol{X}_{\boldsymbol{p_2}}^{-1}\boldsymbol{X}_d\right)\vec{H}_{N_p}, \tag{3.22}$$

$$\vec{\tilde{W}}_{\mathrm{HM}} = \frac{1}{2}(\boldsymbol{X}_{\boldsymbol{p_1}}^{-1} + \boldsymbol{X}_{\boldsymbol{p_2}}^{-1})\vec{W}_{N_p}, \tag{3.23}$$

DFT based filtering and interpolation. Assuming the CIR is always shorter than the CP length while the noise is spread all over the impulse response, the estimated response must always be truncated at the CP length. Such truncation reduces the impact of noise as well as interference and therefore, it provides an improvement in the estimation. Additionally, the channel length L can also be estimated by considering only the most significant taps of the estimation impulse response [HJ12]. Thus, we assume that we have the *a-priori* knowledge of the channel length L. This knowledge, yields the following modification in the LS and HM estimates:

$$\vec{\hat{h}}_{\mathfrak{M}} = \frac{\Delta k}{\sqrt{N}}\mathbf{F}_L{}^H\vec{\tilde{H}}_{\mathfrak{M}}, \tag{3.24}$$

$$\vec{\hat{H}}_{\mathfrak{M}} = \sqrt{N}\boldsymbol{F}_N\vec{\hat{h}}_{\mathfrak{M}}, \tag{3.25}$$

here, $\mathfrak{M} \in \{\mathrm{OP}, \mathrm{HM}\}$ and $\mathbf{F}'_L \in \mathbb{C}^{N_p \times N}$ is the DFT matrix that performs DFT of only the first L samples of its argument i.e. $\mathbf{F}'_L = \boldsymbol{F}_{N_p,N}\mathrm{diag}([\vec{1}_L^T,\ \vec{0}_{N-L}^T]^T)$. The scaling factor Δk is due to the fact that $\boldsymbol{F}_{N_p,N}$ is not necessarily unitary. The above DFT truncation also changes the interference and enhanced noise terms as

$$\vec{\tilde{I}}_{\mathfrak{M}} = \Delta k \boldsymbol{F}_N {\mathbf{F}'_L}^H \vec{\tilde{I}}_{\mathfrak{M}},\ \vec{\tilde{W}}_{\mathfrak{M}} = \Delta k \boldsymbol{F}_N {\mathbf{F}'_L}^H \vec{\tilde{W}}_{\mathfrak{M}},$$

respectively. Note that in (3.24) and (3.25), $\boldsymbol{F}_N {\mathbf{F}'_L}^H$ also performs DFT based interpolation for calculating the channel frequency response at all data subcarriers.

3.2.2 LMMSE Channel Estimation for GFDM

The LMMSE criterion-based estimator provides a higher estimation quality compared to LS by minimizing the mean squared error. To this end, we compute the coefficients of a linear filter as

$$\vec{\tilde{H}}_{\mathrm{LMMSE}} = \boldsymbol{\Sigma}_{\mathrm{HH}} \boldsymbol{X}_p^H (\boldsymbol{X}_p \boldsymbol{\Sigma}_{\mathrm{HH}} \boldsymbol{X}_p^H + \boldsymbol{\Sigma}_{EE})^{-1} \vec{Y} \tag{3.26}$$

where, the channel auto-correlation matrix $\boldsymbol{\Sigma}_{\mathrm{HH}}$ is given by

$$\boldsymbol{\Sigma}_{\mathrm{HH}} = \mathbb{E}\left[\vec{H}_{N_p} \vec{H}_{N_p}^H\right] = N \boldsymbol{F}_{N_p,N} \mathrm{diag}(\vec{P}_N) \boldsymbol{F}_{N_p,N}^H, \tag{3.27}$$

where $\vec{P}_N = [\vec{P}^T\ \vec{0}_{N-L}^T]^T$ is the PDP vector. Moreover, $\boldsymbol{\Sigma}_{EE}$ is the covariance matrix of interference plus noise terms and assuming they are uncorrelated, we calculate it as

$$\boldsymbol{\Sigma}_{EE} = \mathbb{E}\left[\boldsymbol{X}_d \vec{H}_{N_p} (\boldsymbol{X}_d \vec{H}_{N_p})^H + \vec{W}_{N_p} \vec{W}_{N_p}^H\right] = \boldsymbol{\Sigma}_{\Psi\Psi} + \sigma_w^2 \mathbf{I}_{N_p}, \tag{3.28}$$

where σ_w^2 is the noise variance. The interference covariance matrix $\boldsymbol{\Sigma}_{\Psi\Psi}$ that is a function of two independent random variables $\boldsymbol{X}_d$ and $\vec{H}$, follows [RC91]:

$$\begin{aligned} \boldsymbol{\Sigma}_{\Psi\Psi} &= \mathbb{E}_{X_d}\left[\boldsymbol{X}_d\, \mathbb{E}_H\left[\vec{H}_{N_p} \vec{H}_{N_p}^H | \boldsymbol{X}_d\right] \boldsymbol{X}_d^H\right] \\ &= \boldsymbol{\Sigma}_{X_d X_d} \circ \boldsymbol{\Sigma}_{\mathrm{HH}}, \end{aligned} \tag{3.29}$$

where $\boldsymbol{\Sigma}_{X_d X_d}$ is the data covariance matrix, given by

$$\boldsymbol{\Sigma}_{X_d X_d} = \boldsymbol{F}_{N_p,N} \boldsymbol{A} \mathrm{diag}(\vec{\sigma}_d^2) \boldsymbol{A}^H \boldsymbol{F}_{N_p,N}^H, \tag{3.30}$$

here, $\vec{\sigma}_d^2 \in \mathbb{R}^{N \times 1}$ is the vector of data variances σ_d^2 including zeros at pilot positions.

Channel auto-correlation estimation via LS. Since the LMMSE estimation requires the knowledge of the channel autocorrelation, we estimate the corresponding channel covariance matrix (CCM) via LS or HM approach. In order to take advantage of the circularity of the channel, we approximate the average power of each path and

accordingly we calculate the channel matrix. To this end, we approximate the average power through the absolute square of the impulse response $\vec{\hat{h}}_{\mathfrak{M}}$ as

$$\hat{\vec{P}}_N \approx \left|\vec{\hat{h}}_{\mathfrak{M}}\right|^2, \tag{3.31}$$

then, we estimate the channel autocorrelation matrix as

$$\widehat{\boldsymbol{\Sigma}}_{HH} \approx N\boldsymbol{F}_{N_p,N}\mathrm{diag}(\hat{\vec{P}}_N)\boldsymbol{F}_{N_p,N}^H. \tag{3.32}$$

Note that (3.32) is not an unbiased estimation and $\mathbb{E}\left[\widehat{\boldsymbol{\Sigma}}_{\mathrm{HH}}\right] \neq \boldsymbol{\Sigma}_{\mathrm{HH}}$. Nevertheless, averaging the power over multiple channel estimates i.e. $\hat{\vec{P}}_{\mathscr{N}} = \frac{1}{\mathscr{N}}\left|\sum_{\mathscr{N}} \vec{\hat{h}}_{\mathfrak{M}}\right|^2$ is unbiased and provides improvement in estimation of $\widehat{\boldsymbol{\Sigma}}_{\mathrm{HH}}$.

3.2.3 Error Characterization

MSE for the OP case. The Mean Squared Error of the channel estimation is expressed as

$$\begin{aligned}\mathrm{MSE}_{\mathrm{OP}} &= \mathbb{E}\left[\left\|\hat{\vec{H}}_{\mathrm{OP}} - \vec{H}\right\|^2\right] \\ &= \mathbb{E}\left[\left\|\hat{\vec{I}}_{\mathrm{OP}} + \hat{\vec{W}}_{\mathrm{OP}}\right\|^2\right].\end{aligned}$$

Elaborating in the above expression and through the rule of Quadratic forms in random variables [MP92, Ch. 3], we have:

$$\mathrm{MSE}_{\mathrm{OP}} = \mathrm{Tr}\left(\boldsymbol{\Sigma}_{\boldsymbol{X}_p\boldsymbol{X}_p}\boldsymbol{\Sigma}_{\Psi\Psi}\right) + \sigma_w^2\mathrm{Tr}(\boldsymbol{\Sigma}_{\boldsymbol{X}_p\boldsymbol{X}_p}), \tag{3.33}$$

where $\boldsymbol{\Sigma}_{\boldsymbol{X}_p\boldsymbol{X}_p} = \Delta k^2 \boldsymbol{X}_p^{-H}\boldsymbol{F}_L'\boldsymbol{F}_L'^H\boldsymbol{X}_p^{-1}$.

MSE for the HM case. Since the Harmonic Mean channel estimation is equivalent to the mean of multiple OP estimations, the error arises from the sum of the MSE due to the separate estimations plus their correlation i.e.

$$\begin{aligned}\mathrm{MSE}_{\mathrm{HM}} &= \mathbb{E}\left[\left\|\frac{1}{2}\hat{\vec{H}}_{\mathrm{HM}_1} + \frac{1}{2}\hat{\vec{H}}_{\mathrm{HM}_2} - \vec{H}\right\|^2\right] \\ &= \frac{1}{4}\left(\mathrm{MSE}_1 + \mathrm{MSE}_2 + 2\Re\left(\mathbb{E}[\langle\vec{E}_1, \vec{E}_2\rangle]\right)\right),\end{aligned} \tag{3.34}$$

with

$$\begin{aligned}\vec{E}_1 &= \hat{\vec{H}}_{\mathrm{HM}_1} - \vec{H}, \qquad \vec{E}_2 = \hat{\vec{H}}_{\mathrm{HM}_2} - \vec{H}, \\ \hat{\vec{H}}_{\mathrm{HM}_1} &= \boldsymbol{X}_{\boldsymbol{p_1}}^{-1}\vec{Y}, \qquad \hat{\vec{H}}_{\mathrm{HM}_2} = \boldsymbol{X}_{\boldsymbol{p_2}}^{-1}\vec{Y}.\end{aligned}$$

As in the OP case, the MSE due to the first and second estimations follow:

$$\mathrm{MSE}_{m_p} = Z_{m_p} + \mathrm{Tr}\left(\boldsymbol{\Sigma}_{\boldsymbol{X}_p,m_p}\boldsymbol{\Sigma}_{\Psi\Psi}\right) + \sigma_w^2\mathrm{Tr}(\boldsymbol{\Sigma}_{\boldsymbol{X}_p,m_p}), \tag{3.35}$$

where $m_p \in \{1,2\}$ and

$$Z_1 = \mathrm{Tr}\left(\boldsymbol{X}_{\boldsymbol{p_2}}^H\boldsymbol{\Sigma}_{\boldsymbol{X}_p,1}\boldsymbol{X}_{\boldsymbol{p_2}}\boldsymbol{\Sigma}_{\mathrm{HH}}\right), \tag{3.36}$$

$$Z_2 = \mathrm{Tr}\left(\boldsymbol{X}_{\boldsymbol{p_1}}^H\boldsymbol{\Sigma}_{\boldsymbol{X}_p,2}\boldsymbol{X}_{\boldsymbol{p_1}}\boldsymbol{\Sigma}_{\mathrm{HH}}\right), \tag{3.37}$$

$$\boldsymbol{\Sigma}_{\boldsymbol{X}_p,m_p} = \Delta k^2\boldsymbol{X}_{\boldsymbol{p_{m_p}}}^{-H}\boldsymbol{F}_L'\boldsymbol{F}_L'^H\boldsymbol{X}_{\boldsymbol{p_{m_p}}}^{-1}, \tag{3.38}$$

Note that in the above expressions Z_1 and Z_2 are the MSE due to the bias of the estimations. Then, we address the inner product between the two estimations as

$$\begin{aligned}\mathbb{E}[\langle\vec{E}_1,\vec{E}_2\rangle] &= \mathbb{E}\left[\vec{E}_1^H\vec{E}_2\right]\\ &= \mathbb{E}\left[\vec{\Psi}_1^H\boldsymbol{\Sigma}_{\boldsymbol{X}_p,\langle1,2\rangle}\vec{\Psi}_2\right] + \mathbb{E}\left[\vec{W}^H\boldsymbol{\Sigma}_{\boldsymbol{X}_p,\langle1,2\rangle}\vec{W}\right],\end{aligned}$$

here,

$$\boldsymbol{\Sigma}_{\boldsymbol{X}_p,\langle1,2\rangle} = \Delta k^2\boldsymbol{X}_{\boldsymbol{p_1}}^{-H}\mathbf{F}_L'\mathbf{F}_L'^H\boldsymbol{X}_{\boldsymbol{p_2}}^{-1}, \tag{3.39}$$

$$\vec{\Psi}_1 = (\boldsymbol{X}_{\boldsymbol{p_2}} + \boldsymbol{X}_d)\vec{H}, \tag{3.40}$$

$$\vec{\Psi}_2 = (\boldsymbol{X}_{\boldsymbol{p_1}} + \boldsymbol{X}_d)\vec{H}. \tag{3.41}$$

Via the identities (A.1), (A.2) and (A.3) from Appendix A.1 and after some algebraic manipulations, the inner product eventually becomes:

$$\mathbb{E}\left[\langle\vec{E}_1,\vec{E}_2\rangle\right] = Z_{\langle1,2\rangle} + \mathrm{Tr}\left(\boldsymbol{\Sigma}_{\boldsymbol{X}_p,\langle1,2\rangle}\boldsymbol{\Sigma}_{\Psi_1\Psi_2}\right) + \sigma_w^2\mathrm{Tr}\left(\boldsymbol{\Sigma}_{\boldsymbol{X}_p,\langle1,2\rangle}\right), \tag{3.42}$$

where

$$Z_{\langle1,2\rangle} = \mathrm{Tr}\left(\boldsymbol{X}_{\boldsymbol{p_2}}^H\boldsymbol{\Sigma}_{\boldsymbol{X}_p,\langle1,2\rangle}\boldsymbol{X}_{\boldsymbol{p_1}}\boldsymbol{\Sigma}_{\mathrm{HH}}\right), \tag{3.43}$$

$$\boldsymbol{\Sigma}_{\Psi_1\Psi_2} = \boldsymbol{F}_{N_p,N}\boldsymbol{A}\mathrm{diag}(\vec{\sigma}_d^2)\boldsymbol{A}^H\boldsymbol{F}_{N_p,N}^H \circ \boldsymbol{\Sigma}_{\mathrm{HH}}, \tag{3.44}$$

Cramér-Rao lower bound. In the following, we study the CRLB for the one-pilot subsymbol per subcarrier case, which its unbiased estimator was derived in (3.19). The observation vector $\vec{Y}_{N_p} \sim \mathcal{N}_C(\boldsymbol{X}_p\vec{H}_{N_p}, \boldsymbol{\Sigma}_{EE})$ is described by the following complex Gaussian probability density function (pdf):

$$p(\vec{Y}_{N_p}|\vec{H}_{N_p}) = \frac{1}{\pi^N\det(\boldsymbol{\Sigma}_{EE})}\exp\Big[-(\vec{Y}_{N_p} - \boldsymbol{X}_p\vec{H}_{N_p})^H\boldsymbol{\Sigma}_{EE}^{-1}(\vec{Y}_{N_p} - \boldsymbol{X}_p\vec{H}_{N_p})\Big], \tag{3.45}$$

which satisfies the *regularity* condition [Kay93, Sec. 15.7] i.e.

$$\mathbb{E}\left[\frac{\partial\ln p(\vec{Y}_{N_p}|\vec{H}_{N_p})}{\partial\vec{H}_{N_p}^*}\right] = \vec{0}_N, \text{ for } \forall\vec{H}_{N_p}. \tag{3.46}$$

Hence, an unbiased estimator may be found that attains the CRLB in that $\boldsymbol{\Sigma}_{\text{CRLB}} = \boldsymbol{\mathcal{I}}(\boldsymbol{H})^{-1}$, if and only if:

$$\frac{\partial \ln p(\vec{Y}_{N_p}|\vec{H}_{N_p})}{\partial \vec{H}^*_{N_p}} = \boldsymbol{\mathcal{I}}(\boldsymbol{H})(\vec{\check{H}}_{\text{OP}} - \vec{H}_{N_p}), \tag{3.47}$$

for the complex *Fisher Information* matrix $\boldsymbol{\mathcal{I}}(\boldsymbol{H})$. Therefore, from (3.45), we have:

$$\frac{\partial \ln p(\vec{Y}_{N_p}|\vec{H}_{N_p})}{\partial \vec{H}^*_{N_p}} = \boldsymbol{X}_p^H \boldsymbol{\Sigma}_{EE}^{-1} \boldsymbol{X}_p \left(\boldsymbol{X}_p^{-1} \vec{Y}_{N_p} - \vec{H}_{N_p} \right).$$

Thus, the *equality* condition (3.47) is satisfied, and (3.19) is an efficient estimator (i.e. the Minimum Variance Unbiased estimator) with covariance matrix:

$$\boldsymbol{\Sigma}_{\text{CRLB}} = \boldsymbol{\mathcal{I}}(\boldsymbol{H})^{-1} = (\boldsymbol{X}_p^H \boldsymbol{\Sigma}_{EE}^{-1} \boldsymbol{X}_p)^{-1}. \tag{3.48}$$

Note that (3.48) is the covariance matrix of CRLB for an estimation without considering the a-priori knowledge of the channel length L. In order to also take into account the improvement via a-priori knowledge of L, we rearrange the observation vector as $\vec{\mathcal{Y}} = \Delta k \boldsymbol{F}_N \mathbf{F}'^{\,H}_L \boldsymbol{X}_p^{-1} \vec{Y}_{N_p}$. Therefore, the pdf for $\vec{\mathcal{Y}} \sim \mathcal{N}_C(\vec{H}_{N_p}, \widehat{\boldsymbol{\Sigma}}_{EE})$ becomes:

$$p(\vec{\mathcal{Y}}|\vec{H}_{N_p}) = \frac{1}{\pi^N \det(\widehat{\boldsymbol{\Sigma}}_{EE})} \exp\Big[- (\vec{\mathcal{Y}} - \vec{H}_{N_p})^H \widehat{\boldsymbol{\Sigma}}_{EE}^{-1} (\vec{\mathcal{Y}} - \vec{H}_{N_p}) \Big], \tag{3.49}$$

where

$$\widehat{\boldsymbol{\Sigma}}_{EE} = \boldsymbol{\Sigma}_{\boldsymbol{X}_p \boldsymbol{X}_p} \boldsymbol{\Sigma}_{\Psi\Psi} + \sigma_w^2 \boldsymbol{\Sigma}_{\boldsymbol{X}_p \boldsymbol{X}_p}. \tag{3.50}$$

Subsequently, the result of derivation follows:

$$\frac{\partial \ln p(\vec{\mathcal{Y}}|\vec{H}_{N_p})}{\partial \vec{H}^*_{N_p}} = \widehat{\boldsymbol{\Sigma}}_{EE}^{-1} (\Delta k \boldsymbol{F}_N \mathbf{F}'^{\,H}_L \boldsymbol{X}_p^{-1} \vec{Y}_{N_p} - \vec{H}_{N_p}). \tag{3.51}$$

It can be seen from (3.51) that the refined efficient estimation is $\widehat{H}_{\text{eff}} = \Delta k \boldsymbol{F}_N \mathbf{F}'^{\,H}_L \boldsymbol{X}_p^{-1} \vec{Y}_{N_p}$. Furthermore, the CRLB for the refined estimation follows:

$$\widehat{\boldsymbol{\Sigma}}_{\text{CRLB}} = \boldsymbol{\mathcal{I}}_{\text{eff}}(\boldsymbol{H})^{-1} = \widehat{\boldsymbol{\Sigma}}_{EE}. \tag{3.52}$$

An immediate observation from (3.52) depicts that the CRLB covariance matrix coincides with the LS MSE covariance matrix defined in (3.33).

MSE for LMMSE estimation. We compute the mean squared error of the LMMSE estimator through the error covariance matrix which follows a standard form as explained in [Kay93]. Thus, we have:

$$\begin{aligned} \text{MSE}_{\text{LMMSE}} &= \mathbb{E}\left[\left\| \vec{\widehat{H}}_{\text{LMMSE}} - \vec{H}_{N_p} \right\|^2 \right] \\ &= \text{Tr}\left(\boldsymbol{\Sigma}_{\text{HH}} - \widehat{\boldsymbol{\Sigma}}_{HH} \right). \end{aligned} \tag{3.53}$$

where the covariance matrix of the estimated channel follows:

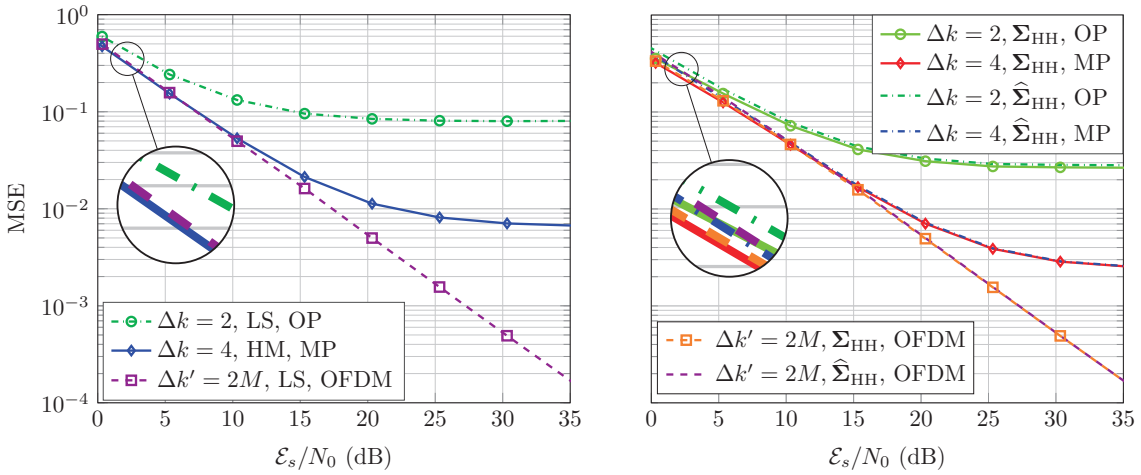
$$\begin{aligned} \widehat{\boldsymbol{\Sigma}}_{HH} &= \mathbb{E}\left[\vec{\widehat{H}}_{\text{LMMSE}} \, \vec{\widehat{H}}^{\,H}_{\text{LMMSE}} \right] \\ &= \boldsymbol{\Sigma}_{\text{HH}} \boldsymbol{X}_p^H (\boldsymbol{X}_p \boldsymbol{\Sigma}_{\text{HH}} \boldsymbol{X}_p^H + \boldsymbol{\Sigma}_{EE})^{-1} \boldsymbol{X}_p \boldsymbol{\Sigma}_{\text{HH}}. \end{aligned} \tag{3.54}$$

3.2.4 Simulation Results

In this section, we evaluate the performance of GFDM receiver under the impact of channel estimation errors. Specifically, we compare the result of two pilot patterns of (semi-) block-type and rectangular grid. The (semi-) block-type pilot pattern inserts one pilot subsymbol into the pilot subcarriers (i.e. OP case) with $\Delta k_{\mathrm{OP}} = 2$. On the other hand, the rectangular grid pilot pattern employs two pilots within the pilot subcarriers while $\Delta k_{\mathrm{MP}} = 4$ and uses MP algorithms for channel estimation. Monte-Carlo method has been adapted for evaluating the bit error rate (BER) for low and high code-rates.

Here, we consider a sequence of information bits with energy per bit $\mathcal{E}_b$ are coded through parallel concatenated convolutional codes (PCCC) with octal generator polynomials (1,15/13) and code-rates of 1/3 and 5/6. The coded bits are mapped into 16-QAM symbols with energy per symbol $\mathcal{E}_s$ and transmitted through $M = 7$ subsymbols on $K = 96$ subcarriers. In addition, the pilots are generated in form of first root *LTE Zadoff-chu* sequence and multiplexed by the QAM symbols. The prototype filter of the GFDM transmitter is configured as a Raised Cosine pulse with roll-off factor $\alpha = 0.3$. After GFDM modulation, the transmit signal $\vec{x}$ is influenced by a random multipath channel including $L = 24$ independent fading gains with exponential power delay profile and Rayleigh distribution. The channel is assumed to be static within one GFDM block. Moreover, the noise spectral density is denoted N_0. At the receiver side, we employ zero forcing channel equalizer as well as zero forcing GFDM demodulator (i.e. $\boldsymbol{A}^{-1}$). The data symbols after GFDM demodulation are soft decoded with 8 turbo decoder iterations. In order to compare the GFDM performance with a conventional OFDM, we consider sufficiently long symbols (with respect to the CP length) are transmitted through K' narrow subcarriers in OFDM. Assuming the duration of one transmission is identical for GFDM and OFDM, GFDM splits the OFDM symbol duration into M subsymbols. Therefore, the subcarrier spacing in GFDM becomes M times broader compared to OFDM. In order to maintain identical bandwidth for GFDM and OFDM, GFDM must have M times fewer subcarriers than OFDM resulting $K' = MK$. For the sake of keeping identical pilots overhead for GFDM and OFDM we set the OFDM pilot subcarrier spacing $\Delta k' = M\Delta k_{\mathrm{OP}}$.

Fig. 3.9 compares the performance of channel estimation in OFDM with two cases of (semi-) block-type and rectangular grid pilots pattern in GFDM. Fig. 3.9a depicts that the rectangular grid has significantly better performance than the block type pilot pattern. In Fig. 3.9a *zero* MSE has been observed for the bias of HM estimation, which is due to the fact that identical pilot subsymbols have been used within the pilot subcarriers. Moreover, note that the interference from data symbols in GFDM depends to the amount of data transmitted on pilot subcarriers. In block type, $M - 1$ data subsymbols in every pilot subcarrier are the source of interference within the channel estimation. However, in rectangular grid, $M - M_p$ data subsymbols interfere the M_p pilots in pilot subcarriers, i.e. pilot power to interference ratio is proportional to $\frac{M_p}{M-M_p}$. In addition, since the noise enhancement in HM channel estimation is smaller than the case of LS, the estimation



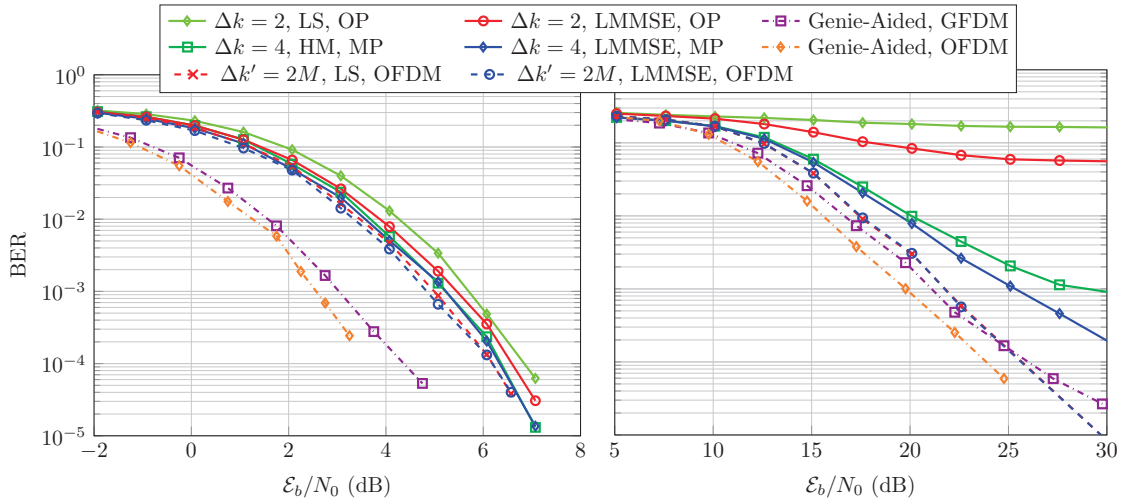
(a) MSE for LS and HM estimations.

(b) LMMSE estimation with perfect $\mathbf{\Sigma}_{\mathrm{HH}}$ and imperfect $\widehat{\mathbf{\Sigma}}_{\mathrm{HH}}$ knowledge of channel covariance matrix.

Fig. 3.9: Channel estimation MSE for OP vs. MP cases in GFDM. Pilots overhead is 7%. Theoretic and simulated values are shown with *marks* and *lines* respectively.

error due to noise is smaller for rectangular grid. As a result, both interference and noise are better rejected in HM channel estimation. A similar behavior can also be observed for LMMSE channel estimation in Fig. 3.9b. Note that both noise and interference are at a lower level in MP (i.e. rectangular grid) compared to the OP (i.e. block type) case. Fig. 3.9b also shows that although LMMSE channel estimation with imperfect knowledge of CCM has less than 0.8 dB difference at low SNR, it becomes almost identical with perfect CCM knowledge at high SNR, specially for the case of MP. The theoretical MSE calculations do not hold when using the imperfect knowledge of CCM and thus they are omitted in Fig. 3.9b. Comparing both the LMMSE and LS/HM GFDM channel estimations with OFDM, we observe that GFDM MP is marginally better than OFDM at low SNR due to the smaller noise enhancement in MP case. On the contrary, looking at the high SNR, GFDM still suffers from the error floor due to interference from data symbols while OFDM channel estimation improves as SNR increases because of its clear pilots observation. Looking into the complexity of the estimation techniques, LMMSE has a complexity which increases with $\mathcal{O}(N^3)$ due to the matrix inversion of size $\lfloor \frac{N}{\Delta k} \rfloor \times \lfloor \frac{N}{\Delta k} \rfloor$. However, the complexity of the LS/HM estimation is dominated by N size FFT transform i.e. $\mathcal{O}(N \log N)$.

The result of Monte-Carlo simulation for coded performance of GFDM and OFDM under the influence of channel estimation is illustrated in Fig. 3.10, where we have simulated the coded bit error rate for different code-rates. The x-axis of the BER curves in Fig. 3.10 can be linked to the MSE curves (i.e. Fig. 3.9) by $\mathcal{E}_b/N_0 = \mathcal{E}_s/N_0 - 10\log_{10}(\mu r_{\mathrm{code}})$ where μ is the modulation order and r_{code} is the code-rate. Both cases are also compared to Genie-Aided (GA) receivers where we assumed the perfect knowledge of the CSI at the



(a) 1/3 code-rate **(b)** 5/6 code-rate

Fig. 3.10: Bit error rate performance for OP vs. MP cases in GFDM. Pilots overhead is 7%.

receiver side. Due to the similarity of the channel estimation performances with perfect and imperfect knowledge of channel covariance matrix, here, we present the results only for the LMMSE channel estimation with imperfect knowledge of CCM. As expected from the channel estimation errors, the BER performance of all GFDM cases are very similar to OFDM at 1/3 code-rate which is suitable for low SNR scenarios. On the other hand, if we employ a high code-rate of 5/6, we observe that the large error floor in channel estimation for the OP case yields an unacceptable high code-rate performance. However, GFDM MP cases achieve a significantly better performance compared to the OP case because of better interference suppression, though, due to some remaining error floor in rectangular grid pilot pattern, the MP cases become roughly 5 dB worse than OFDM at high SNR.

To sum up, we theoretically studied the choice of pilots pattern for GFDM CE and we found out that rectangular grid pilot pattern better suppresses the interference in GFDM channel estimation than the block type pilots. We also evaluated the coded performance of GFDM and OFDM transceivers under imperfect channel estimation and CCM knowledge. As a remark, the performance of both transceivers are similar when a robust code-rates is adopted. But for a high code-rate, GFDM transceiver falls behind OFDM because of the self-interference in GFDM signal itself as well as contaminated pilots in its channel estimation. In the next section, by having a closer look on GFDM modulation via frequency domain processing, we propose a technique to insert interference-free pilots.

3.3 Interference-Free Pilot Insertion for MIMO GFDM Channel Estimation

In this section, in contrast to the approaches used in Sec. 3.2 and Sec. 3.1, we aim to insert orthogonal pilots into the pilot subcarriers. In GFDM modulation, it is possible to reserve some frequency bins specifically for the pilots without any influence from the data symbols.

3.3.1 Interference-Free Pilot Insertion

In order to insert interference-free pilots, we modify the expression (2.5) by the following:

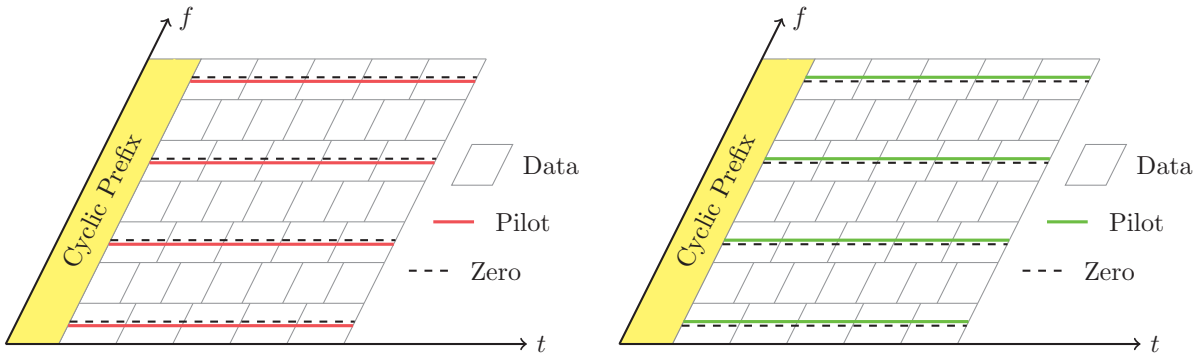
$$\vec{x} = \boldsymbol{F}_N^H \sum_{k=0}^{K-1} \boldsymbol{P}^{(k)} \boldsymbol{G}^{(\delta)} \boldsymbol{R}^{(\delta)} \left(\boldsymbol{\Gamma} \vec{d}_{k,p} + \boldsymbol{F}_M \vec{d}_{k,d} \right), \tag{3.55}$$

where $\boldsymbol{\Gamma} = \boldsymbol{P'}\text{blkdiag}(\lambda \boldsymbol{I}_{\mathfrak{I}}, \boldsymbol{F}_{M-\mathfrak{I}})$. Here, $\boldsymbol{P'}$ can be any permutation matrix of compatible size which allocates the pilots to any frequency bin within the pilot subcarriers. The parameter λ is a scaling factor that normalizes the pilots energy to one. Note that in expression (3.55) the plus sign does not superimpose the information i.e. if k is a pilot subcarrier $\boldsymbol{F}_M \vec{d}_{k,d}$ becomes $\vec{0}_M$ and if it belongs to data subcarriers $\boldsymbol{\Gamma} \vec{d}_{k,p}$ would be $\vec{0}_M$. Additionally, the signal $\vec{\tilde{x}}$ at pilot subcarriers[2] can be obtained from the first term of (3.55), i.e.

$$\vec{\tilde{x}} = \boldsymbol{F}_N^H \sum_{k=0}^{K-1} \boldsymbol{P}^{(k)} \boldsymbol{G}^{(\delta)} \boldsymbol{R}^{(\delta)} \boldsymbol{\Gamma} \vec{d}_{k,p}. \tag{3.56}$$

In (3.56), if as an example, we set $\boldsymbol{P'} = \boldsymbol{I}_M$, the first $\mathfrak{I}$ subsymbols of the pilot subcarriers (i.e. $\{\tilde{d}_k[0], \tilde{d}_k[1], \cdots, \tilde{d}_k[\mathfrak{I} - 1]\}$ which are filled with pilots) are processed directly in frequency domain being isolated from the rest of subsymbols. Although, such isolation holds if and only if the pilots are located at the frequency bins where no inter-carrier interference is present. In this case, each orthogonal subsymbol can be reserved for a specific Tx antenna such that the $\mathfrak{I} \times \mathfrak{Q}$ MIMO channel can be processed in terms of $\mathfrak{I}\mathfrak{Q}$ SISO channels. The approach can be considered as a variation of cell-specific reference signal mapping in LTE [Acc09]. Fig. 3.11 shows an example how the pilot subsymbols in the GFDM data block are mapped into the time-frequency grid of the resources for a 2×2 MIMO channel. Here, two frequency bins of the pilot subcarriers are reserved only for the pilots while at each Tx antenna only one pilot is being transmitted. Thus, the pilot is being transmitted during the whole GFDM symbol, while also the energies of the data subsymbols are no longer concentrated at equally spaced M peaks. Fig. 3.12 shows an example of the signal filtering in frequency domain where the pilots at different antennas are orthogonal to one another as well as to the data samples.

[2] Note that at pilot subcarriers, we have $\mathfrak{I}$ pilot subsymbols and $M - \mathfrak{I}$ data subsymbols in the $\boldsymbol{D}$ matrix.



(a) Antenna 1 **(b)** Antenna 2

Fig. 3.11: Data subsymbols and interference-free pilots in time-frequency resources.

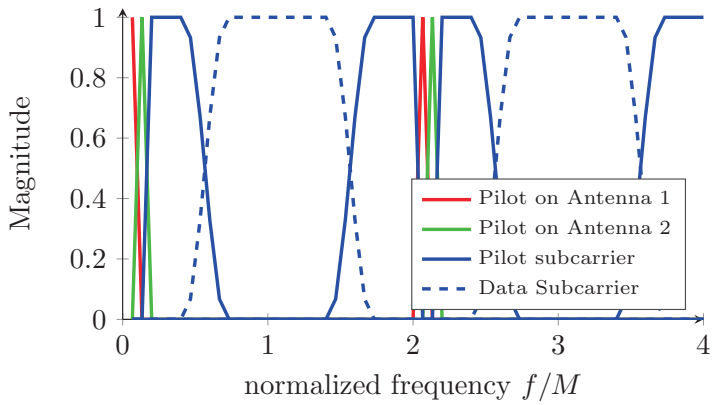


Fig. 3.12: Signal DFT for $M = 15$, $K = 4$, $\boldsymbol{P}' = \boldsymbol{I}_M$. Since two frequency bins at the pilot subcarriers are reserved only for pilot bins, interference-free pilot observation can be achieved in DFT domain.

3.3.2 Pilot Observation

Since the pilots from each Tx antenna are orthogonal to the rest of subsymbols, the channel between each Tx-Rx antenna pair can be separately estimated using the existing estimation techniques. For such separate estimations, we need to extract the pilot frequencies from the received signal at each Rx antenna. Define the matrix $\boldsymbol{F}^{(i)} \subset \boldsymbol{F}_N$ which contains only the rows of the DFT matrix associated to the frequency bins of the pilots from antenna i. The matrix $\boldsymbol{F}^{(i)}$ yields to non-zero values in frequency domain of the pilots signal $\vec{\tilde{X}}'_i = \boldsymbol{F}^{(i)}(\vec{\tilde{x}}_0)_i$ from antenna i. Here, $\vec{\tilde{x}}_0$ is a known time domain pilots signal and it is generated in a similar fashion as $\vec{\tilde{x}}$ (i.e. (3.56)) with the only difference that zeros are inserted at data positions of $\vec{d}_{k,p}$. Due to $\boldsymbol{F}^{(i)}$, also the frequency domain signal $\vec{\tilde{X}}_i = \boldsymbol{F}^{(i)}\vec{x}_i$ consists of only pilot values (of antenna i) with zero influence from data subsymbols.

Once the DFT matrix $\boldsymbol{F}^{(i)}$ associated to the frequencies of interest is specified, the signal

due to the received pilots from Tx antenna i at Rx antenna q is given by

$$\vec{Y}_{i,q} = \boldsymbol{F}^{(i)}\vec{y}_q. \tag{3.57}$$

Here, since each received pilot sequence $\vec{Y}_{i,q}$ between antenna pairs i-q can be treated separately in form of multiple SISO channels, we refer the readers to Sec. 2.4 for the LS and LMMSE estimations of the channel.

3.3.3 Complexity

Due to the orthogonality of the pilots and data in frequency domain, the complexity of the channel estimation for interference-free pilot insertion (IFPI) GFDM turns out to be identical to that of OFDM and hence, it becomes less complex than the channel estimation in basic GFDM. In basic GFDM, since the interference from data subsymbols is treated as extra correlated noise, further signal processing operations are required to reduce the impact of such interference. Moreover, since the pilots are spread over all frequency bins of the pilot subcarriers, more frequency domain pilot samples (i.e. $N_p = K_pM$ samples) need to be handled in basic GFDM. Considering the matrix inversion in LMMSE estimator, the complexity growth of channel estimation in OFDM as well as IFPI GFDM is proportional to $\mathcal{O}(\mathfrak{I}\mathfrak{Q}K_p^3)$, whereas the basic GFDM requires $\mathcal{O}(\mathfrak{Q}N_p^3)$ for its LMMSE channel estimation. On the other hand, since the IFPI GFDM employs the frequency domain GFDM modulation for orthogonal pilot insertion, its complexity increases with respect to the basic GFDM modulated in time domain. The time domain GFDM modulation [MMG+16] requires an implementation effort of $M(N + K\log_2 K)$ in terms of number of complex multiplications. The frequency domain GFDM modulation needs $N\log_2 N + K\delta M + KM\log_2 M$ complex multiplications [GMN+13]. Though, for IFPI GFDM it slightly reduces to $N\log_2 N + K\delta M + K_p(M-\mathfrak{I})\log_2(M-\mathfrak{I}) + K_dM\log_2 M$. In general, considering a receiver that employs a channel estimator, the complexity of the data transmission in IFPI GFDM becomes smaller than that of basic GFDM with the channel estimation proposed in Sec. 3.1.

3.3.4 Simulation Results

In order to evaluate the performance of GFDM with interference-free pilot design, we adopt a 2×2 MIMO block fading multipath channel with Rayleigh distribution. Here, by considering the performance of OFDM channel estimation as a benchmark, we compare the performance of GFDM under the impact of channel estimation with the proposed orthogonal pilot insertion versus our prior approach (i.e. Sec. 3.1) which adapted the basic GFDM pilot insertion. The PAPR, OOB radiation, MSE and the coded BER performances are evaluated through Monte-Carlo simulations with the parameters summarized in Table 3.1. The signal to noise ratio is denoted by $\mathcal{E}_s/N_0$ and considering also the coding

Table 3.1: Simulation Parameters for a channel model similar to LTE extended vehicular A (EVA) power-delay-profile. As in SER results of Sec 3.1, we observed that GFDM zero forcing receiver was subject to less interference than the MF receiver, we solely evaluate the performance of ZF receiver here.

Parameter	*Symbol*	*GFDM*	*OFDM*
Modulation order	μ	4 (16-QAM)	
No. of subsymbols	M	7,15,21	1
No. of subcarriers	K	4,96	84, 96, 672
Pilot subcarrier spacing	Δk	3	21
Roll-off factor of RC filter	α	0.3	0
GFDM demodulation	-	Zero Forcing	-
Pilot sequence - Basic	-	Complex randoms	-
Pilot sequence - IFPI, OFDM	-	1^{st} root Zadoff-Chu	
Channel equalization	-	MMSE	
Channel coding	-	PCCC	
Octal generator polynomial	-	(1,15/13)	
Decoder decision metric	-	log-MAP	
No. of turbo iterations	-	8	
No. of channel taps	L	24	
Power delay profile	-	Exponential	

gain we have $\mathcal{E}_b/N_0 = \mathcal{E}_s/N_0 - 10\log_{10}(\mu r_{\text{code}})$. The receive signal constellations are detected via GFDM zero forcing demodulation and MMSE frequency domain equalization i.e. $\hat{\vec{d}} = (\boldsymbol{I}_\Omega \otimes \boldsymbol{A}^{-1})\boldsymbol{\Xi}[\vec{Y}_1^T \cdots \vec{Y}_\Omega^T]^T$, where, $\boldsymbol{\Xi} = (\hat{\boldsymbol{H}}^H\hat{\boldsymbol{H}} + \sigma_w^2\boldsymbol{I}_N)^{-1}\hat{\boldsymbol{H}}^H$ is the equalization matrix using the estimated channels matrix $\hat{\boldsymbol{H}}$. The detected data symbols are transferred into maximum likelihood (ML) symbol log-likelihoods and they are then fed into the soft demapper with 8 turbo decoder iterations.

The PAPR of the proposed signal is compared to the original GFDM beside OFDM in Fig. 3.13. One can see that due to orthogonal pilot insertion, the PAPR of IFPI GFDM increases with respect to the basic GFDM. However, it still has more than one dB difference with the PAPR of an OFDM signal. On the other hand, comparing the power spectral densities of the signals in Fig. 3.14, we observe that IFPI GFDM has slightly larger OOB compared to the original GFDM signal. In fact, the orthogonal pilot insertion causes the time domain signal to have slightly more abrupt changes, resulting on further OOB radiation. Nonetheless, the windowed case of IFPI GFDM, i.e. W-IFPI achieves almost the same OOB radiation as in original windowed GFDM (W-GFDM). Here, the window function is configured in form of Raised Cosine (RC) window with a ramp length of a quarter subsymbol. For further details regarding the windowing process, we refer the interested readers to [MMG+14].

Figure 3.15a compares the MSE performances of channel estimation with the proposed IFPI technique, OFDM and the basic GFDM channel estimation (i.e. Sec. 3.1). Due to the

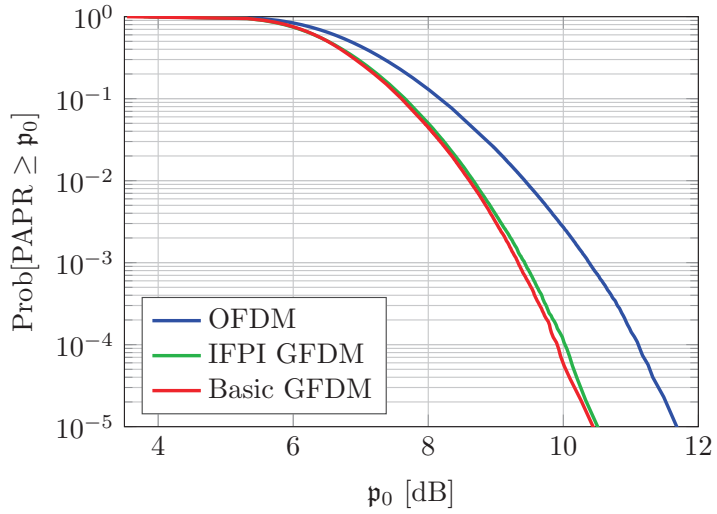


Fig. 3.13: Signal PAPR for $M = 21$, $K = 4$.

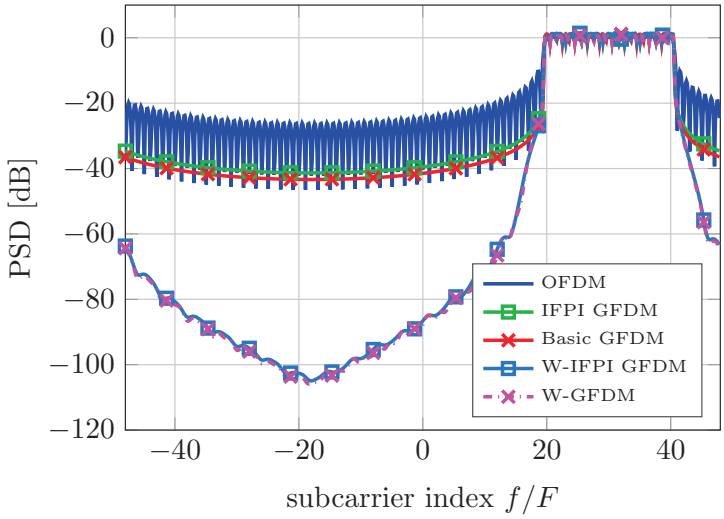
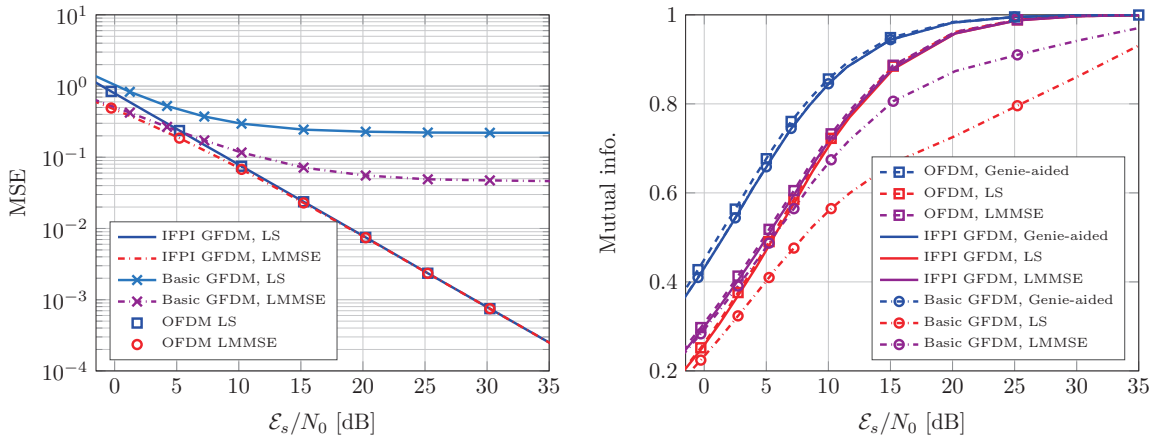


Fig. 3.14: Out of band emission for $K = 96$, $M = 15$ and the total number of frequency samples F.

orthogonality of the pilots to the data symbols in IFPI, one can see that the performance of such channel estimation becomes identical with the performance of OFDM. However, the basic GFDM channel estimation which suffers from significant interference has a large error floor at high SNR regions. Similar behavior can also be observed in Fig. 3.15b. The genie-aided receivers have perfect knowledge of the channel state information without transmitting pilots. Here, the OFDM genie-aided receiver has slightly larger mutual information with respect to the both GFDM cases due to the orthogonality between the data symbols. With the current GFDM configuration the self-interference is treated as extra noise at the receiver slightly limiting its outcome, although, such performance loss is marginal when we consider the GFDM gains in PAPR and OOB emission. Also note that the genie-aided receiver of IFPI GFDM has slightly smaller mutual information with respect to the basic GFDM genie-aided receiver which is due to the fact that in basic



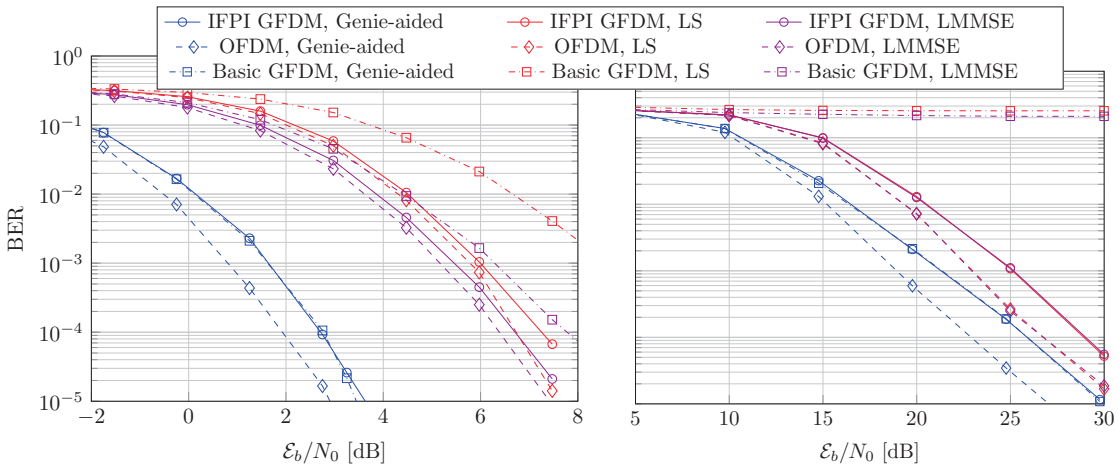
(a) Channel estimation performance for $M = 7, K = 96$.

(b) Mutual information of transmitted bitstream vs received bit LLRs for $M = 7, K = 96$ (after encoder and before decoder).

Fig. 3.15: MSE and MI Performances with 5% pilots overhead over 2×2 MIMO channel.

GFDM the energies of the data subsymbols are spread over all frequency bins of the pilot subcarriers gaining more frequency diversity. However, if we employ the channel estimation, IFPI GFDM achieves a significant gain with respect to basic GFDM. Because of the interference from data symbols in basic GFDM channel estimation, full information transfer cannot be achieved even at high SNR regions. On the other hand, IFPI GFDM with imperfect channel knowledge (e.g. LS and LMMSE estimations) follows the OFDM performance with slight difference. Note that again here, the slight performance degradation of GFDM with respect to OFDM is due to the noise enhancement of GFDM ZF receiver. Nevertheless, with more advanced receiver structures that we analyze later in Sec. 3.4 and 3.5, we show that GFDM with its broader subcarrier spacing can benefit the frequency diversity more efficiently.

The coded performances of the three receiver types are provided in Fig. 3.16. Employing a robust code-rate of 1/3, OFDM, IFPI GFDM and basic GFDM receivers obtain almost similar BER, though, basic GFDM with LS estimation has 2-3 dB worse BER performance than the rest of receivers with imperfect channel knowledge. Here, due to around 1 dB gap of the genie-aided receivers of OFDM and IFPI GFDM, the latter receiver stays around 0.5 dB behind OFDM when the channel is estimated through pilot transmission. On the other hand, Fig. 3.16b shows that the basic GFDM channel estimation with non-orthogonal pilots becomes entirely unqualified for a high code-rate of 5/6 which is due to its large error floor in channel estimation. Comparing OFDM and IFPI GFDM, we observe that the performance loss in GFDM which is a non-orthogonal waveform is not significant compared to OFDM. Furthermore, in Fig. 3.16b the BER for LS and LMMSE estimations in OFDM as well as IFPI GFDM are identical due to identical channel estimation performances at high SNR regions.



(a) 1/3 code-rate **(b)** 5/6 code-rate

Fig. 3.16: Bit error rate performance for 5% pilots overhead over 2×2 MIMO channel ($M = 7, K = 96$).

Briefly, with the intention to achieve interference-free pilot observation at the receiver side, we proposed a technique to insert orthogonal pilots into a MIMO non-orthogonal multi-carrier. Although, we applied our approach to the GFDM system with its circular filtering, the same concept can also be employed for any non-orthogonal multi-carrier that uses frequency domain channel estimation. Our simulation results confirmed that the IFPI-GFDM maintains the advantages of basic GFDM in terms of its signal characteristics, while it also achieves identical channel estimation performance as in OFDM. In the following, we show that pilot time localization in non-orthogonal MIMO multi-carriers can be exploited for reusing the CP energy in channel estimation, and thus, achieve a better CE performance than CP-less orthogonal pilots in OFDM.

3.4 Bayesian Pilot- and CP-aided Channel Estimation in MIMO Non-Orthogonal Multi-Carriers

The idea to use CP information for enhanced channel estimation was investigated in [QS10] and [KK16] via Kalman filters. However, their approaches are applicable only to orthogonal waveforms. In this work, we take advantage of symbol-time-localization in non-orthogonal waveforms in order to improve the CE performance by reusing the pilots' information from CP. In other words, if we properly localize the energy of the pilots to the end of the block, they would also consist of high energy when copied into the CP. Unlike [QS10] and [KK16], this work presents the approach for any interference-limited linear modulation including their MIMO application.

In this section, our contribution is three-fold:

- Extending our prior works in the previous sections, we derive the well-accepted

LMMSE PIC for close-to-optimal joint channel estimation and equalization of MIMO non-orthogonal multi-carriers. In our derivations, we take into account the pilots' information from the whole block, including its CP. For a non-orthogonal receiver, the pilots' information are contaminated with the data symbols transmitted on the same subcarrier. The contamination from the data symbols is not easily removable, because the interference statistics are correlated with the channel statistics. Thus, in order to derive the MIMO LMMSE estimator with a tractable manner of calculating interference statistics, we propose a new technique to simplify the equivalent channel model. Next, focusing on GFDM waveform and its block-circularity structure, we derive the sequential LMMSE (SLMMSE) estimator of the channel for a reduced complexity version.

- We thoroughly analyze the complexity of each algorithm by counting the number of floating point operations (FLOPs). The complexity comparison of the proposed algorithms, allows us to distinguish and find a trade-off between the algorithms for different applications. For instance, in URLLC a fast and lower complexity transmission is usually preferred to high data-rate while in an application like video-streaming complexity is less noteworthy.

- Through extensive simulations, we evaluate the performance of each algorithm in different scenarios. Given the choice of modulation and coding scheme (MCS), we distinguish the most favorable algorithm in terms of complexity or robustness. In our simulations, we consider MIMO-OFDM transmission as a baseline for comparison.

3.4.1 Review on System Model

Linear Modulation with CP. We consider an $\mathfrak{I} \times \mathfrak{Q}$ MIMO system with digital baseband signals linearly modulated by a modulation matrix $\boldsymbol{\mathcal{M}} \in \mathbb{C}^{N \times N}$. In case of orthogonal modulation as in OFDM, $\boldsymbol{\mathcal{M}}$ would be defined by a DFT matrix, i.e. $\boldsymbol{\mathcal{M}}_{\text{OFDM}} = \boldsymbol{F}^H$. In non-orthogonal systems, $\boldsymbol{\mathcal{M}}$ can perform any linear operation not limited to $\boldsymbol{F}^H$, e.g. a *GFDM* waveform performs circular pulse shaping of individual subcarriers, whereas *single carrier frequency domain multiple access (SC-FDMA)* performs M-point DFT followed by N-point IDFT.

We define the N dimensional Tx signal on each Tx antenna i as

$$\vec{x}_{i,N} = \boldsymbol{\mathcal{M}} \vec{d}_i, \tag{3.58}$$

where $\vec{d}_i \in \mathbb{C}^N$ is a sequence of complex values that need to be transmitted. The vector $\vec{d}_i$ may contain data $\vec{d}_{d,i} \in \mathbb{C}^N$ from a modulation alphabet (e.g. 2^μ-QAM) or reference signal (i.e. pilots) $\vec{d}_{p,i} \in \mathbb{C}^N$ for channel estimation. Therefore, $\vec{d}_i = \vec{d}_{d,i} + \vec{d}_{p,i}$, though, the data symbols and pilots are multiplexed and do not overlap, i.e. $\vec{d}_{d,i} \circ \vec{d}_{p,i} = \vec{0}_N$.

Once the signal is modulated, we insert the CP by copying the last N_{CP} samples of $\vec{x}_i$ to its beginning. The CP insertion can also be integrated to the modulation matrix $\boldsymbol{\mathcal{M}}$ by considering its last N_{CP} rows. Thus, we write the signal associated to block[3] b as

$$\vec{x}_{b,i,N'} = [\boldsymbol{\mathcal{M}}_{\mathrm{CP}}^T \ \boldsymbol{\mathcal{M}}^T]^T \vec{d}_i, \tag{3.59}$$

where $\boldsymbol{\mathcal{M}}_{\mathrm{CP}} \in \mathbb{C}^{N_{\mathrm{CP}} \times N}$ is the last N_{CP} rows of $\boldsymbol{\mathcal{M}}$, N_{CP} denotes the CP length, and $N' = N + N_{\mathrm{CP}}$.

Pilot's influence on CP. The influence of the pilots on CP highly depends on the allocation vector $\vec{d}_{p,i}$, ratio of the CP length over the CP-less block length N_{CP}/N, as well as the modulation matrix $\boldsymbol{\mathcal{M}}$. In order to take into consideration the amount of influence pilots would have on CP, we obtain the time domain signal $\vec{x}_{p,b,i,N'}$ associated to the pilots of block b and antenna i by substituting $\vec{d}_i$ with $\vec{d}_{p,i}$, i.e.

$$\vec{x}_{p,b,i,N'} = [\boldsymbol{\mathcal{M}}_{\mathrm{CP}}^T \ \boldsymbol{\mathcal{M}}^T]^T \vec{d}_{p,i}. \tag{3.60}$$

From (3.60), it is clear that the pilots' energy inside CP does not depend only on $\vec{d}_{p,i}$, but also on the modulation matrix $\boldsymbol{\mathcal{M}}$ and particularly its last N_{CP} rows $\boldsymbol{\mathcal{M}}_{\mathrm{CP}}$. Thus, we define the ratio

$$\kappa \triangleq \frac{\|\vec{x}_{p,b,i,N'}\|^2}{\|\boldsymbol{\mathcal{M}}\vec{d}_{p,i}\|^2}, \tag{3.61}$$

which measures the amount of gain, we could expect from using CP for channel estimation. Intuitively, one can expect a performance gain with respect to the noise floor. Nevertheless, (3.61) can serve only as an upper bound, because the ISI produced—via multi-path—from block a into block b, as well as the self-interference generated within the block b would reduce the gain. A particular case of pilot allocation with high reference signal energy of the CP will be discussed in Sec. 3.4.4.

3.4.2 Single-Input-Single-Output Systems

Assume the channel is block fading over two consecutive Tx signals $\vec{x}_{a,N'}$ and $\vec{x}_{b,N'}$ with each containing N' samples and indexed by a and b, respectively. Since the ISI that is leaked from block a into the CP portion of block b must be taken into account, we assume the first block a transmits only information bits and the second block b contains scattered pilots multiplexed with data within its resource grid. Note that such assumption is a general form, which also includes the case where block a contains pilots as well; although, in order to lower the pilots' overhead, we assume block a transmits no pilot. If the channel is block fading, the associated channel matrix $\boldsymbol{\mathcal{T}}_{2N'} \in \mathbb{C}^{2N' \times 2N'}$ becomes a lower triangular

[3] A block refers to a set of complex symbols modulated by $\boldsymbol{\mathcal{M}}$ and protected by a CP.

Toeplitz matrix with impulse response $\vec{h} \in \mathbb{C}^L$ at the beginning of its first column. At the receiver side, the Rx signal $\vec{y}_{2N'} = [\vec{y}_{a,N'}^T \vec{y}_{b,N'}^T]^T$ can be represented in a linear form as

$$\vec{y}_{2N'} = \boldsymbol{\mathcal{T}}_{2N'} \vec{x}_{2N'} + \vec{w}_{2N'}, \tag{3.62}$$

where $\vec{x}_{2N'} = [\vec{x}_{a,N'}^T \vec{x}_{b,N'}^T]^T$ and $\vec{w}_{2N'} \sim \mathcal{N}_C(\vec{0}, \boldsymbol{I}_{2N'} \sigma_w^2)$ is the AWGN vector.

Since the pilots are inserted into the second signal, one should observe parts of $\vec{y}_{2N'}$ where the pilots have energy. Such task should not only include the CP-less signal, but the whole $\vec{x}_{b,N'}$ that contains the CP part as well (if CP contains sufficient energy of the pilots). Moreover, in order to take into account the ISI leakage from $\vec{x}_{a,N'}$ into $\vec{x}_{b,N'}$, we consider $\vec{y}_{N''} \in \mathbb{C}^{N''}$ as the last $N'' = N + 2N_{\text{CP}}$ samples of $\vec{y}_{2N'}$. Hence, the observed Rx signal follows

$$\vec{y}_{N''} = \boldsymbol{\mathcal{T}}_{N''} \vec{x}_{N''} + \vec{w}_{N''}, \tag{3.63}$$

where $\boldsymbol{\mathcal{T}}_{N''} \in \mathbb{C}^{N'' \times N''}$ is the Toeplitz channel matrix with $\vec{h}$ on its first column. $\vec{x}_{N''}$ and $\vec{w}_{N''}$ contain the last N'' samples of $\vec{x}_{2N'}$ and $\vec{w}_{2N'}$, respectively. Further, $\vec{x}_{N''}$ can be expressed as

$$\begin{aligned} \vec{x}_{N''} &= \vec{x}_{p,N''} + \vec{x}_{d,N''} \\ &= \boldsymbol{\mathcal{M}}_{N''} ([\vec{0}_N^T \ \vec{d}_p^T]^T + [\vec{d}_{d,a}^T \ \vec{d}_{d,b}^T]^T), \end{aligned} \tag{3.64}$$

where $\vec{d}_p$ is the vector containing pilots, $\vec{d}_{d,a}$ and $\vec{d}_{d,b}$ are the vectors associated to the information sequence in blocks a and b, respectively. Moreover, the matrix $\boldsymbol{\mathcal{M}}_{N''}$ is defined as

$$\boldsymbol{\mathcal{M}}_{N''} \triangleq \left(\boldsymbol{\mathcal{M}}_{\text{CP}} \oplus [\boldsymbol{\mathcal{M}}_{\text{CP}}^T \ \boldsymbol{\mathcal{M}}^T]^T \right) \in \mathbb{C}^{N'' \times 2N}. \tag{3.65}$$

In order to observe parts of $\vec{y}_{N''}$ where pilots have energy, we define the windowing matrix $\boldsymbol{J} \in \mathbb{R}^{N' \times N''}$ as

$$\boldsymbol{J} \triangleq [\boldsymbol{0}_{N' \times N_{\text{CP}}} \ \boldsymbol{I}_{N'}]. \tag{3.66}$$

Hence, the measurement signal of interest follows:

$$\vec{y}_{N',0} = \boldsymbol{J} \vec{y}_{N''} = \boldsymbol{J} (\boldsymbol{\mathcal{T}}_{N''} \vec{x}_{N''} + \vec{w}_{N''}). \tag{3.67}$$

Note that here, by considering $\vec{y}_{N''}$ and then windowing it into $\vec{y}_{N',0}$, not only we observe the parts of signal that pilots have energy, but also, we include the interference terms that leak from the last N_{CP} samples of block a into N'-length block b. The linear model (3.67) contains a defective channel matrix $\boldsymbol{\mathcal{T}}_{N''}$, which is not unitary diagonalizable. If $\boldsymbol{\mathcal{T}}_{N''}$ were to be unitary-diagonalizable, it had to be a normal matrix. However, a matrix is simultaneously triangular and normal if and only if it is diagonal, which is not true for a multi-tap channel matrix $\boldsymbol{\mathcal{T}}_{N''}$. In addition, Appendix A.2 also provides a mathematical proof of $\boldsymbol{\mathcal{T}}_{N''}$ being a defective matrix. On the other hand, if $\boldsymbol{\mathcal{T}}_{N''}$ could be diagonalized, not only the complexity could be reduced, but also the theoretical calculations of the interference statistics could be done in a tractable and straightforward fashion. Thus, in

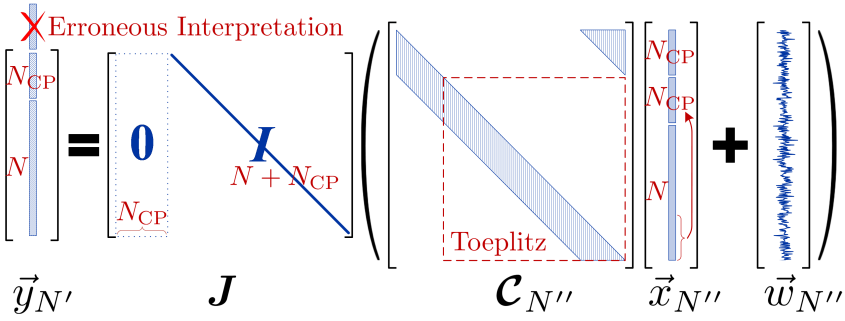


Fig. 3.17: Matrix structures in the modified system model (3.68). The above structure allows to treat the Toeplitz channel matrix as circulant, while its observation $\vec{y}_{N'}$ is identical to the standard system model (3.67).

order to overcome with the defectiveness of $\boldsymbol{\mathcal{T}}_{N''}$, we treat the channel matrix as circulant, despite the fact that such treatment would lead to erroneous interpretation. Therefore, we redefine the observed signal as

$$\vec{y}_{N'} \triangleq \boldsymbol{J}(\boldsymbol{\mathcal{C}}_{N''}\vec{x}_{N''} + \vec{w}_{N''}), \tag{3.68}$$

where $\boldsymbol{\mathcal{C}}_{N''} \in \mathbb{C}^{N''\times N''}$ is a circulant matrix with respect to $\vec{h}$. Note that in (3.68), the matrix $\boldsymbol{\mathcal{C}}_{N''}$ is not the actual channel matrix, but a unitary-diagonalizable matrix, which allows to model the actual Toeplitz channel matrix of size N' (as depicted in Fig. 3.17). In fact, treating $\boldsymbol{\mathcal{T}}_{N''}$ as $\boldsymbol{\mathcal{C}}_{N''}$ would cause erroneous interpretation to the first N_{CP} samples of $\vec{y}_{N''}$. But since the windowing matrix $\boldsymbol{J}$ selects the last N' samples of $\vec{y}_{N''}$ and discards the first N_{CP} samples, such erroneous interpretation is pulled out and thus, the observation $\vec{y}_{N'}$ would be identical to the result of (3.67), i.e. $\vec{y}_{N'} = \vec{y}_{N',0}$.

The new model (3.68) provides the benefit to diagonalize the channel matrix. Thus, we have:

$$\vec{y}_{N'} = \boldsymbol{J}\left(\boldsymbol{F}_{N''}^{H}\mathrm{diag}(\vec{H}_{N''})\boldsymbol{F}_{N''}\vec{x}_{N''} + \vec{w}_{N''}\right), \tag{3.69}$$

with $\vec{H}_{N''} = \sqrt{N''}\boldsymbol{F}_{N'',L}\vec{h}$. Since we aim to estimate the parameters vector $\vec{h}$, we reorder (3.69) as

$$\vec{y}_{N'} = \boldsymbol{J}\left(\sqrt{N''}\boldsymbol{F}_{N''}^{H}\boldsymbol{X}_{N''}\boldsymbol{F}_{N'',L}\vec{h} + \vec{w}_{N''}\right), \tag{3.70}$$

where $\boldsymbol{X}_{N''} = \mathrm{diag}(\boldsymbol{F}_{N''}\vec{x}_{N''})$. Defining the matrix $\boldsymbol{Q} \in \mathbb{C}^{N'\times L}$ as

$$\boldsymbol{Q} \triangleq \sqrt{N''}\boldsymbol{J}\boldsymbol{F}_{N''}^{H}\boldsymbol{X}_{N''}\boldsymbol{F}_{N'',L}, \tag{3.71}$$

we obtain a general linear model in which, we can calculate the LMMSE estimate of $\vec{h}$ via Bayesian Gauss-Markov theorem [Kay93]. Thus,

$$\hat{\vec{h}}_{\mathrm{CPLMMSE}} = \boldsymbol{\Sigma}_{hh}\boldsymbol{Q}_p^{H}(\boldsymbol{Q}_p\boldsymbol{\Sigma}_{hh}\boldsymbol{Q}_p^{H} + \boldsymbol{\Sigma}_{EE,1}^{\boldsymbol{J}})^{-1}\vec{y}_{N'}, \tag{3.72}$$

$$\widehat{\boldsymbol{\Sigma}}_{hh} = \boldsymbol{\Sigma}_{hh} - \boldsymbol{\Sigma}_{hh}\boldsymbol{Q}_p^{H}(\boldsymbol{Q}_p\boldsymbol{\Sigma}_{hh}\boldsymbol{Q}_p^{H} + \boldsymbol{\Sigma}_{EE,1}^{\boldsymbol{J}})^{-1}\boldsymbol{Q}_p\boldsymbol{\Sigma}_{hh}, \tag{3.73}$$

where $\boldsymbol{Q}_p = \sqrt{N''}\boldsymbol{J}\boldsymbol{F}_{N''}^H \text{diag}(\boldsymbol{F}_{N''}\vec{x}_{p,N''})\boldsymbol{F}_{N'',L}$ is the observation matrix with respect to the pilots signal $\vec{x}_{p,N''}$ defined in (3.64). $\boldsymbol{\Sigma}_{EE}^{\boldsymbol{J}} \in \mathbb{C}^{N'\times N'}$ is the covariance matrix of noise plus interference term with respect to data signal $\vec{x}_{d,N''}$ as well as noise vector $\boldsymbol{J}\vec{w}_{N''}$. We calculate $\boldsymbol{\Sigma}_{EE,1}^{\boldsymbol{J}}$ in analogous way as in Sec. 3.2 while we also take into account the additional matrices $\boldsymbol{J}\boldsymbol{F}_{N''}^H$ at the left hand side of the observed signal. Nevertheless, in Appendix A.3, we provide the derivations for the MIMO version of $\boldsymbol{\Sigma}_{EE,1}^{\boldsymbol{J}}$.

3.4.3 Extension to MIMO

Similar to (2.18), the MIMO model for N'' samples of each Rx antenna is given by

$$\vec{y}_{\mathfrak{Q}N''} = \overline{\boldsymbol{\mathcal{T}}}_{N''}\vec{x}_{\mathfrak{I}N''} + \vec{w}_{\mathfrak{Q}N''}, \tag{3.74}$$

where $\vec{y}_{\mathfrak{Q}N''} \in \mathbb{C}^{\mathfrak{Q}N''}$ and $\vec{w}_{\mathfrak{Q}N''} \in \mathbb{C}^{\mathfrak{Q}N''}$ are the vectors of $\mathfrak{Q}$ concatenated Rx signals and noises on $\mathfrak{Q}$ antennas, respectively. $\vec{x}_{\mathfrak{I}N''} = [\vec{x}_{1,N''}^T \ \cdots \ \vec{x}_{\mathfrak{I},N''}^T]^T \in \mathbb{C}^{\mathfrak{I}N''}$ is the vector of $\mathfrak{I}$ concatenated Tx signals of size N''. Note that here, each Tx signal $\vec{x}_{i,N''}$ is generated in the same manner as of $\vec{x}_{N''}$ in (3.64), though, for the Tx antenna i. Further, the matrix $\overline{\boldsymbol{\mathcal{T}}}_{N''} \in \mathbb{C}^{\mathfrak{Q}N''\times\mathfrak{I}N''}$ is the MIMO version of $\mathfrak{Q}\times\mathfrak{I}$ Toeplitz matrices $\boldsymbol{\mathcal{T}}_{N'',qi}$ of the wireless channels between antennas i and q. In a similar fashion as explained in Sec. 3.4.2, we introduce the windowing matrix $\overline{\boldsymbol{J}} = \boldsymbol{I}_{\mathfrak{Q}} \otimes \boldsymbol{J}$, and we treat the triangular channels as circulant ones i.e.

$$\vec{y}_{\mathfrak{Q}N'} = \overline{\boldsymbol{J}}\left(\overline{\boldsymbol{\mathcal{C}}}_{N''}\vec{x}_{\mathfrak{I}N''} + \vec{w}_{\mathfrak{Q}N''}\right), \tag{3.75}$$

where $\overline{\boldsymbol{\mathcal{C}}}_{N''}$ is the MIMO version of $\mathfrak{Q}\times\mathfrak{I}$ circulant channels $\boldsymbol{\mathcal{C}}_{N'',qi} \in \mathbb{C}^{N''\times N''}$ with the same structure of (2.19), though, for N'' samples. The vector $\vec{y}_{\mathfrak{Q}N'} = [\vec{y}_{1,N'}^T \ \cdots \ \vec{y}_{\mathfrak{Q},N'}^T]^T$ consists of $\mathfrak{Q}$ concatenated Rx signals of $\mathfrak{Q}$ Rx antennas. Moreover, since the product of $\overline{\boldsymbol{J}}\vec{w}_{\mathfrak{Q}N''}$ selects only the last N' samples of individual noise vectors of each Rx antenna, it can be simplified into a shorter vector $(\vec{w}_{q,N'}^T)_{q=1,\cdots,\mathfrak{Q}}^T \in \mathbb{C}^{\mathfrak{Q}N'}$, in which, its samples are still white.

Diagonalizing the channel matrices and reshaping the expression (3.75) for estimation of the parameters matrix $\boldsymbol{h} \in \mathbb{C}^{\mathfrak{I}L\times\mathfrak{Q}}$, we obtain the observation matrix $\boldsymbol{y}_{N'}$ as

$$\boldsymbol{y}_{N'} = \boldsymbol{J}\boldsymbol{F}_{N''}^H\boldsymbol{X}_{\mathfrak{I}N''}\bar{\mathbf{F}}_{N'',L}\boldsymbol{h} + \boldsymbol{w}_{N'}, \tag{3.76}$$

with

$$\boldsymbol{y}_{N'} \triangleq (\vec{y}_{1,N'} \ \cdots \ \vec{y}_{\mathfrak{Q},N'}) \in \mathbb{C}^{N'\times\mathfrak{Q}}, \tag{3.76a}$$

$$\boldsymbol{X}_{s,\mathfrak{I}N''} \triangleq (\boldsymbol{X}_{s,1,N''} \ \cdots \ \boldsymbol{X}_{s,\mathfrak{I},N''}) \in \mathbb{C}^{N''\times\mathfrak{I}N''}, \tag{3.76b}$$

$$\bar{\boldsymbol{F}}_{N'',L} \triangleq \boldsymbol{I}_{\mathfrak{I}} \otimes \sqrt{N''}\mathbf{F}_{N'',L} \in \mathbb{C}^{\mathfrak{I}N''\times\mathfrak{I}L}, \tag{3.76c}$$

$$\boldsymbol{w}_{N'} \triangleq (\vec{w}_{1,N'} \ \cdots \ \vec{w}_{\mathfrak{Q},N'}) \in \mathbb{C}^{N'\times\mathfrak{Q}}, \tag{3.76d}$$

where $\boldsymbol{X}_{\mathfrak{I}N''} = \boldsymbol{X}_{p,\mathfrak{I}N''} + \boldsymbol{X}_{d,\mathfrak{I}N''}$ and $\boldsymbol{X}_{s,i,N''} = \text{diag}(\boldsymbol{F}_{N''}\vec{x}_{s,i,N''})$ for $s \in \{p,d\}$.

Similar to Sec. 3.4.2, we define the observation matrix $\boldsymbol{\mathcal{Q}}_p$ as

$$\boldsymbol{\mathcal{Q}}_p \triangleq \boldsymbol{J}\boldsymbol{F}_{N''}^H \boldsymbol{X}_{p,\mathfrak{I}N''} \bar{\mathbf{F}}_{N'',L} \in \mathbb{C}^{N' \times \mathfrak{I}L} \tag{3.77}$$

for the linear MIMO expression (3.76). Subsequently, we calculate the LMMSE estimation of the q-th column of $\boldsymbol{h}$ as

$$\hat{\vec{h}}_{\text{CPLMMSE},q} = \boldsymbol{\Sigma}_{hh}^q \boldsymbol{\mathcal{Q}}_p^H (\boldsymbol{\mathcal{Q}}_p \boldsymbol{\Sigma}_{hh}^q \boldsymbol{\mathcal{Q}}_p^H + \boldsymbol{\Sigma}_{EE,q}^{\boldsymbol{J}})^{-1} \vec{y}_{q,N'}, \tag{3.78}$$

$$\hat{\bar{\boldsymbol{\Sigma}}}_{\text{HH}} = (\boldsymbol{I}_{\mathfrak{Q}} \otimes \bar{\mathbf{F}}_{N'',L}) \hat{\bar{\boldsymbol{\Sigma}}}_{hh} (\boldsymbol{I}_{\mathfrak{Q}} \otimes \bar{\mathbf{F}}_{N'',L}^H), \tag{3.79}$$

$$\hat{\bar{\boldsymbol{\Sigma}}}_{hh} = \bar{\boldsymbol{\Sigma}}_{hh} - \left(\bar{\boldsymbol{\Sigma}}_{hh} \bar{\boldsymbol{\mathcal{Q}}}_p^H (\bar{\boldsymbol{\mathcal{Q}}}_p \bar{\boldsymbol{\Sigma}}_{hh} \bar{\boldsymbol{\mathcal{Q}}}_p^H + \bar{\boldsymbol{\Sigma}}_{EE}^{\boldsymbol{J}})^{-1} \bar{\boldsymbol{\mathcal{Q}}}_p \bar{\boldsymbol{\Sigma}}_{hh} \right), \tag{3.80}$$

where $\bar{\boldsymbol{\mathcal{Q}}}_p = \boldsymbol{I}_{\mathfrak{Q}} \otimes \boldsymbol{\mathcal{Q}}_p$. The noise-plus-interference covariance $\bar{\boldsymbol{\Sigma}}_{EE}^{\boldsymbol{J}} = \boldsymbol{\Sigma}_{EE,1}^{\boldsymbol{J}} \oplus \cdots \oplus \boldsymbol{\Sigma}_{EE,\mathfrak{Q}}^{\boldsymbol{J}}$ is block diagonal, and the calculation of its individual matrix blocks is provided in Appendix A.3. The frequency domain CE covariance matrix $\hat{\bar{\boldsymbol{\Sigma}}}_{\text{HH}}$ has a block-based structure, which follows

$$\hat{\bar{\boldsymbol{\Sigma}}}_{\text{HH}} = \hat{\boldsymbol{\Sigma}}_{\text{HH},q=1} \oplus \ \cdots \ \oplus \hat{\boldsymbol{\Sigma}}_{\text{HH},q=\mathfrak{Q}}, \tag{3.81}$$

and

$$\hat{\boldsymbol{\Sigma}}_{\text{HH},q} = \begin{pmatrix} \hat{\boldsymbol{\Sigma}}_{\text{HH},q}^{i=1} & \hat{\boldsymbol{\Sigma}}_{\text{HH},q}^{12} & \cdots & \hat{\boldsymbol{\Sigma}}_{\text{HH},q}^{1\mathfrak{I}} \\ \hat{\boldsymbol{\Sigma}}_{\text{HH},q}^{21} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \hat{\boldsymbol{\Sigma}}_{\text{HH},q}^{(\mathfrak{I}-1)\mathfrak{I}} \\ \hat{\boldsymbol{\Sigma}}_{\text{HH},q}^{\mathfrak{I}1} & \cdots & \hat{\boldsymbol{\Sigma}}_{\text{HH},q}^{\mathfrak{I}(\mathfrak{I}-1)} & \hat{\boldsymbol{\Sigma}}_{\text{HH},q}^{i=\mathfrak{I}} \end{pmatrix}, \tag{3.82}$$

Note that the individual matrices in (3.82) are already calculated via (3.79). Fig. 3.18 shows an example of $\hat{\boldsymbol{\Sigma}}_{\text{HH},q}$ for two Tx antennas. In this case, if we assume $\boldsymbol{\mathcal{Q}}_p = [\boldsymbol{Q}_{p,1},\ \boldsymbol{Q}_{p,2}]$, where $\boldsymbol{Q}_{p,i}$ denotes the observation matrix associated to the Tx antenna i, we would have $\hat{\boldsymbol{\Sigma}}_{\text{HH},q}^{i=1} = \boldsymbol{Q}_{p,1}^H (\boldsymbol{Q}_{p,1} \boldsymbol{\Sigma}_{hh,1q} \boldsymbol{Q}_{p,1}^H + \boldsymbol{Q}_{p,2} \boldsymbol{\Sigma}_{hh,2q} \boldsymbol{Q}_{p,2}^H + \boldsymbol{\Sigma}_{EE,q}^{\boldsymbol{J}})^{-1} \boldsymbol{Q}_{p,1}$ on the first diagonal entry, whereas $\hat{\boldsymbol{\Sigma}}_{\text{HH},q}^{12} = \boldsymbol{Q}_{p,1}^H (\boldsymbol{Q}_{p,1} \boldsymbol{\Sigma}_{hh,1q} \boldsymbol{Q}_{p,1}^H + \boldsymbol{Q}_{p,2} \boldsymbol{\Sigma}_{hh,2q} \boldsymbol{Q}_{p,2}^H + \boldsymbol{\Sigma}_{EE,q}^{\boldsymbol{J}})^{-1} \boldsymbol{Q}_{p,2}$ would be on the upper off-diagonal of $\hat{\boldsymbol{\Sigma}}_{\text{HH},q}$. For simplicity of the iterative joint channel estimation and equalization in Sec. 3.4.5, we will only consider the diagonal matrix blocks, although, ignoring the off-diagonal matrices may cost some performance degradation.

3.4.4 Application to GFDM

Pilot insertion. Consider the GFDM modulation described in Sec. 2.1. In addition, we multiplex the pilots with data in block b, while block a transmits only data. In order to include a large portion of pilots within the CP of block b, we set the CP duration N_{CP} to the length of one subsymbol, i.e. $N_{\text{CP}} = K$ samples. As depicted in Fig. 3.19a, we insert one pilot into the first and one pilot into the last subsymbol of the pilot subcarriers. Pilot-subcarrier-spacing is denoted by Δk and in order to have sufficient number of observations

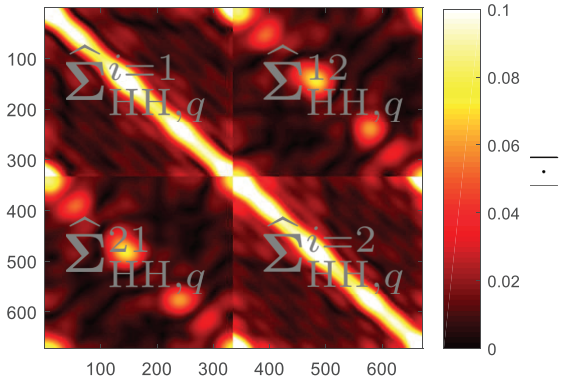


Fig. 3.18: Example of the frequency domain CE covariance matrix $\widehat{\boldsymbol{\Sigma}}_{\mathrm{HH},q}$ for GFDM with $M = 7$ and $K = 48$ over $2 \times \mathfrak{Q}$ MIMO channel and 0 dB SNR. Pilots are generated in form of complex randoms with unit energy and they are inserted into the first and into the last subsymbols of GFDM with $\Delta k = 2$. The diagonal matrix blocks are the CE covariance matrices that depend on each Tx antenna separately. The off-diagonal matrix blocks are the CE covariance matrices that occur due to the contamination of pilots and data transmitted from different Tx antennas and received on the antenna q.

for estimating L channel taps, we consider $\lfloor \frac{\mathfrak{J}K}{\Delta k} \rfloor \geq \mathfrak{J}\mathfrak{Q}L$. Such rectangular grid pilot pattern reduces the self-interference (see Sec. 3.2), and in addition, it allows to achieve considerable amount of pilots' power inside CP[4].

If $L < N_{\mathrm{CP}}$, block a would produce ISI into the CP portion of block b. Fig. 3.19b shows an example of the two GFDM blocks including their CP.

Linear Model. In Appendix A.4, we show that by block-diagonalizing the effective channel of the N-dimensional receive signal $\vec{y}_{q,N}$ on antenna q, we can split $\vec{y}_{q,N}$ into M sequences $\vec{Y}_{m,q,K} \in \mathbb{C}^K$ of K-dimension, and for the sake of complexity reduction, process them separately.

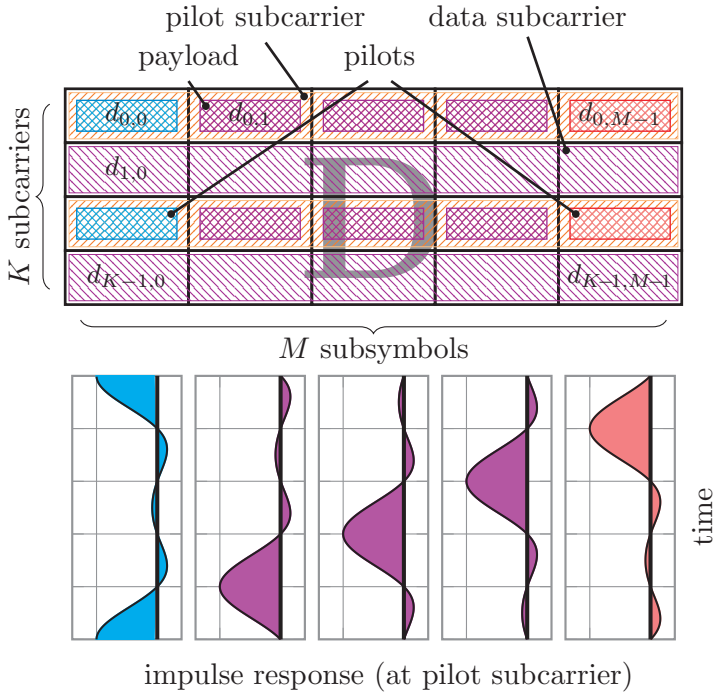
Let $\boldsymbol{Z}_m \triangleq \vec{\omega}_m^T \otimes \boldsymbol{I}_K \in \mathbb{C}^{K \times N}$ with $\vec{\omega}_m^T$ being the m-th row of $\boldsymbol{F}_M$. The K dimensional m-th sequence $\vec{Y}_{m,q,K}$ follows

$$\vec{Y}_{m,q,K} = \sum_{i=1}^{\mathfrak{J}} \boldsymbol{Z}_m \boldsymbol{\mathcal{C}}_{N,qi} \boldsymbol{A} \vec{d}_i + \boldsymbol{Z}_m \vec{w}_{q,N}, \tag{3.83}$$

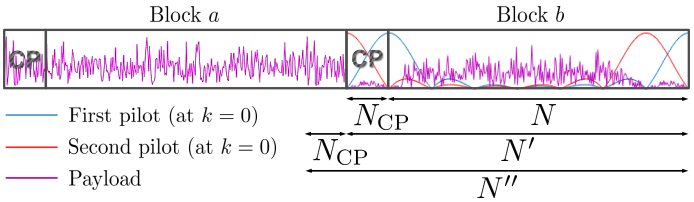
with the observation matrix

$$\tilde{\boldsymbol{X}}_{p,m} = \sqrt{N} \boldsymbol{Z}_m \boldsymbol{F}_N^H \boldsymbol{X}_{p,\mathfrak{J}N} \bar{\mathbf{F}}_{N,L} \in \mathbb{C}^{K \times \mathfrak{J}L}, \tag{3.84}$$

[4] In order to use the pilots' information from cyclic prefix, they should possess high energy at the end of the time-domain *block* (also so-called *symbol* in OFDM terminology) and therefore, subsymbol-time-localization is significantly important here. The interference-free pilot design that was discussed in Sec. 3.3 is not beneficial in this context, because the orthogonal pilots cannot be localized inside the time domain symbol. Indeed, they have identical properties of OFDM pilots, which have uniform energy over the symbol duration in time domain.



(a) Pilot insertion in $\mathbf{D}$ matrix of block b ($M = 5$).



(b) Transmit signal over two blocks ($M = 7$), while a large portion of pilots' power is copied into the CP. Note that the power of the first and the second pilots would be summed in the observation signal.

Fig. 3.19: Example of pilot insertion into the second GFDM block while the first block transmits only information bits.

to be used in SLMMSE CE. Here, we obtain $\boldsymbol{X}_{p,\mathfrak{I}N}$ and $\bar{\mathbf{F}}_{N,L}$ by substituting N'' with N in (3.76b) and (3.76c), respectively.

Sequential LMMSE Estimation. The SLMMSE is a recursive approach that updates the LMMSE estimator when new information becomes available. However, such method is desirable if the statistical properties of the channel do not vary significantly in time, which is also the case in our initial assumption in Sec. 2.3. Moreover, the SLMMSE estimation requires *diagonal* error covariance of the observations [Kay93, Sec. 12A], which is not true for non-orthogonal waveforms, i.e. interference limited. In this case, enforcing the SLMMSE to observations with non-diagonal covariance matrix would cause the interference to propagate on every iteration. Thus, the following possibilities would remain:

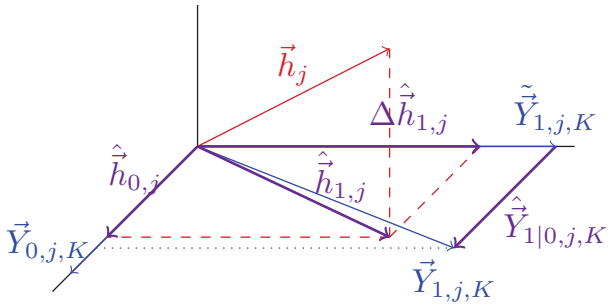


Fig. 3.20: SLMMSE concept with one iteration, i.e. $m = [0, 1]$.

a. If the estimation error due to the interference $\boldsymbol{\mathcal{C}}_{N,qi}\boldsymbol{A}\vec{d}_{d,i}$ is much smaller than the noise variance σ_w^2, ignore the interference term and set its covariance to zero.

b. If the estimation error due to the interference $\boldsymbol{\mathcal{C}}_{N,qi}\boldsymbol{A}\vec{d}_{d,i}$ is *not* much smaller than the noice variance σ_w^2,

1. Pre-whiten the observation $\vec{y}_{q,N}$ by $\boldsymbol{\Sigma}_{EE}^{-1/2}$. However, due to prohibitively complex approach, we do not consider it here.
2. Estimate the interference and cancel it recursively via LMMSE-PIC.

If the selected MCS is a robust one, e.g. QPSK 1/2, or the Rx signal power is small, i.e. noise is dominant with respect to the interference term, we ignore the interference by setting its covariance to zero. In this case, we first estimate an initial guess of the channel and its covariance from the first splitted sequence of length K via LMMSE estimation. Setting $m = 0$, for each Rx antenna q, we calculate

$$\hat{\vec{h}}_{0,q} = \left((\boldsymbol{\Sigma}_{hh}^{q})^{-1} + \frac{\tilde{\boldsymbol{X}}_{p,0}^{H}\tilde{\boldsymbol{X}}_{p,0}}{\sigma_w^2}\right)^{-1} \frac{\tilde{\boldsymbol{X}}_{p,0}^{H}}{\sigma_w^2}\vec{Y}_{0,q,K}, \tag{3.85}$$

$$\hat{\boldsymbol{\Sigma}}_{hh,0}^{q} = \left((\boldsymbol{\Sigma}_{hh}^{q})^{-1} + \frac{\tilde{\boldsymbol{X}}_{p,0}^{H}\tilde{\boldsymbol{X}}_{p,0}}{\sigma_w^2}\right)^{-1}. \tag{3.86}$$

Once an initial estimate of the channel is available, we update $\hat{\vec{h}}_{0,q}$ by adding a component $\Delta\hat{\vec{h}}_{m,q}$ that is orthogonal to $\hat{\vec{h}}_{0,q}$ (see Fig. 3.20). Thus, incrementing $m \leftarrow m + 1$ for the new pilot data, we find the LMMSE estimator of $\vec{Y}_{m,q,K}$ based on $\vec{Y}_{m-1,q,K}$, i.e.

$$\hat{\vec{Y}}_{m|m-1,q,K} = \tilde{\boldsymbol{X}}_{p,m}\hat{\vec{h}}_{m-1,q}, \tag{3.87}$$

and subtract out from the next sequence parts that could be anticipated from the result of past observations, i.e.

$$\tilde{\vec{Y}}_{m,q,K} = \vec{Y}_{m,q,K} - \hat{\vec{Y}}_{m|m-1,q,K}. \tag{3.88}$$

procedure SLMMSE($\tilde{\boldsymbol{X}}_{p,m}, \boldsymbol{\Sigma}^q_{hh}, \vec{Y}_{m,q,K}$)
  $m \leftarrow 0$
  Calculate the covariance $\widehat{\boldsymbol{\Sigma}}^q_{hh,0}$. ▷ i.e. (3.86)
  Estimate $\hat{\vec{h}}_{0,q}$. ▷ i.e. (3.85)
  while $m < M$ **do**
    $m \leftarrow m + 1$
    Estimate $\hat{\vec{Y}}_{m|m-1,q,K}$. ▷ i.e. (3.87)
    Obtain $\widetilde{\vec{Y}}_{m,q,K}$ by subtracting out the old information. ▷ i.e. (3.88)
    Find the covariances $\boldsymbol{\Sigma}_{h\tilde{Y},m}, \boldsymbol{\Sigma}_{\tilde{Y}\tilde{Y},m}$ and $\widehat{\boldsymbol{\Sigma}}^q_{hh,m}$ via the updated matrix $\tilde{\boldsymbol{X}}_{p,m}$. ▷ i.e. (3.90), (3.91) and (3.93)
    Update $\hat{\vec{h}}_{m,q}$ via (3.92).
  end while
  return $\hat{\vec{h}}_{m,q}$
end procedure

Fig. 3.21: Procedure for sequential LMMSE of $\vec{h}_q$ when noise is dominant.

Calling $\widetilde{\vec{Y}}_{m,q,K}$ the error of observation, it will be orthogonal to $\vec{Y}_{m-1,q,K}$ and it represents the new information that $\vec{Y}_{m,q,K}$ contributes to the estimation of $\vec{h}_q$. Therefore, the projection of $\vec{h}_q$ along the vector $\widetilde{\vec{Y}}_{m,q,K}$ is the desired correction

$$\Delta\hat{\vec{h}}_{m,q} = \boldsymbol{\Sigma}_{h\tilde{Y},m}\boldsymbol{\Sigma}^{-1}_{\tilde{Y}\tilde{Y},m}\widetilde{\vec{Y}}_{m,q,K}, \tag{3.89}$$

with

$$\boldsymbol{\Sigma}_{h\tilde{Y},m} = \widehat{\boldsymbol{\Sigma}}^q_{hh,m-1}\tilde{\boldsymbol{X}}^H_{p,m}, \tag{3.90}$$

$$\boldsymbol{\Sigma}_{\tilde{Y}\tilde{Y},m} = \tilde{\boldsymbol{X}}_{p,m}\widehat{\boldsymbol{\Sigma}}^q_{hh,m-1}\tilde{\boldsymbol{X}}^H_{p,m} + \boldsymbol{I}_K\sigma^2_w. \tag{3.91}$$

Adding the correction term $\Delta\hat{\vec{h}}_{m,q}$ to the old estimate, we get a new finer estimate

$$\hat{\vec{h}}_{m,q} = \hat{\vec{h}}_{m-1,q} + \Delta\hat{\vec{h}}_{m,q}, \tag{3.92}$$

and its covariance

$$\widehat{\boldsymbol{\Sigma}}^q_{hh,m} = \left(\boldsymbol{I}_{\jmath L} - \boldsymbol{\Sigma}_{h\tilde{Y},m}\boldsymbol{\Sigma}^{-1}_{\tilde{Y}\tilde{Y},m}\tilde{\boldsymbol{X}}_{p,m}\right)\widehat{\boldsymbol{\Sigma}}^q_{hh,m-1}. \tag{3.93}$$

Summarizing the above description, Fig. 3.21, presents the procedure for sequential LMMSE estimation of $\vec{h}_q$.

Low Complexity CP-Aided Channel Estimation. In Section 3.4.3, the complexity of the LMMSE estimator grows with $\mathcal{O}(N'^3)$. Through the discrete ZAK transform (see Appendix A.4), we managed to block-diagonalize the effective channel and correspondingly, split the GFDM signal into smaller sequences. Such block-diagonalization is

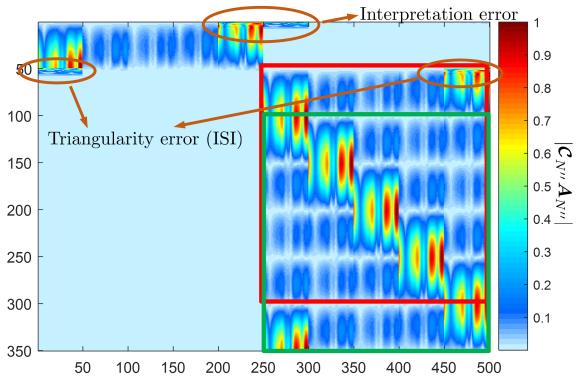


Fig. 3.22: Example of the effective channel matrix $|\mathcal{C}_{N''}\boldsymbol{A}_{N''}|$ (red: $\boldsymbol{S}_{\mathrm{CP}}$, green: $\mathcal{C}_N\boldsymbol{A}$) for $M = 5, K = 50$. The green rectangle is block circulant, whereas the red rectangle has some irregularity at the top-right corner of the block due to ISI. Such irregularity yields off-diagonal elements when block-diagonalization of $\boldsymbol{S}_{\mathrm{CP}}$ is attempted.

relying on block-circularity of the effective channel. However, if we consider $\vec{y}_{N'}$ with $N + N_{\mathrm{CP}}$ samples in order to include the CP energy, the effective channel matrix is no longer square nor block-circulant. Hence, in order to have the effective channel as a square matrix and therefore, approximate a lower complexity version of (3.72), we consider the first N samples of $\vec{y}_{N'}$ (i.e. N_{CP} samples of CP plus $N - N_{\mathrm{CP}}$ samples at the beginning of $\vec{y}_N$) and we use $\hat{\vec{h}}_{M-1,q}$ ($\hat{\vec{h}}_{m,q}$ with $m = M - 1$) of (3.92) as the initial estimate of CP-SLMMSE. In this case, we replace $\boldsymbol{J}$ by $\boldsymbol{J}' = [\boldsymbol{0}_{N\times N_{\mathrm{CP}}}\ \boldsymbol{I}_N\ \boldsymbol{0}_{N\times N_{\mathrm{CP}}}] \in \mathbb{C}^{N\times N''}$ and approximately block-diagonalize the effective channel matrix via DZT. The effective channel matrix for two GFDM blocks is

$$\boldsymbol{S}'_{\mathrm{CP}} = \boldsymbol{J}'\mathcal{C}_{N''}\boldsymbol{A}_{N''} \in \mathbb{C}^{N\times 2N}, \tag{3.94}$$

with $\boldsymbol{A}_{N''} = \left(\boldsymbol{A}_{\mathrm{CP}} \oplus [\boldsymbol{A}_{\mathrm{CP}}^T\ \boldsymbol{A}^T]^T\right) \in \mathbb{C}^{N''\times 2N}$ and $\boldsymbol{A}_{\mathrm{CP}}$ being the last N_{CP} rows of $\boldsymbol{A}$. The left matrix $\boldsymbol{A}_{\mathrm{CP}}$ in $\boldsymbol{A}_{N''}$ is associated to the last N_{CP} samples of the transmit signal from GFDM block a. Also, note that replacing $\boldsymbol{J}$ by $\boldsymbol{J}'$ would cause the lower triangular channel matrix of size $N \times N$ located at the center of $\mathcal{C}_{N''}$ to be selected.

Next, we take the last N columns of $\boldsymbol{S}'_{\mathrm{CP}}$ which allows to focus on the second block where pilots are inserted

$$\boldsymbol{S}_{\mathrm{CP}} = \boldsymbol{S}'_{\mathrm{CP}}[\boldsymbol{0}_N\ \boldsymbol{I}_N]^T \in \mathbb{C}^{N\times N}. \tag{3.95}$$

As a consequence, the approximately block-diagonalized effective channel $\mathfrak{B}'$ follows:

$$\mathfrak{B}' \approx \boldsymbol{Z}\boldsymbol{S}_{\mathrm{CP}}\boldsymbol{Z}^H. \tag{3.96}$$

Note that the product of $\boldsymbol{Z}\boldsymbol{S}_{\mathrm{CP}}\boldsymbol{Z}^H$ is not exactly block-diagonal because the channel matrix considered in $\boldsymbol{S}_{\mathrm{CP}}$ is triangular and not circulant. Moreover, considering the interference that leaks from the previous block, the effective channel matrix $\boldsymbol{S}_{\mathrm{CP}}$ is not entirely block-circulant. Fig. 3.22 shows an example of the effective channel matrix for N' observations. The green and red rectangles are $\mathcal{C}_N$ and $\boldsymbol{S}_{\mathrm{CP}}$, respectively. One can see that the

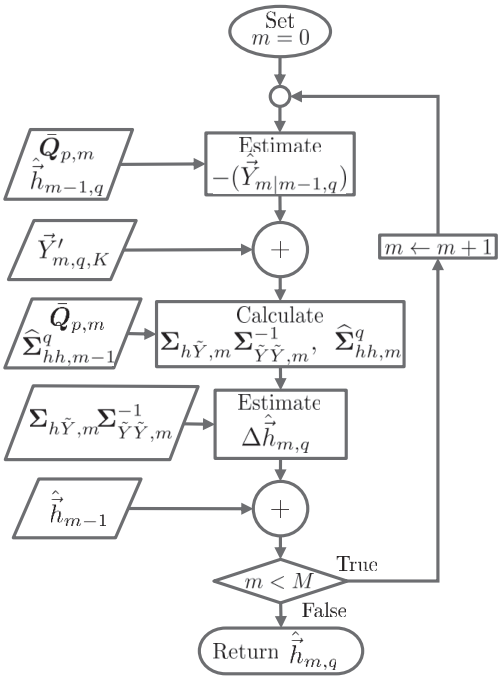


Fig. 3.23: Block diagram of CP-SLMMSE of $\vec{h}_q$ when noise is dominant. Note that the CP-SLMMSE reuses the $\hat{\vec{h}}_{M-1,q}$ and $\widehat{\boldsymbol{\Sigma}}^q_{hh,M-1}$ results from SLMMSE procedure in Fig. 3.21

matrix within the green rectangle has a regular block-circulant structure, i.e. the columns with their index larger than K are $M-1$ times circularly shifted with respect to the first K columns. On the other hand, the matrix within the red rectangle that is equivalent to $\boldsymbol{S}_{\mathrm{CP}}$ has some irregular structure at the top-right corner of the matrix. Such irregular structure leads to off-diagonal elements when block-diagonalization is attempted.

Similar to Sec. 3.4.3, the observation matrix $\bar{\boldsymbol{Q}}_{p,m}$ in MIMO scenario becomes

$$\bar{\boldsymbol{Q}}_{p,m} = \sqrt{N''}\boldsymbol{Z}_m\boldsymbol{J}'\boldsymbol{F}^H_{N''}\boldsymbol{X}_{p,\mathfrak{I}N''}\bar{\boldsymbol{F}}_{N'',L} \in \mathbb{C}^{K\times\mathfrak{I}L}, \tag{3.97}$$

with which, we can reuse the SLMMSE estimation to improve the channel estimation quality. Here, taking the SLMMSE procedure of Fig. 3.21, we replace $\tilde{\boldsymbol{X}}_{p,m}$ by $\bar{\boldsymbol{Q}}_{p,m}$ and $\vec{Y}_{m,q,K}$ by $\vec{Y}'_{m,q,K} = \boldsymbol{Z}_m(\vec{y}_{q,N'}[n])^T_{n=0:N-1}$ whereas, instead of estimating an initial guess of the channel, we reuse $\hat{\vec{h}}_{M-1,q}$ and $\widehat{\boldsymbol{\Sigma}}^q_{hh,M-1}$ from the final SLMMSE estimation obtained in (3.92) for CP-less GFDM. Thus, for CP-SLMMSE, we run the SLMMSE estimation of CP-less GFDM as the initial step, and afterwards, via the block diagram of Fig. 3.23, we update the estimations using the $\vec{Y}'_{m,q,K}$ sequences.

3.4.5 Joint Channel Estimation and Equalization via LMMSE Parallel Interference Cancellation

If the signal reception is in a relatively high SNR region, it is more likely that the interference from transmitted data symbols causes an error floor in channel estimation.

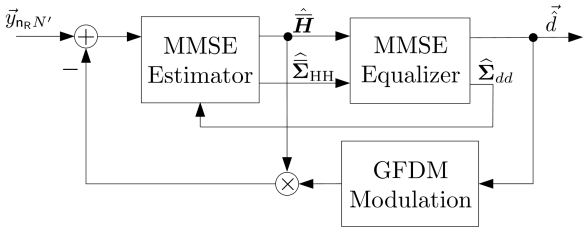


Fig. 3.24: CP-aided channel estimation combined with parallel interference cancellation.

Thus, with the knowledge of equalized payload $\vec{\hat{d}}$ one can cancel such interference on all subcarriers in parallel. In this section, we describe a two step iterative approach to first equalize the payload sequence and then use them for interference cancellation of channel estimation in an LMMSE fashion. Fig. 3.24 depicts the overall concept of CP-aided channel estimation which takes advantage of PIC.

Equalization with Imperfect Channel Estimation. Consider we are equalizing the payload in frequency domain. Thus, the observation signal after removing the known terms is given by

$$\begin{aligned}\vec{Y}'_{\mathfrak{Q}N} &= \vec{Y}_{\mathfrak{Q}N} - \hat{\overline{\boldsymbol{H}}}\bar{\boldsymbol{F}}\overline{\boldsymbol{\mathcal{M}}}\vec{d}_p \\ &= \hat{\overline{\boldsymbol{H}}}\bar{\boldsymbol{F}}\overline{\boldsymbol{\mathcal{M}}}\vec{d}_d + \tilde{\overline{\boldsymbol{H}}}\bar{\boldsymbol{F}}\overline{\boldsymbol{\mathcal{M}}}\vec{d}_d + \tilde{\overline{\boldsymbol{H}}}\bar{\boldsymbol{F}}\overline{\boldsymbol{\mathcal{M}}}\vec{d}_p + \vec{W}_{\mathfrak{Q}N},\end{aligned} \tag{3.98}$$

where $\hat{\overline{\boldsymbol{H}}} \in \mathbb{C}^{\mathfrak{Q}N\times\mathfrak{I}N}$ is the $\mathfrak{Q}\times\mathfrak{I}$ matrix of estimated channels $\hat{\boldsymbol{H}}_{qi} \in \mathbb{C}^{N\times N}$ in frequency domain, where each $\hat{\boldsymbol{H}}_{qi} = \mathrm{diag}(\sqrt{N}\boldsymbol{F}_{N,L}\vec{\hat{h}}_{i,q})$. Here, $\vec{\hat{h}}_{i,q}$ shall be obtained via any of the proposed estimations, e.g. equations (3.78), (3.92) or CP-SLMMSE of Sec. 3.4.4 via observation matrix $\bar{\boldsymbol{Q}}_{p,m}$ defined in (3.97). The matrix $\tilde{\overline{\boldsymbol{H}}} = \overline{\boldsymbol{H}} - \hat{\overline{\boldsymbol{H}}}$ is the error of channel estimation with covariance matrix $\hat{\overline{\boldsymbol{\Sigma}}}_{\mathrm{HH}}$ given in (3.81). The component-wise conditionally unbiased (CWCU) LMMSE [HL14] equalizer that solves (3.98) for $\vec{d}_d$ is given by

$$\vec{\hat{d}} = \frac{\boldsymbol{\Sigma}_{dY}(\boldsymbol{\Sigma}_{d_d\hat{H}} + \boldsymbol{\Sigma}_{d_d\tilde{H}} + \boldsymbol{\Sigma}_{d_p\tilde{H}} + \sigma_w^2\boldsymbol{I}_{\mathfrak{Q}N})^{-1}\vec{Y}'_{\mathfrak{Q}N}}{\mathrm{diag}^{-1}(\tilde{\boldsymbol{\Sigma}}_{dd})}, \tag{3.99}$$

$$\hat{\boldsymbol{\Sigma}}_{dd} = \boldsymbol{\Sigma}_{dd} - \tilde{\boldsymbol{\Sigma}}_{dd}, \tag{3.100}$$

$$\tilde{\boldsymbol{\Sigma}}_{dd} = \boldsymbol{\Sigma}_{dY}(\boldsymbol{\Sigma}_{d_d\hat{H}} + \boldsymbol{\Sigma}_{d_d\tilde{H}} + \boldsymbol{\Sigma}_{d_p\tilde{H}} + \sigma_w^2\boldsymbol{I}_{\mathfrak{Q}N})^{-1}\boldsymbol{\Sigma}_{dY}^H, \tag{3.101}$$

where we calculate the individual matrices as:

$$\boldsymbol{\Sigma}_{d_d\hat{H}} = \hat{\overline{\boldsymbol{H}}}\bar{\boldsymbol{F}}\overline{\boldsymbol{\mathcal{M}}}\overline{\boldsymbol{\Sigma}}_{dd}\overline{\boldsymbol{\mathcal{M}}}^H\bar{\boldsymbol{F}}^H\hat{\overline{\boldsymbol{H}}}^H, \tag{3.101a}$$

$$\boldsymbol{\Sigma}_{d_d\tilde{H}} = \boldsymbol{I}_{\mathfrak{Q}} \otimes \sum_{i=1}^{\mathfrak{I}} \left(\boldsymbol{F}_N\boldsymbol{\mathcal{M}}\boldsymbol{\Sigma}_{dd,i}\boldsymbol{\mathcal{M}}^H\boldsymbol{F}_N^H \circ \hat{\boldsymbol{\Sigma}}^i_{\mathrm{HH},q}\right), \tag{3.101b}$$

$$\boldsymbol{\Sigma}_{d_p\tilde{H}} = \boldsymbol{I}_{\mathfrak{Q}} \otimes \left(\boldsymbol{X}_p\hat{\boldsymbol{\Sigma}}_{\mathrm{HH},q=1}\boldsymbol{X}_p^H\right), \tag{3.101c}$$

$$\boldsymbol{\Sigma}_{dY} = \overline{\boldsymbol{\Sigma}}_{dd}\overline{\boldsymbol{\mathcal{M}}}^H\bar{\boldsymbol{F}}^H\hat{\overline{\boldsymbol{H}}}^H. \tag{3.101d}$$

Herein, for simplicity, we assumed PDP is the same between all Tx-Rx antennas, thus, $\boldsymbol{\Sigma}^q_{hh} = \boldsymbol{\Sigma}^1_{hh}$. Further, $\overline{\boldsymbol{\Sigma}}_{dd} = \boldsymbol{\Sigma}_{dd,1} \oplus \cdots \oplus \boldsymbol{\Sigma}_{dd,\mathfrak{I}}$ and each $\boldsymbol{\Sigma}_{dd,i}$ is the covariance matrix of

transmit data from Tx antenna i. Assuming the Tx data are i.i.d. the matrix $\mathbf{\Sigma}_{dd,i}$ becomes diagonal, with entries of $\text{diag}^{-1}(\mathbf{\Sigma}_{dd,i})$ being 1 at data indexes and 0 at pilot indexes of $\vec{d}_i$. Moreover, with proper permutations, the matrices $\mathbf{\Sigma}_{d_d\hat{H}}$ and $\mathbf{\Sigma}_{d_d\tilde{H}}$ have either one of the following properties:

- Banded with periodic boundary conditions in case of full subcarrier allocation.
- Purely banded matrices in case of guard band insertion.

The lower and upper bandwidth of such matrices is given by $B_l = B_u = \mathfrak{I}(M+N_\alpha)$ where $N_\alpha < M$ is the number of overlapping frequency bins with neighboring subcarriers and it depends on the filter roll-off factor α. The individual matrix blocks of $\mathbf{\Sigma}_{d_p\tilde{H}}$ are the covariance matrices between the pilots and the error of channel estimation and they are non-sparse matrices. Using $\mathbf{\Sigma}_{d_p\tilde{H}}$ for the equalization may provide some marginal gain, but at a cost of intensive matrix inversion complexity. Therefore, for an efficient calculation of (3.99), we neglect the covariance $\mathbf{\Sigma}_{d_p\tilde{H}}$ by setting it to zero.

MMSE Interference Cancellation for the CP-aided CE. Given the knowledge of the equalized data subsymbols $\vec{\hat{d}}$, one can cancel out the CE interference from the observed signal $\vec{y}_{\mathfrak{Q}N'}$. Thus, we have

$$\vec{y}'_{\mathfrak{Q}N'} = \vec{y}_{\mathfrak{Q}N'} - \hat{\boldsymbol{H}}_{\text{eff}}\vec{\hat{d}}_{ab}, \tag{3.102}$$

with $\hat{\boldsymbol{H}}_{\text{eff}} = (\boldsymbol{I}_{\mathfrak{Q}} \otimes \boldsymbol{J}\boldsymbol{F}^H_{N''})\hat{\overline{\boldsymbol{H}}}_{N''}\bar{\boldsymbol{F}}_{N''}\overline{\boldsymbol{\mathcal{M}}}_{N''}$, $\vec{\hat{d}}_{ab} = [\vec{\hat{d}}^T_{a,1} \;\; \vec{\hat{d}}^T_{b,1} \;\; \cdots \;\; \vec{\hat{d}}^T_{a,\mathfrak{I}} \;\; \vec{\hat{d}}^T_{b,\mathfrak{I}}]^T$, $\overline{\boldsymbol{\mathcal{M}}}_{N''} = \boldsymbol{I}_{\mathfrak{I}} \otimes \boldsymbol{\mathcal{M}}_{N''}$, and $\hat{\overline{\boldsymbol{H}}}_{N''}$ being $\mathfrak{Q}\times\mathfrak{I}$ matrix of $\hat{\boldsymbol{H}}_{N'',qi} = \text{diag}(\sqrt{N''}\boldsymbol{F}_{N'',L}\vec{\hat{h}}_{i,q})$. $\vec{\hat{d}}_{a,i}$ and $\vec{\hat{d}}_{b,i}$ are the equalized data associated to blocks a and b, respectively. Expression (3.102) can be decomposed into

$$\vec{y}'_{\mathfrak{Q}N'} = \boldsymbol{H}_{\text{eff}}\vec{d}_{p,ab} + \hat{\boldsymbol{H}}_{\text{eff}}\vec{\tilde{d}}_{ab} + \tilde{\boldsymbol{H}}_{\text{eff}}\vec{d}_{d,ab} + \vec{w}_{\mathfrak{Q}N'} \tag{3.103}$$

where $\vec{\tilde{d}}_{ab} = \vec{d}_{d,ab} - \vec{\hat{d}}_{ab}$ is equalization error of the two blocks a and b with covariance matrix $\hat{\mathbf{\Sigma}}_{dd,ab}$ and $\tilde{\boldsymbol{H}}_{\text{eff}} = (\boldsymbol{I}_{\mathfrak{Q}} \otimes \boldsymbol{J}\boldsymbol{F}^H_{N''})\tilde{\overline{\boldsymbol{H}}}_{N''}\bar{\boldsymbol{F}}_{N''}\overline{\boldsymbol{\mathcal{M}}}_{N''}$. For estimating the channel from observation $\vec{y}'_{\mathfrak{Q}N'}$, we reuse (3.78) by only updating the noise-plus-interference covariance matrix $\mathbf{\Sigma}^{\boldsymbol{J}}_{EE} = \mathbf{\Sigma}^{\boldsymbol{J}}_{EE,\hat{H}\tilde{d}} + \mathbf{\Sigma}^{\boldsymbol{J}}_{EE,\tilde{H}d}$. Hence, the covariance for the second and third terms of (3.103) must be calculated as

$$\mathbf{\Sigma}^{\boldsymbol{J}}_{EE,\hat{H}\tilde{d}} = \mathbb{E}_{\vec{\tilde{d}}_{\text{ab}}}\left[\hat{\boldsymbol{H}}_{\text{eff}}\vec{\tilde{d}}_{ab}\vec{\tilde{d}}^H_{ab}\hat{\boldsymbol{H}}^H_{\text{eff}}\right] = \hat{\boldsymbol{H}}_{\text{eff}}\hat{\mathbf{\Sigma}}_{dd,ab}\hat{\boldsymbol{H}}^H_{\text{eff}}, \tag{3.104}$$

$$\begin{aligned}\mathbf{\Sigma}^{\boldsymbol{J}}_{EE,\tilde{H}d} = (\boldsymbol{I}_{\mathfrak{Q}} \otimes \boldsymbol{J}\boldsymbol{F}^H_{N''})\mathbb{E}_H\Big[&\tilde{\overline{\boldsymbol{H}}}_{N''}\bar{\boldsymbol{F}}_{N''}\overline{\boldsymbol{\mathcal{M}}}_{N''} \\ &\mathbb{E}_{\vec{d}_d}[\vec{d}_{d,ab}\vec{d}_{d,ab}|\tilde{\overline{\boldsymbol{H}}}_{N''}]\overline{\boldsymbol{\mathcal{M}}}^H_{N''}\bar{\boldsymbol{F}}^H_{N''}\tilde{\overline{\boldsymbol{H}}}^H_{N''}\Big]\end{aligned} \tag{3.105}$$

$$\begin{aligned}&(\boldsymbol{I}_{\mathfrak{Q}} \otimes \boldsymbol{F}_{N''}\boldsymbol{J}^H) \\ = \boldsymbol{I}_{\mathfrak{Q}} \otimes &\left(\boldsymbol{J}\boldsymbol{F}^H_{N''}\left(\sum_{i=1}^{\mathfrak{I}}\mathbf{\Sigma}^i_{\text{XX},N''} \circ \hat{\mathbf{\Sigma}}^i_{\text{HH},q}\right)\boldsymbol{F}_{N''}\boldsymbol{J}^H\right),\end{aligned}$$

where $\widehat{\boldsymbol{\Sigma}}_{dd,ab}$ is given by

$$\widehat{\boldsymbol{\Sigma}}_{dd,ab} = \operatorname{diag}\Big([\operatorname{diag}^{-1}(\widehat{\boldsymbol{\Sigma}}_{dd,a,1})^T \ \operatorname{diag}^{-1}(\widehat{\boldsymbol{\Sigma}}_{dd,b,1})^T \\ \cdots \ \operatorname{diag}^{-1}(\widehat{\boldsymbol{\Sigma}}_{dd,a,\mathfrak{I}})^T \ \operatorname{diag}^{-1}(\widehat{\boldsymbol{\Sigma}}_{dd,b,\mathfrak{I}})^T]^T\Big), \tag{3.106}$$

herein, $\widehat{\boldsymbol{\Sigma}}_{dd,a,i}$ and $\widehat{\boldsymbol{\Sigma}}_{dd,b,i}$ are obtained separately via (3.101) for blocks a and b, respectively. Note that in (3.106), since we have assumed that the transmit data symbols in $\vec{d}_d$ are uncorrelated, we would only consider the diagonal elements of $\widehat{\boldsymbol{\Sigma}}_{dd,a,i}$ and $\widehat{\boldsymbol{\Sigma}}_{dd,b,i}$, ignoring the off-diagonal elements. Also note that, should the calculation of interference covariance matrix be ignored for a lower complexity version (as in Sec. 3.4.4), we directly use (3.78) with $\boldsymbol{\Sigma}^{\boldsymbol{J}}_{EE,q} = \sigma_w^2 \boldsymbol{I}_N$.

Interference Cancellation for Sequential Estimators. In Sec. 3.4.4, we described the SLMMSE and CP-SLMMSE approaches, in which, we ignored the interference terms by setting their covariance to zero. We have also mentioned that the approaches are effective only if noise is dominant with respect to the interference level. On the other hand, if the Rx signal power is sufficiently large, the noise level is no longer dominant with respect to the interference level. At such high SNR regions, the SLMMSE and CP-SLMMSE estimators—with zero interference covariance matrix—expect a clear observation signal and based on that, they aim on a reliable estimation. However, the estimation might become worse with increase of SNR, because of the interference not being considered in the MMSE-based estimators. In this case, applying the PIC directly to the sequential estimators might not even be thoughtful, because a wrong estimation would lead to wrong interference cancellation. Therefore, if the Rx signal power is sufficiently large with respect to the noise power, we apply the sequential estimations in a more conservative manner by setting the input SNR of the estimators to a lower value with respect to the true SNR value. The interference cancellation should also be done in a conservative way by taking into account that, on every iteration of PIC the interference level becomes less dominant. Thus, we are allowed to increase the input SNR of the estimators until the point that the final PIC iteration applies the SLMMSE (or CP-SLMMSE) using either a value close to the true SNR or the true SNR value itself as the input SNR of the estimator. The input SNR of the estimators γ_{in} (in dB) on every iteration n_{PIC} of the PIC is calculated as

$$\gamma_{\text{in}} = \gamma_0 + (n_{\text{PIC}} - 1)\frac{\gamma - \gamma_0}{N_{\text{PIC}} + \zeta}, \tag{3.107}$$

where $\sigma^2_{w,\text{in}} = 10^{-0.1\gamma_{\text{in}}}\mathcal{E}_s^2$ is to be used instead of σ_w^2 in (3.85), (3.86), (3.91) as well as in equalization (3.99). γ and $\mathcal{E}_s$ are the true SNR and the signal energy, respectively. γ_0 is the starting SNR value. $\zeta \geq -1$ is an integer determining how conservative the SLMMSE and CP-SLMMSE estimators should perform the estimations. For instance, setting $\zeta = -1$, the true SNR γ would be used on the final PIC iteration while setting ζ to $+\infty$ keeps $\gamma_{\text{in}} = \gamma_0$ for all iterations.

Table 3.2: Complexity growth of the given algorithms

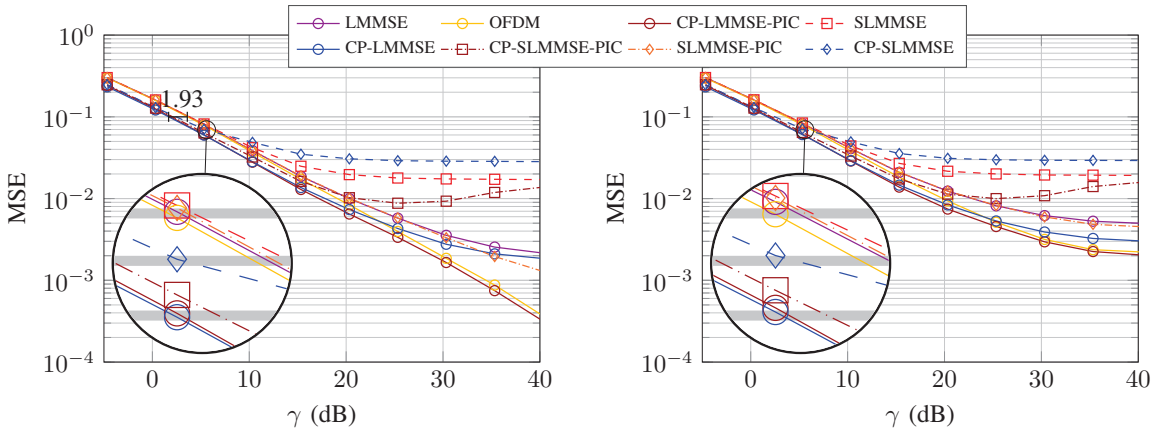
Algorithm	Complexity growth for GFDM with $N_{cp} = K$ samples
CP-LMMSE	$\mathcal{O}(\mathsf{n_R n_T} M^2K^2) + \mathcal{O}(\mathsf{n_R}M^3K^3)$
SLMMSE	$\mathcal{O}(\mathsf{n_R}MK\log_2 M) + \mathcal{O}(\mathsf{n_R n_T}^2K) + \mathcal{O}(\mathsf{n_T}^3) + \mathcal{O}(\mathsf{n_T}MK^2) + \mathcal{O}(MK^3)$
CP-SLMMSE	$2\left(\mathcal{O}(\mathsf{n_R}MK\log_2 M) + \mathcal{O}(\mathsf{n_R n_T}^2K) + \mathcal{O}(\mathsf{n_T}^3) + \mathcal{O}(\mathsf{n_T}MK^2) + \mathcal{O}(MK^3)\right)$
CP-LMMSE-PIC	$N_{\mathrm{PIC}}\big(\mathcal{O}(\mathsf{n_R}^2\mathsf{n_T}M^2K^2) + \mathcal{O}(\mathsf{n_R n_T}^2M^3K^2) + \mathcal{O}(\mathsf{n_T}MK\log_2 K)$ $+\mathcal{O}(\mathsf{n_T}MK\log_2(MK)) + \mathcal{O}(\mathsf{n_R}MK\log_2(MK)) + \mathcal{O}(\mathsf{n_T}M^2K^2) + \mathcal{O}(M^3K^3)\big)$
SLMMSE-PIC	$N_{\mathrm{PIC}}\big(\mathcal{O}(\mathsf{n_R}^2\mathsf{n_T}M^2K^2) + \mathcal{O}(\mathsf{n_R n_T}^2M^3K^2) + \mathcal{O}(\mathsf{n_T}MK\log_2 K)$ $+\mathcal{O}(\mathsf{n_T}MK\log_2(MK)) + \mathcal{O}(\mathsf{n_R}MK\log_2(MK)) + \mathcal{O}(MK\log_2 M) + \mathcal{O}(\mathsf{n_T}^3) + \mathcal{O}(MK^3)\big)$
CP-SLMMSE-PIC	$N_{\mathrm{PIC}}\Big(\mathcal{O}(\mathsf{n_R}^2\mathsf{n_T}M^2K^2) + \mathcal{O}(\mathsf{n_R n_T}^2M^3K^2) + \mathcal{O}(\mathsf{n_T}MK\log_2 K)$ $+\mathcal{O}(\mathsf{n_T}MK\log_2(MK)) + \mathcal{O}(\mathsf{n_R}MK\log_2(MK)) + 2\big(\mathcal{O}(MK\log_2 M) + \mathcal{O}(\mathsf{n_T}^3) + \mathcal{O}(MK^3)\big)\Big)$

3.4.6 Complexity Analysis

The detailed complexity analysis of the algorithms proposed in this section are provided in Appendix A.5. Summarizing such computational efforts in terms of big $\mathcal{O}$ notation, Table 3.2 compares the complexity of presented algorithms for GFDM with $N_{cp} = K$ and $B_l = B_u = \mathfrak{I}(M + N_\alpha)$. One can see that the computational growth of CP-LMMSE, i.e. (3.78) that was cubic in M has been reduced to linear-logarithmic in SLMMSE, although, it would still remain cubic in K because the pilots' energy for all M sequences distributes on all bins in DZT domain. Moreover, since the number of subcarriers K is always larger than the number of estimation parameters $L\mathfrak{I}$, the computational growth $\mathcal{O}(\mathfrak{I}^3)$ that comes from (3.86) would be dominated by $\mathcal{O}(K^3)$ in (3.89). In CP-SLMMSE, despite its computational burden is double of SLMMSE, its growth remains exactly the same as in SLMMSE. In a more advanced receiver that is equipped with parallel interference cancellation, complexity increases due to N_{PIC} times equalization and removal of estimated interference from the observed signal. Here, the overall computational growth reduces in SLMMSE-PIC and CP-SLMMSE-PIC with respect to CP-LMMSE-PIC. But due to the cubic growth of equalization in M, the overall complexity also remains cubic in M. Note that for other multi-carriers the complexity of PIC has to be individually investigated as different multi-carriers might have different equalization and modulation computational efforts.

3.4.7 Simulation Results

In this section, we compare the performance of the proposed algorithms via Monte-Carlo simulations while we consider the performance of OFDM channel estimation as a benchmark. We adopt a spatial multiplexing 2×2 MIMO transceiver that transmits the signal over a Rayleigh fading multipath channel with exponential PDP $\vec{P} = \frac{P_\ell}{\sum_{\ell=0}^{L-1} P_\ell}$ where, $P_\ell = [10^{P_\ell^{\mathrm{dB}}/10}]_{\ell=0,\cdots,L-1}^{T}$ and P_ℓ^{dB} decreases linearly from 0 to -20 dB. We assume 9 taps of the channel which corresponds to $4.7\mu s$ length of impulse response for sampling frequency of 1.92 MHz. The SNR per bit $\mathcal{E}_b/N_0$ which includes the gain of modulation and coding is obtained via $\mathcal{E}_b/N_0 = \gamma - 10\log_{10}\mu r$ where, μ denotes the modulation order



(a) Block fading

(b) $f_d = 70$ Hz

Fig. 3.25: Channel estimation performance of 2×2 MIMO GFDM in block-fading and time-varying conditions.

and r denotes the code rate. For GFDM modulation, we transmit two GFDM blocks of $M = 7$ subsymbols over $K = 48$ subcarriers in which only $K_{\text{on}} = 37$ subcarriers transmit data while $K_{\text{off}} = 11$ subcarriers are considered as guard-bands. The first GFDM block transmits only useful information while the second block multiplexes 2 pilots (first and the last subsymbols) at every second subcarrier i.e. pilot subcarrier spacing of $\Delta k = 2$. This yields $N_p = 2\lceil \frac{K_{\text{on}}}{\Delta k} \rceil$ pilots in the second block and thus, 7.3% pilot overhead over the two GFDM blocks. The pilots are generated as random complex values with unit power. The CP length is set to $N_{\text{CP}} = K$ samples. The prototype filter of GFDM is chosen to be a RC filter with roll-off factor 0.3. With the above given pilot pattern, $\frac{N_{\text{CP}}}{N} = 0.14$, and the RC 0.3 prototype filter, we would have the parameter $10\log_{10}(\kappa_{\text{GFDM}}) = 2.16$ dB. For comparison purpose, we consider an OFDM symbol over $K_{\text{OFDM}} = MK$ subcarriers with time duration of one GFDM block i.e. M subsymbols. Assuming the same bandwidth for both systems, OFDM subcarriers become $1/M$ times narrower than GFDM subcarriers, whereas each GFDM subcarrier carries M data symbols. In order to maintain the same pilot overhead of GFDM also for OFDM, we transmit two OFDM symbols with N_p number of pilots being equispaced over the MK_{on} subcarriers of the second symbol. The parameter κ_{OFDM} with $\frac{N_{\text{CP}}}{N} = 0.14$ becomes $10\log_{10}(\kappa_{\text{OFDM}}) = 0.5$ dB. For proper evaluation of detection performance, we encode the information bits through *Parallel Concatenated Convolutional Codes* with octal generator polynomial (1,15/13). At the receiver side we employ CWCU LMMSE equalization as explained in Sec. 4.5. Afterwards, we transform the equalized signal constellations into maximum likelihood (ML) log-likelihoods by taking into account the fact that each element of the received constellations $\hat{\vec{d}} = \vec{d}_d + \vec{w}$ subjects to an effective noise enhancement approximately distributed as $\vec{w} \sim \mathcal{N}_C(\vec{0}_N, \text{diag}(\frac{1}{\text{diag}^{-1}(\widetilde{\mathbf{\Sigma}}_{dd})} - \sigma_w^2 \vec{1}_N))$ [MZF18]. The ML log-likelihoods are then fed into the soft decoder with 8 turbo iterations.

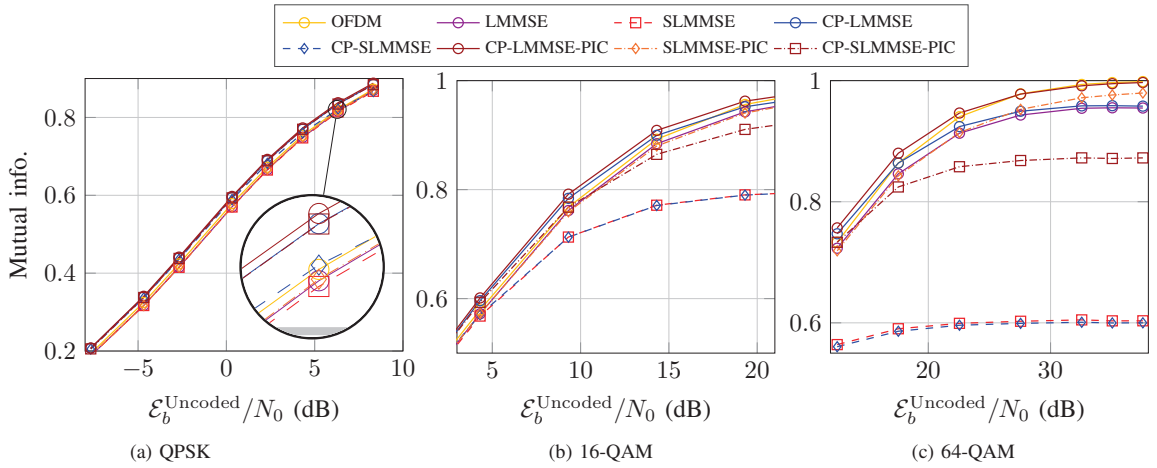
Fig. 3.25 compares the performance of channel estimation in terms of MSE in two scenarios of block fading and time variant channels. In Fig. 3.25a we assumed the channel responses remain constant during transmission of two GFDM blocks a and b. For the SLMMSE and CP-SLMMSE cases, we set $\zeta_{\text{SLMMSE}} = 0$ and $\zeta_{\text{CP-SLMMSE}} = 7$, respectively. One can see that at low SNR regions (e.g. <5 dB) all CP-aided channel estimations outperform the normal LMMSE for both GFDM and OFDM with around 1.93 dB SNR gap. On the other hand, at high SNR values where interference becomes dominant, CP-LMMSE outperforms the normal LMMSE channel estimation of GFDM due to its larger pilot energy, though, the error floor due to the remaining interference does not allow a better performance than OFDM. However, if we afford the extra expense of PIC with $N_{\text{PIC}} = 3$ iterations, the MSE performance becomes entirely smaller than all other cases. Comparing the sequential LMMSE approaches where we ignore the interference covariance matrix, it can be observed that at low SNR regions CP-SLMMSE outperforms OFDM as well as GFDM cases where we ignore the pilot's energy in CP. Though, it is also slightly worse than GFDM's CP-LMMSE which is due to the off-diagonal elements when we enforce block-diagonalization to the first N samples of $y_{N'}$ i.e. (3.96). At high SNR regions, SLMMSE shows a better MSE performance than CP-SLMMSE, because in CP-SLMMSE, the additional interference that leaks from block a into the CP part of block b causes a larger amount of error floor. For both cases, interference cancellation with $N_{\text{PIC}} = 5$ iterations reduces such error floors, although, for SNR values > 25 dB, CP-SLMMSE-PIC starts an upward trend with SNR. Such performance results mean that the conservativeness of the estimation with $\zeta_{\text{CP-SLMMSE}} = 7$ is not sufficient for $\gamma > 25$ dB. In other words, the observed signal is not as reliable as the CP-SLMMSE-PIC estimator expects, and therefore, the MSE results become worse at such SNR regions. This upward trend is not observed for SLMMSE and CP-SLMMSE, because their estimators are operated with $\gamma_{\text{in}} \leq \gamma_0$ as their input SNR.

The performance of channel estimations in a time-variant situation in which the channel impulse responses vary during the transmission of two blocks are depicted in Fig. 3.25b. The channels are simulated according to Jake's model [JBS00] with maximum Doppler shift $f_d = 70$ Hz, and OFDM symbol duration (GFDM block duration) $T = 175\mu$s. In this case, MSE is calculated by considering the mean of all channel realizations i.e.

$$\text{MSE} = \frac{1}{\mathfrak{Q}\mathfrak{I}N} \sum_{q=1}^{\mathfrak{Q}} \sum_{i=1}^{\mathfrak{I}} \sum_{n=0}^{N-1} \left| \sqrt{N} \vec{\omega}_{n,L}^{T} (\vec{h}_{iq}[n] - \hat{\vec{h}}_{iq}) \right|^2 , \tag{3.108}$$

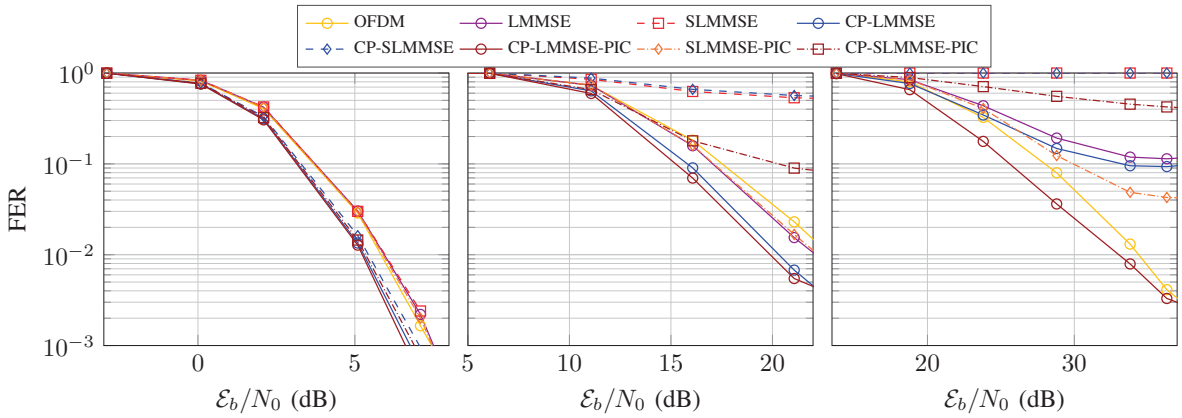
with $\vec{\omega}_{n,L}^{T}$ being the $(n+1)$-th row of $\boldsymbol{F}_{N,L}$, and $\vec{h}_{iq}[n]$ being the channel realization $\vec{h}_{iq}$ defined in (2.10) at time sample n. As can be seen, an error floor adds up to all estimation techniques at high SNR regions, which is due to the additional ISI of time-variant channel.

In order to evaluate the detection performance of the tranceivers with different channel estimation techniques in a block-fading situation, the simulation results of mutual information (MI)—between transmitted bit-stream and received bit log-likelihood-ratios



(a) QPSK (b) 16-QAM (c) 64-QAM

Fig. 3.26: Mutual information for different modulation orders in block fading situation.



(a) QPSK 1/3 (b) 16-QAM 2/3 (c) 64-QAM 3/4

Fig. 3.27: Transceiver performance for different MCS in block fading situation.

(LLR)—and frame error rate (FER) for different MCS are provided in Figures 3.26 and 3.27, respectively. One can observe that for a robust transmission with QPSK 1/3 code rate, a small gap of $\mathcal{E}_b^{\text{Uncoded}}/N_0$ in MI (Fig. 3.26a) leads to approximately 0.5 dB gain of $\mathcal{E}_b/N_0$ at target FER of 10^{-1}. As expected from MSE of channel estimations, all the CP-aided techniques outperform the non-CP-aided cases of the channel estimation, which is due to the additional energy of pilots inside CP as well as dominance of noise over interference in such $\mathcal{E}_b/N_0$ value. Moving forward to the higher MCS, we observe that the $\mathcal{E}_b/N_0$ gain of FER results becomes larger up to nearly 2.4 dB compared to OFDM, e.g. at a target FER of 10^{-1}, CP-LMMSE-PIC outperforms OFDM with 2.32 dB (16-QAM 2/3 Fig. 3.27b) and 2.4 dB (64-QAM 3/4 Fig. 3.27c), in which, the behavior of such FER performances are also being validated through the MI results in Figures 3.26b and 3.26c, respectively. We also note that, for an intermediate MCS e.g. 16-QAM 2/3, CP-LMMSE

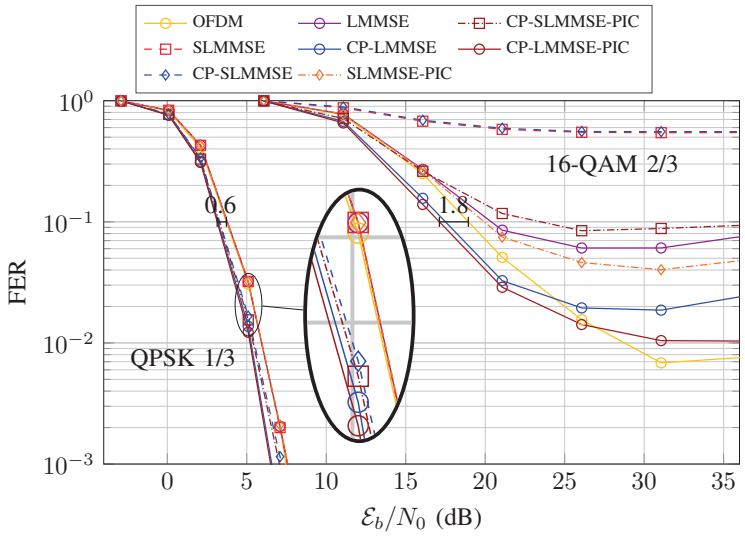
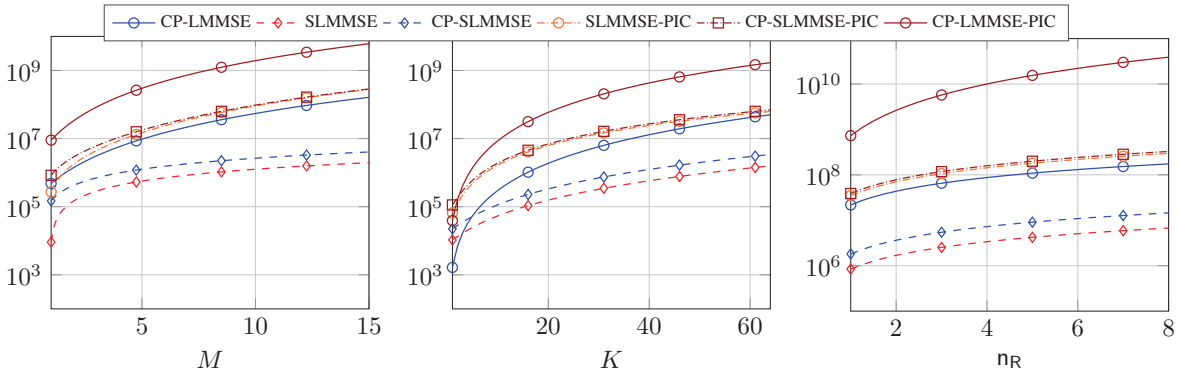


Fig. 3.28: Transceiver performance for QPSK 1/3 and 16-QAM 2/3 in time-variant situation with $f_d = 70$ Hz.

has nearly the same performance of CP-LMMSE-PIC, which makes it possible to avoid further complexity of PIC. However, for a high MCS e.g. 64-QAM 3/4, we suggest not to abandon CP-LMMSE-PIC due to almost entirely dominance of interference over noise levels.

In a time-variant situation where $f_d = 70$ Hz, the information bearing symbols subject to additional ISI which results to increase of error floor in symbol detection. This is illustrated in Fig. 3.28, where, we compare FER performances for QPSK 1/3 and 16-QAM 2/3. As expected, for a robust MCS of QPSK 1/3, the FER degradation in Fig. 3.28 is not significant compared to Fig. 3.27a, in which, the channel was constant. However, for a higher order of MCS, i.e. 16-QAM 2/3, noise is no longer dominant over interference and thus, CP-LMMSE-PIC in Fig. 3.28 degrades by 1.45 dB with respect to its performance in block fading situation, i.e. Fig. 3.27b. For OFDM, the degradation with respect to its block fading scenario is 1.63 dB. We also observe that at very high SNR regions, i.e. $\mathcal{E}_b/N_0 > 30$ dB, FERs of LMMSE, SLMMSE-PIC, CP-LMMSE and OFDM increase with SNR. We justify this behavior by considering the fact that in time-variant channel additional ICI is produced into the detected symbols while the MMSE equalizer as well as decoder do not consider such ICI. Formally, by ignoring the off-diagonal elements of $\widetilde{\mathbf{\Sigma}}_{dd}$ for decoder, the LLR outputs are too confident and thus, the detection performance degrades. In SLMMSE and CP-SLMMSE such behavior does not appear, because the input SNR of the equalizer and decoder is limited to $\gamma_{\text{in}} \leq \gamma_0$. We also note that, the low complexity approaches have larger error floors, because of ignoring interference terms in SLMMSE estimation. In addition, comparing OFDM and GFDM's CP-LMMSE-PIC, when the self-interference of GFDM adds up to the ICI produced by Doppler effects, the FER performance of GFDM slightly degrades with respect to OFDM at SNR values of



(a) FLOPs vs. no. of subsymbols (b) FLOPs vs. no. of subcarriers (c) FLOPs vs. no. of Rx antennas

Fig. 3.29: Complexity of the proposed approaches in terms of number of FLOPs.

$>$ 29 dB, because OFDM only subjects to ICI and it does not suffer from self-interference generation.

The complexity comparison of the proposed approaches in terms of number of FLOPs are depicted in Fig 3.29. The default simulation parameters are chosen as $M = 7$, $K = 48$, $\mathfrak{I} = \mathfrak{Q} = 1$, $L = 9$ and $N_{\text{PIC}} = 3$. Clearly, the SLMMSE estimation which divides the N dimensional signal into M sequences of K dimension, has the lowest complexity in all three figures, i.e. Fig. 3.29a, Fig. 3.29b and Fig. 3.29c, although, in Fig. 3.29b for $K < 3$, CP-LMMSE requires smaller number of FLOPs because the dominant term $2L^3$ in $C_{(3.93)}$ becomes larger than $\frac{(N')^3}{3}$ in $C_{(3.78)}$. Taking advantage of pilots' energy from CP, the complexity of CP-SLMMSE almost doubles with respect to SLMMSE, but they remain almost in the same order of magnitude. If we adopt PIC, we observe that the complexity increases with about 1.5 order of magnitude in SLMMSE and CP-SLMMSE and almost 2 orders of magnitude for CP-LMMSE.

Based on the FER results and complexity comparisons, we summarize that CP-SLMMSE fits best for a robust MCS, i.e. QPSK 1/3, whether with or without mobility. CP-LMMSE should be chosen for an intermediate MCS, i.e. 16-QAM 2/3, whether with or without mobility. CP-LMMSE-PIC must be chosen for high data rate transmissions, i.e. 64-QAM 3/4.

So far, the channel estimation techniques that were proposed in previous sections, were solely relying on block-fading assumptions of the wireless channel. In realistic channel conditions where the users have mobility, the channel transfer function starts to vary during the signal transmission. Under such circumstances, a more accurate channel estimation is needed to combat the ICI posed by the Doppler shifts. In the next section, having in mind that non-orthogonal multi-carriers take advantage of pilot time localization, we derive the *Wiener-Hopf* filters for improved channel estimation in time-varying channel conditions.

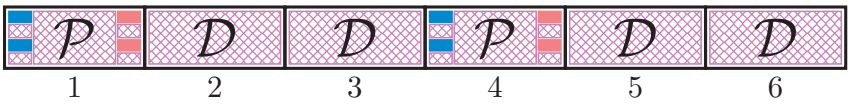


Fig. 3.30: Pilot $\mathcal{P}$ and data $\mathcal{D}$ blocks over $B = 6$ number of GFDM blocks.

3.5 Pilot- and CP-aided Channel Estimation in Time-Varying Scenarios

Consider a digital baseband signal $\vec{x}_b \in \mathbb{C}^N$ associated to GFDM block $b \in \{1, \cdots, B\}$ is linearly modulated by $\boldsymbol{A} \in \mathbb{C}^{N \times N}$, i.e.

$$\vec{x}_b = \boldsymbol{A}\vec{d}_b, \tag{3.109}$$

where $\vec{d}_b = \text{vec}(\boldsymbol{D}_b)$, and $\boldsymbol{D}_b \in \mathbb{C}^{K \times M}$ denotes the allocation matrix for K subcarriers and M subsymbols at block b.

Similar to Sec. 3.4.4, we adopt rectangular grid pilot pattern as in Fig. 3.19. Here, we consider multiple GFDM blocks in which, every Δb block, pilots are multiplexed with data. Fig. 3.30 illustrates an example of the pilot insertion over $B = 6$ blocks. According to Fig. 3.30, the channel should be estimated at $\mathcal{P}$-type GFDM blocks, i.e. $b = \{1, 4\}$, while at the $\mathcal{D}$-type blocks it must be either interpolated, i.e. for $b = \{2, 3\}$, or be predicted, i.e. for $b = \{5, 6\}$. Note that at $\mathcal{P}$-type GFDM blocks, scattered pilots (blue and red) are multiplexed with payloads (magenta) as was depicted in Fig. 3.19, whereas the $\mathcal{D}$-type GFDM blocks carry only payloads. In order to take advantage of pilots' energy from CP, we set the CP length equal to the duration of one subsymbol, i.e. $N_{\text{CP}} = K$ samples, and we use the same CE algorithm proposed in Sec. 3.4.2.

Under the assumption of perfect time and frequency synchronization, the received signal $\vec{y}_b$ associated to block b becomes

$$\vec{y}_b = \boldsymbol{\mathcal{C}}_N \vec{x}_b + \vec{w}_b, \tag{3.110}$$

where $\vec{w}_b \in \mathbb{C}^N$ is the vector of AWGN samples. $\boldsymbol{\mathcal{C}}_N \in \mathbb{C}^{N \times N}$ is the matrix of CIRs with $\vec{h}$ at the beginning of its first column. In a block-fading scenario, the CIR $\vec{h}_n$ remains constant on every column of $\boldsymbol{\mathcal{C}}_N$ (i.e. $\vec{h}_n = \vec{h}_{n'}$), and therefore, $\boldsymbol{\mathcal{C}}_N$ becomes a circulant matrix (due to the CP insertion). In a time-variant situation, every tap of $\vec{h}_n$ varies with n depending on a normalized Doppler frequency given by $\nu_D = f_d/F_b$, where f_d denotes the maximum Doppler shift and F_b denotes the block rate in Hertz (Hz). The temporal correlation of the elements of $\vec{h}_n$ with their corresponding taps on $\vec{h}_{n'}$ becomes

$$r_{n,n'} = \mathbb{E}(h_n[\ell]^* h_{n'}[\ell]), \tag{3.111}$$

which immediately, follows with the channel autocorrelation matrix

$$[\mathbf{R}'_h]_{n,n'} = r_{n,n'} \text{ for } n, n' \in \{0, \cdots, BN' - 1\}. \tag{3.112}$$

Note that here, the linear model (3.110) considers a SISO channel and also our simulations in Sec. 3.5.2 are applied to SISO systems. But since the filtering is a straight-forward approach that considers individual CIRs, it can directly be applied to MIMO systems as well.

3.5.1 Adaptive Filtering based on Wiener-Hopf Approach

In transmission of B blocks with Rx signal model (3.110), the channel not only varies from one block to another, but also within each block. If the frequency-dispersion of the channel is considerable, i.e. $\nu_D > 1\%$, equalizing the $\mathcal{D}$-type blocks with the channel estimations at $\mathcal{P}$-type blocks is not thoughtful. In fact, the channel has to be interpolated for the data blocks between two $\mathcal{P}$-type blocks and be predicted for those which follow the last pilot-scattered block. Furthermore, given multiple channel estimations at different $\mathcal{P}$-type blocks, such estimations can be further improved.

Define the auto-correlation matrix $\mathbf{R}_h \in \mathbb{C}^{B\times B}$ being a sub-matrix of $\mathbf{R}'_h$ with each of its rows/columns being the $n_{p,b}$-th row/column of $\mathbf{R}'_h$. $n_{p,b}$ denotes a sample index within the duration of block b, in which, the channel estimation is expected to have its best performance. As will be further justified in Sec. 3.5.2, $n_{p,b}$ is obtained via

$$n_{p,b} = (b-1)N' + \mathfrak{N}_0 + \frac{\sum_{n=1}^{N_{x_p}} nP_{x_p}[n]}{\sum_{n=1}^{N_{x_p}} P_{x_p}[n]}, \tag{3.113}$$

which is the weighted average of the sample indexes with the pilot's power $P_{x_p}[n]$ as the weighting coefficient. $\mathfrak{N}_0$ denotes the number of initial samples that are not used for CE, N_{x_p} denotes the length of pilots' signal that is used for CE e.g. If the whole block including its CP is used for CE, we have $N_{x_{p,N'}} = N', \mathfrak{N}_0 = 0$ and $\vec{P}_{x_{p,N'}} = |\vec{x}_{b,p,N'}|^2$, whereas for a CE with CP-less pilots' signal, we have $N_{x_{p,N}} = N, \mathfrak{N}_0 = N_{\text{CP}}$ and $\vec{P}_{x_{p,N}} = |\vec{x}_{b,p,N}|^2$. $\vec{x}_{b,p,N'}$ and $\vec{x}_{b,p,N}$ denote the last N' and the last N elements of $\vec{x}_{b,p,N''}$, respectively.

Here, we use the *Wiener-Hopf* approach in order to derive MMSE based adaptive filters for joint *smoothing-interpolation-prediction* of the channel. Thus, the ℓ-th channel tap at b-th block based on Wiener-Hopf approach [Kay93] follows:

$$\hat{h}_{\ell,b} = [\vec{R}_h]^H_{\mathcal{P},b}(\mathbf{R}_h^{(\mathcal{P})} + \hat{\sigma}^2_\ell \boldsymbol{I}_{B_p})^{-1}\vec{\hat{h}}_\ell, \tag{3.114}$$

where $[\vec{R}_h]_{\mathcal{P},b} \subset \mathbf{R}_h$ is the b-th column of $\mathbf{R}_h$ with row indexes corresponding to $\mathcal{P}$-type block indexes. The sub-matrix $\mathbf{R}_h^{(\mathcal{P})} \subseteq \mathbf{R}_h$ contains the rows and columns of $\mathbf{R}_h$ associated to $\mathcal{P}$-type blocks. $\hat{\sigma}^2_\ell$ denotes the ℓ-th diagonal element of $\widehat{\boldsymbol{\Sigma}}_{hh}$ defined in (3.73). B_p denotes the number of pilot blocks. $\vec{\hat{h}}_\ell \in \mathbb{C}^{B_p}$ is the vector of channel estimations at tap ℓ and at $\mathcal{P}$-type blocks. Note that the channel estimation in (3.72) is L-dimensional and corresponds to the sample time index $n_{p,b}$, whereas $\vec{\hat{h}}_\ell$ contains the ℓ-th element of $\hat{\vec{h}}_{b,\text{CPLMMSE}}$ defined in (3.72) for multiple estimations at $\mathcal{P}$-type blocks.

In the above filtering approach, knowledge of PDP $\vec{P}$ as well as maximum Doppler shift f_d (and correspondingly the temporal auto-correlation matrix $\mathbf{R}'_h$) is required. Moreover, in a stationary channel condition where the CIR taps are fully correlated in time, i.e. $\mathbf{R}'_h = \mathbf{1}_{BN'}$, the expression (3.114) acts similar to taking the average of all estimations[5], whereas in an extreme channel condition with $f_d \to \infty$, we would have $\mathbf{R}'_h = \boldsymbol{I}_{BN'}$ and thus, each estimation is treated independently.

3.5.2 Simulation Results

In this section, we examine the performance of the proposed CP-aided CE and the adaptive filtering for time varying channels via Monte-Carlo simulations. Consider a Rayleigh fading multipath channel with exponential PDP where, $P_\ell = [10^{P_\ell^{\mathrm{dB}}/10}]^T_{\ell=0,\cdots,L-1}$ and P_ℓ^{dB} decreases linearly from 0 to -20 dB. We consider $L = 9$ taps of the channel which corresponds to 4.7µs length of impulse response for sampling frequency of 1.92 MHz. Each tap of the impulse response $\vec{h}[n]$ is temporally correlated according to [Cla68] $r_{n,n'} = J_0(2\pi(n-n')\nu_D)$, where $J_0(\cdot)$ is the zeroth order Bessel function of the first kind. The transmit information bits are encoded through *Parallel Concatenated Convolutional Codes* with octal generator polynomial (1, 15/13) and code-rate r. The number of transmission blocks is set to $B = 15$, and from $b = 1$, every $\Delta b = 3$-rd block, we multiplex pilots with the data symbols. For GFDM, we consider blocks $M = 7$ subsymbols are transmitted over $K = 48$ subcarriers. Pilots in form of first root Zadoff-Chu sequence are inserted according to Fig. 3.19 with pilot subcarrier spacing $\Delta k = 2$. The above GFDM configuration yields 4.76% pilots overhead. RC with roll-off factor 0.3 has been chosen as the GFDM prototype filter. As a comparison benchmark, we adopt $B_{\mathrm{OFDM}} = 15$ OFDM symbols over $K_{\mathrm{OFDM}} = MK$ subcarriers while one OFDM symbol has the same time duration of a GFDM block. Hence, with the assumption of having identical bandwidth for both systems, the subcarriers for GFDM become M times broader than the OFDM subcarriers. In order to maintain the same pilot overhead also for OFDM, we set $\Delta k_{\mathrm{OFDM}} = 2M$ and $\Delta b_{\mathrm{OFDM}} = 3$. The SNR is denoted by $\mathcal{E}_s/N_0$ and adding the gain of modulation and coding we have $\mathcal{E}_b/N_0 = \mathcal{E}_s/N_0 - 10\log_{10}\mu r$, where μ is the modulation-order. The receivers for both systems adopt Component-Wise Conditionally Unbiased LMMSE estimation [MZF18, Sec. II.B] for joint equalization of the wireless channel and the GFDM/OFDM matrix $\boldsymbol{\mathcal{M}}$. Thereupon, the equalized signal constellations are transformed into ML log-likelihoods and they are then soft-decoded with 8 turbo iterations.

Fig. 3.31 illustrates an example of pilots signal in time domain, i.e. $\vec{x}_{p,N'} = [\boldsymbol{\mathcal{M}}_{\mathrm{CP}}^T \ \boldsymbol{\mathcal{M}}^T]^T \vec{d}_{b,p}$ (the upper figure), and the MSE of the CE $\hat{\vec{H}} = \boldsymbol{F}_{N,L}\hat{\vec{h}}_b$ with respect to each channel realization $\vec{H}_n$ at sample time instant n and $\mathcal{E}_s/N_0 = 30$ dB (the lower figure). One can see that the minimum error of the channel estimation—with

[5] Note that here, the power of each tap ℓ and thus, its estimation accuracy is also taken into account via $\hat{\sigma}^2_\ell$.

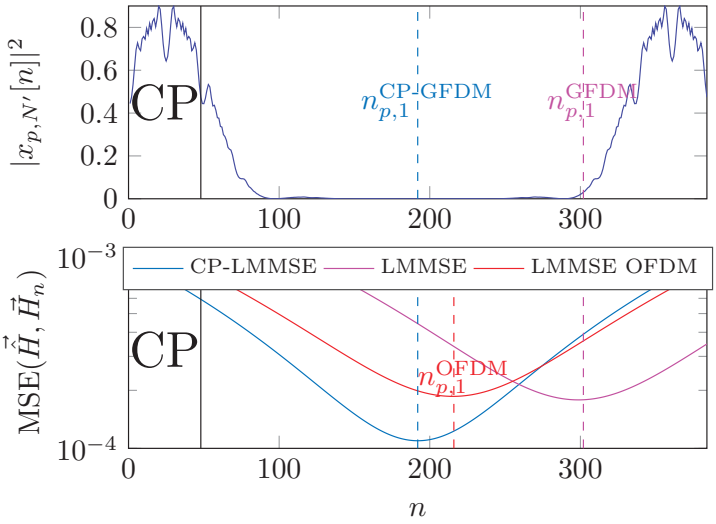
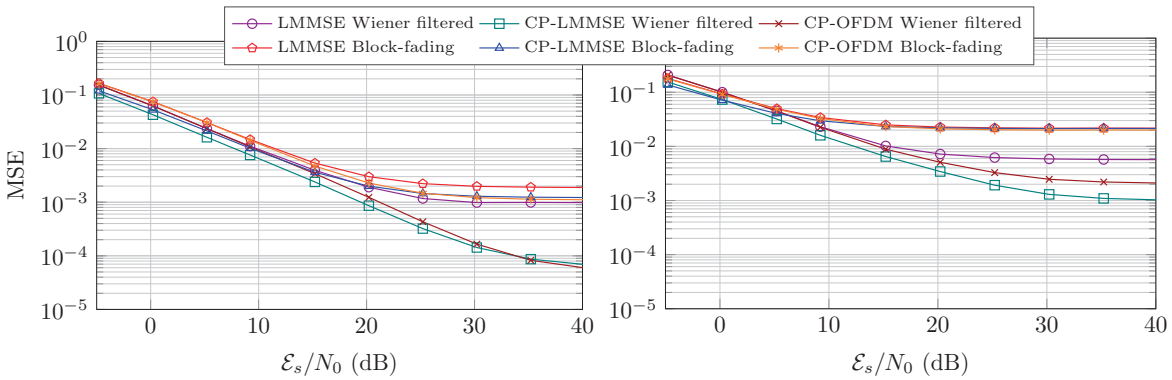


Fig. 3.31: MSE of channel estimation within one block duration at $\mathcal{E}_s/N_0 = 30$ dB and $f_d = 70$ Hz. Dashed lines are the theoretical values of $n_{p,b}$ defined in (3.113).

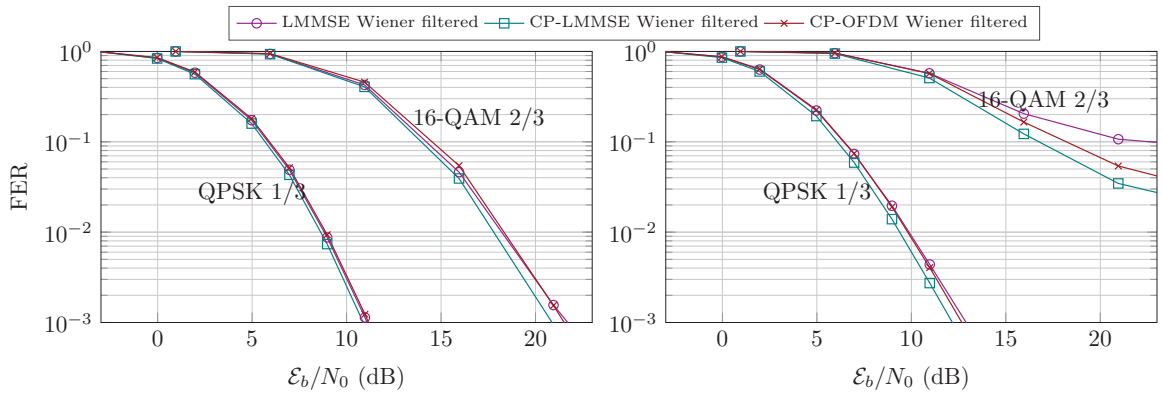


(a) $f_d = 70$ Hz **(b)** $f_d = 300$ Hz

Fig. 3.32: Channel estimation performance over time-varying channels.

block-fading assumption—within the block duration happens at the weighted average of the pilot signal $n_{p,1}^{\text{CP-GFDM}}$ (with CP) and $n_{p,1}^{\text{GFDM}}$ (without CP). For OFDM, the minimum error happens at the center of the N-length CP-less signal because the pilots in OFDM have uniform energy over the whole duration of N samples. Fig. 3.31 confirms the choice of $n_{p,b}$ in (3.113) for defining the autocorrelation matrix $\mathbf{R}_h$ out of $\mathbf{R}'_h$ because the CE has its best performance at sample instance $n_{p,b}$.

Fig. 3.32 compares the performance of channel estimation in terms of MSE for GFDM LMMSE and CP-LMMSE estimations as well as OFDM CP-LMMSE. We consider two time-variant channel scenarios with $f_d = 70$ Hz and $f_d = 300$ Hz. The corresponding normalized Doppler frequency for each scenario becomes $\nu_D = 1.23\%$ and $\nu_D = 5.25\%$, respectively. In addition, the block-fading assumption calculates the MSE of the estimated channel at $\mathcal{P}$-type blocks with respect to the channel realizations within the duration of $\mathcal{P}$-type block itself and also its following two $\mathcal{D}$-type blocks. The Wiener filtering performs



(a) $f_d = 70$ Hz **(b)** $f_d = 300$ Hz

Fig. 3.33: Frame error rate simulation results of GFDM and OFDM over time-varying channels for QPSK 1/3 and 16-QAM 2/3.

channel smoothing at $\mathcal{P}$-type blocks $b^{(\mathcal{P})} = \{1, 4, 7, 10, 13\}$, channel interpolation at $\mathcal{D}$-type blocks $b^{(\mathcal{D})} = \{2, 3, 5, 6, 8, 9, 11, 12\}$ and channel prediction at the last two $\mathcal{D}$-type blocks $b^{(\mathcal{D})} = \{14, 15\}$. From Fig. 3.32a and Fig. 3.32b, it can be observed that the optimum CE performance is achieved via GFDM's CP-LMMSE with Wiener filter, which is due to the fact that the total pilot's energy including the CP in GFDM becomes larger than the one in OFDM as well as CP-less GFDM. Although, at very high SNR regions in Fig. 3.32a, OFDM becomes slightly better than GFDM's CP-LMMSE because of the self interference from data symbols in GFDM. Such behavior does not happen in Fig. 3.32b because the error floor due to the self-interference in GFDM is at a much lower lever compared to the error floor due to the channel variations in OFDM. We also note that using the proposed CP-LMMSE approach does not considerably improve the CE performance in OFDM because the pilot's energy is uniformly distributed over the whole symbol duration and thus an OFDM system lacks of an efficient pilot time-localization. Comparing the simulation curves with block-fading assumption, we observe that at very low SNR regions, CP-LMMSE has smaller MSE with respect to OFDM and CP-less GFDM, while at high SNR regions the error floor due to the channel variations becomes significant for three approaches.

The transceiver performance for two MCS of QPSK 1/3 and 16-QAM 2/3 is provided in Fig. 3.33. Since in channel estimation MSE results, we observed that the approaches with Wiener filtering has outperformed the results with block-fading assumptions, here we compare the performance for Wiener filter approaches. As can be expected from the CE MSE results, GFDM's CP-LMMSE outperforms OFDM and CP-less GFDM in terms of frame error rate (FER) for both MCS as well as both time-variant scenarios of $f_d = 70$ Hz and $f_d = 300$ Hz. Although, the performance improvement is almost negligible for a robust MCS of QPSK 1/3, the SNR gap becomes much larger up to 1.45 dB for a target FER of 10^{-1} in 16-QAM 2/3 and $f_d = 300$ Hz, i.e. Fig. 3.33b. Comparing the CP-less GFDM's LMMSE with OFDM, we observe that albeit they almost have the same

MSE performance for $\mathcal{E}_s/N_0 < 15$ dB, GFDM's LMMSE slightly outperforms OFDM in Fig. 3.33a, which is due to the higher frequency diversity of GFDM in frequency selective situations. In a scenario where the channel is too much time-varying i.e. Fig. 3.33b, OFDM outperforms GFDM's LMMSE because of its smaller error floor in CE MSE results.

3.6 Summary

In this chapter, we initially studied the pilot-aided channel estimation in MIMO GFDM systems where we derived the classical and Bayesian estimators for interference-limited MIMO GFDM systems. The simulation results confirmed the closed form expressions of our theoretical MSE derivations. Moreover, we have also studied the pilot design in GFDM systems in which, we observed that rectangular grid pilot pattern outperforms the block-type pilots. Indeed, rectangular grid pilot pattern better suppresses the error floor—from self-interference—because at pilot subcarriers the ratio of the pilots' power to the data subsymbols' power is much larger. We have also further compared the coded error rates of GFDM and OFDM under imperfect CE and CCM knowledge. In this regard, via extensive simulations, we showed that the BER performance of GFDM via various CE techniques is very similar to OFDM when a robust code-rate is employed. Although, by adopting a high code-rate of e.g. 5/6, the self-interference dominates the CE performance and GFDM MP cases fall roughly 5 dB behind OFDM at high SNR regions.

Since GFDM's low complexity frequency domain processing allows orthogonal pilot insertion, we proposed a precoding technique so that interference-free pilots can be observed at the receiver. Evaluating the signal characteristics, we showed that IFPI GFDM achieves a PAPR and OOB radiation that is very close to basic GFDM, but on the other hand, it also achieves identical CE performance when compared to an OFDM system.

In sections 3.4 and 3.5, we proposed different solutions to exploit a multi-carrier's pilot time localization, and achieve a higher quality CE performance via using the whole transmission block including its CP. There, we derived an iterative approach for joint channel estimation and equalization of MIMO non-orthogonal multi-carriers with imperfect CIR knowledge. Applying the approach to the non-orthogonal GFDM, we showed that low complexity approaches of SLMMSE and CP-SLMMSE achieve almost the same performance of their high complexity variations for robust MCS. However, when high orders of MCS are of interest, ignoring the interference terms leads to significant performance degradation, and thus, it necessitates the algorithms to adopt interference-cancellation. Considering transmission of multiple GFDM blocks, we have also derived the adaptive Wiener-Hopf filters for joint smoothing-interpolation-prediction of the channel. From simulation results, we have observed that using adaptive filters for channel estimation has significantly reduced the error floor, when it is compared to the case where the CE relies solely on block-fading assumptions.

Chapter 4

Design of UW-Based Transmission for MIMO Multi-Carriers

Recently, unique words transmission has attracted attention in the research on single-carrier as well as multi-carrier systems. The idea of transmitting a deterministic sequence instead of a CP was initially proposed by [DGE00] and [WMS+02] for single-carrier systems. The main concept is to exploit the knowledge of the UW sequences at the receiver side to improve synchronization [DGE00], channel estimation [CSBM06], and equalization [WMS+02], while at the same time, maintaining the advantages of CP in terms of robustness against multi-path channels. In CP based OFDM or single carrier frequency domain equalization (SC-FDE) systems, the last part of time domain payload block is copied to its beginning. Having the CP length longer than the channel delay spread, allows the receiver to emulate the wireless channel transfer function as a circular filter. Thus, DFT based channel estimation and equalization techniques, significantly reduce the complexity of the receivers for OFDM, SC-FDE, as well as other multi-carrier systems (e.g. [MMG+14]). Nevertheless, since CP is being generated by random data, it provides very limited advantages in terms of parameter estimation (e.g. CIR, carrier frequency offset (CFO), etc.), and thus it is usually discarded at the receiver and in that sense, its energy is being wasted. Unlike CP that is random, UW sequences are deterministic and besides bringing the possibility to the receiver to emulate circular CIR, they allow per-block synchronization and channel estimation.

Our contributions in this chapter are mainly published in [1], [5] and [4]. Here, in order to enable a robust UW-aided data transmission of MIMO transceivers in broadband mobile channels, the following topics have been addressed:

- By taking into account the UW frame design of [DGE00] as a benchmark, we extend the frame structure to MIMO applications [5]. Considering a doubly-dispersive channel condition, we analyze the overhead requirements of a UW-based system from synchronization and channel estimation perspectives and we compare its efficiency with a pilot-aided CP-OFDM system design.

- Considering a non-orthogonal multi-carrier that not only focuses on robust data transmission, but also controls the transmit signal impairments, we analyze the MIMO UW sequences in terms of OOB radiation for GFDM [1].

- We extend the SoA UW-based frame structure to centralized MIMO systems, and correspondingly, we derive the timing and frequency synchronization metrics [4]. The proposed synchronization techniques would be applicable to any MIMO system (e.g. OFDM, SC, non-orthognal [MMG+14], etc.). Having multiple synchronization metrics from different antenna pairs, we propose a new technique to combine the metrics for a significantly more robust timing synchronization. By deriving the Cramér-Rao lower bound (CRLB) for CFO estimation of MIMO UW sequences, we theoretically show that MIMO UW-based systems have smaller bounds of CFO estimation, in comparison to preamble-and-CP based systems.

- Considering that the *Payload-UW* blocks are being transmitted through a highly frequency selective and time-variant channel, we derive the Wiener-Hopf filters for the corresponding MIMO LMMSE CE. Proposing a frame structure that considers circularly rotating the UW slots of the MIMO antennas, we also analyze its CE performance with respect to the SoA solutions [1].

- Proposing an EQ technique that emulates the circular channel condition over the payload (and thus, employs an FFT-size only over the payload length within the *Payload-UW* block), we derive its *component-wise conditionally unbiased* (CWCU) LMMSE equalizer [HL14] and correspondingly, we analyze the transceiver performance of different proposed techniques with respect to the SoA approaches [1].

4.1 Frame Design, Efficiency and Overhead Analysis

4.1.1 Illustrative Scenario

For clarity, consider an extreme scenario of a spatial multiplexing 2×2 MIMO system, in which, 100 complex data[1] (i.e. 50 per Tx. antenna) must be transmitted by $K = 18$ subcarriers over a frequency selective—with $L = 4$ taps—and fast fading channel. In [CT08], it is recommended that for a CP-OFDM system in a time-variant channel, it is necessary to employ space-frequency pilots pattern, i.e. $\mathfrak{I}L$ pilot insertion into every OFDM symbol. In this case $\mathfrak{I}L = 8$ channel taps must be estimated and therefore, 8 pilot subcarriers must be inserted into every OFDM symbol. From a total number of 18 subcarriers, each OFDM symbol can carry 10 data subcarriers. Hence, because of the two streams being simultaneously transmitted over a 2×2 MIMO channel, five OFDM

[1] Complex data is referred to any complex value from a modulation alphabet, e.g. 2^{μ}-QAM. For instance, one may assume 100 bits via 1/2 code-rate QPSK.

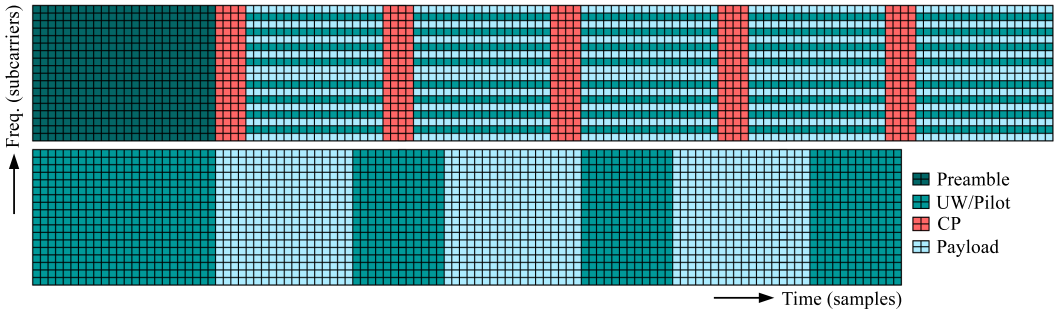


Fig. 4.1: Example of the frame structure in time-frequency resource grid. Here, each resource element is associated to the bandwidth of one subcarrier and the duration of one time sample. Without loss of generality, the impact of guard band insertion has been neglected in this figure. As one may see, the UW-based frame design, saves the time-frequency resources of one OFDM symbol for transmission of 100 complex data over a 2×2 MIMO channel.

symbols would be needed to carry $10 \times 2 \times 5 = 100$ complex symbols over $5(18+4) = 110$ time resources. The frame design for such an OFDM system is depicted in the upper part of Fig. 4.1. Considering that the power is doubled over the two MIMO streams, the energy efficiency of this frame design is $\eta_{\mathcal{E},\text{Pilots}} = \frac{100}{2\times 110} = 45\%$.

Now, consider the same explanatory scenario, in which, the knowledge of UW sequences shall be exploited for channel estimation (and also synchronization), and therefore, no pilot insertion is necessary. In this case, we design the length of the UW sequence to be $N_\text{u} \geq (\mathfrak{I}+1)L$, such that the first L samples shall be ignored (due to the inter-block-interference (IBI) from the previous block), and the rest of $(N_\text{u} - L) \geq \mathfrak{I}L$ samples would be used to estimate $\mathfrak{I}L$ channel parameters. Unlike the approach in [DGE00], which considers the length of *Payload-UW* as the FFT size and thus reduces the payload size N_d, we keep the FFT size for N_d samples without including the UW length in it, and for its corresponding equalization, we will reconstruct a cyclic signal at the receiver side and elaborate further in Sec. 4.5. In addition, at the beginning of a transmission, we consider a double UW sequence for primary synchronization, which is equivalent to the preamble of the CP-OFDM system. For the rest of the blocks, we consider only a single UW sequence. Note that here, if we use double UW sequence for all blocks, the overhead becomes significant, and if we use double half-length UW sequences, the SNR for synchronization becomes half. The frame design of such a UW-based system is depicted in lower part of Fig. 4.1. There, we can see that 12 time samples are considered for each UW sequence, while all the 18 subcarriers of the data symbol are actively used for data transmission. In this case, for three UW-based symbols, $18 \times 2 \times 3 = 108$ complex symbols are transmitted over $3(18+12) = 90$ time resources. Therefore, the energy efficiency of the UW-based frame design becomes $\eta_{\mathcal{E},\text{UW}} = \frac{108}{2\times 90} = 60\%$, and thus the time resource of one OFDM symbol is being saved. The detailed efficiency analysis of such frame design is further discussed in the following.

4.1.2 CP vs. UW Efficiency Analysis

Consider that for primary synchronization, a preamble consisting of two repetitive deterministic sequences shall be used for both CP-OFDM and UW-OFDM. The CP-OFDM employs a preamble for B_{CP} number of OFDM blocks, while in UW-OFDM the preamble is used for $B_{\mathrm{UW}} \geq B_{\mathrm{CP}}$. Thanks to the unique word sequences after each payload block in UW-OFDM, per-block-synchronization allows B_{UW} to be much larger than B_{CP}.

Summing up the overhead requirements that were discussed in Sec. 4.1.1 and beginning of this section, the numbers of overhead samples for pilot-aided CP-OFDM and UW-OFDM follow[2]:

$$\xi_{\mathrm{CP+Pilots}} = \underbrace{N_p}_{\text{Preamble}} N_d + B_{\mathrm{CP}}(\underbrace{L}_{\text{CP}} N_d + \underbrace{\mathfrak{I}L}_{\text{Pilots}} N_d), \tag{4.1}$$

$$\xi_{\mathrm{UW}} = N_p N_d + B_{\mathrm{UW}} N_d N_{\mathrm{u}}, \tag{4.2}$$

respectively. Here, N_p is the preamble size, and for a UW-based system, we set it to $N_p = 2N_{\mathrm{u}}$ for two repeated UW sequences. Assuming $N_{\mathrm{u}} = (\mathfrak{I}+1)L$ and comparing (4.1) and (4.2), one may notice that if the number of transmission blocks B_{CP} and B_{UW} are equal, the number of overhead samples for CP-OFDM and UW-OFDM becomes equal too, i.e. $\xi_{\mathrm{CP+Pilots}} = \xi_{\mathrm{UW}}$ for $B_{\mathrm{CP}} = B_{\mathrm{UW}}$.

On the other hand, the total number of resources for both systems follow:

$$S_{\mathrm{CP}} = N_p N_d + B_{\mathrm{CP}} N_d (L + N_d), \tag{4.3}$$

$$S_{\mathrm{UW}} = N_p N_d + B_{\mathrm{UW}} N_d \big(N_{\mathrm{u}} + N_d\big). \tag{4.4}$$

Clearly, the resource size S_{UW} for UW-OFDM is always larger than S_{CP} for CP-OFDM, even if the number of blocks B_{UW} and B_{CP} are equal. Consequently, we calculate the energy efficiency by taking the ratio of the number of resources dedicated to useful data transmission with respect to the total number of resources, i.e.,

$$\eta_{\mathcal{E},\mathrm{Pilots}} = \frac{S_{\mathrm{CP}} - \xi_{\mathrm{CP+Pilots}}}{S_{\mathrm{CP}}}, \tag{4.5}$$

$$\eta_{\mathcal{E},\mathrm{UW}} = \frac{S_{\mathrm{UW}} - \xi_{\mathrm{UW}}}{S_{\mathrm{UW}}}. \tag{4.6}$$

The above energy efficiency of the two systems does not take into account the gain with respect to the modulation order μ and spatial multiplexing with $\mathfrak{I}$ Tx antennas. Therefore, assuming high SNR conditions, (4.5) and (4.6) shall be mapped into the upper bound of spectral efficiency (in bit/second/Hertz) via

$$\eta_{S,\mathrm{Pilots}} = \frac{\mathfrak{I}\mu B_{\mathrm{CP}}(N_d - \mathfrak{I}L)}{T_{F,\mathrm{CP}}\mathcal{B}}, \tag{4.7}$$

[2] Assuming full subcarrier allocation for N_d-point DFT size, L time domain samples of CP in (4.1) occupy N_d frequency bins, whereas, $\mathfrak{I}L$ frequency domain pilots are spread over N_d time samples.

$$\eta_{S,\mathrm{UW}} = \frac{\mathfrak{I}\mu B_{\mathrm{UW}} N_d}{T_{F,\mathrm{UW}}\mathcal{B}}, \tag{4.8}$$

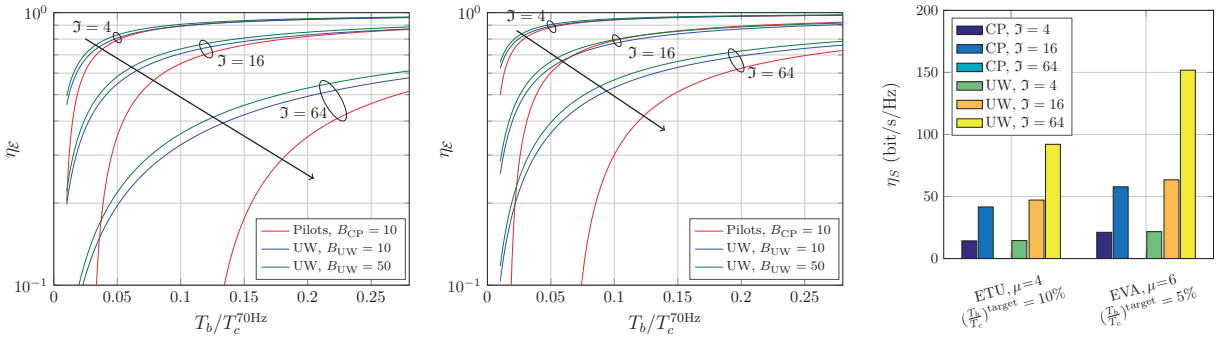
where $T_{F,\mathfrak{M}} = S_{\mathfrak{M}} T_s$ denotes the frame duration for $\mathfrak{M} \in \{\mathrm{CP}, \mathrm{UW}\}$. Furthermore, $\mathcal{B}$ denotes the bandwidth of the system and in case of full subcarrier allocation we have $\mathcal{B} = F_s$. Here, $F_s = \frac{1}{T_s}$ denotes the sampling frequency. Note that (4.7) and (4.8) describe only an upper bound to the spectral efficiency, because, due to further impairments, e.g. guard band insertion, channel codes and FER, the efficiency might degrade. Nevertheless, if the modulation order μ has been chosen appropriately by taking into account the coherence bandwidth and coherence time of the underlying scenario, $\eta_{S,\mathrm{Pilots}}$ and $\eta_{S,\mathrm{UW}}$ provide fair figures of merit for efficiency comparison.

4.1.3 Numerical Results

In this section, we make a numerical comparison of the energy and spectral efficiency of UW-based systems vs. CP-based systems in different MIMO setups. We assume that due to the mobility of the transceivers space-frequency pilots are necessary. The sampling frequency is set to $F_s = 1.92\mathrm{MHz}$ and the PDP is chosen to be similar to either LTE extended typical urban (ETU) or EVA, i.e. $L_{\mathrm{ETU}} = 9$ and $L_{\mathrm{EVA}} = 5$ channel taps, which correspond to 4.7µs and 2.6µs maximum delay spread, respectively. The UW length $N_{\mathrm{u}}^{\text{EVA-}\mathfrak{I}}$ for different antenna configurations and channel models is chosen to be $N_{\mathrm{u}}^{\text{EVA-4}} = 5^2$, $N_{\mathrm{u}}^{\text{ETU-4}} = 7^2$, $N_{\mathrm{u}}^{\text{EVA-16}} = 11^2$, $N_{\mathrm{u}}^{\text{ETU-16}} = 13^2$, $N_{\mathrm{u}}^{\text{EVA-64}} = 19^2$, $N_{\mathrm{u}}^{\text{ETU-64}} = 29^2$, and for the sake of primary synchronization a double length UW is used as preamble. In the CP-based system we also consider a preamble length of $N_p = 2N_u$ to be used for synchronization and afterwards, $B_{\mathrm{CP}} = 10$ blocks are continuously being transmitted. The CP length is set to $N_{\mathrm{cp}} = L$ samples, and also $\mathfrak{I}L$ orthogonal pilots are inserted into every block.

Fig. 4.2 compares the energy and spectral efficiency of the UW- vs. CP-based systems for two channel models and different number of Tx antennas. Here, we consider an intermediate mobility condition with maximum Doppler shift $f_d = 70\mathrm{Hz}$ and thus, the coherence time becomes $T_c^{70\mathrm{Hz}} \approx \frac{9}{16\pi f_d} = 2.56\mathrm{ms}$. In order to design the transmit blocks such that they experience near block-fading conditions, the block duration $T_b = N_d/F_s$ is commonly chosen to be smaller than 10% of the coherence time. Thus, in Fig. 4.2, we divide the x-axis by T_c in order to clearly find the corresponding target[3] $\frac{T_b}{T_c}$. As can be seen in Fig. 4.2, the energy efficiency of the UW-based systems with $B_{\mathrm{UW}} = 10$ and $B_{\mathrm{UW}} = 50$ is always larger than its corresponding CP-based system for short block duration, because the total resource size in UW-based systems is larger (see Sec. 4.1.2). On the other hand, if the coherence time T_c becomes larger, also larger block duration T_b can be chosen, and in that case, the efficiency curves of the UW- vs. CP-based systems tend to merge. We also note

[3] As a matter of fact, if the ratio $\frac{T_b}{T_c}$ is large, the channel varies too quickly within the block duration, and consequently, due to the channel estimation and equalization imperfections, it limits the choice of modulation order μ.



(a) ETU like channel, i.e. $L = 9$. **(b)** EVA like channel, i.e. $L = 5$. **(c)** Spectral efficiency.

Fig. 4.2: Energy and spectral Efficiency of the systems for different number of Tx antennas $\mathfrak{I}$ while $T_c^{70\text{Hz}} \approx 2.56$ ms.

that in each channel scenario, by increasing the number of Tx. antennas $\mathfrak{I}$, the energy efficiency of the systems decreases. Indeed, large number of Tx. antennas yield in increase of the channel parameters that need to be estimated, and therefore, more resources must be dedicated to reference signals (either UW or pilots) to estimate the $\mathfrak{I}L$ parameters. Although, comparing ETU and EVA like channels, $\eta_\mathcal{E}$ decreases with a slower pace for EVA, due to the smaller channel length L. The spectral efficiency of the MIMO systems has been depicted in Fig. 4.2c where we have assumed that the data is spatially multiplexed over all Tx antennas. Here, assuming a very selective channel, i.e. ETU, and 10% target $\frac{T_b}{T_c}$, i.e. channel varies slightly within the block, we choose 16-QAM, i.e. $\mu = 4$, as a sufficiently robust modulation order to decode the data. On the other hand, for an EVA like channel, we set the target $\frac{T_b}{T_c}$ to 5% in order to achieve near block-fading conditions, although, smaller $\frac{T_b}{T_c}$ yields increase of overhead. In this case, as the channel is less selective and nearly block-fading, choosing higher modulation order, e.g. $\mu = 6$ becomes feasible. Note that the 10% and 5% $(\frac{T_b}{T_c})^{\text{target}}$ correspond to 0.1 and 0.05 $T_b/T_c{}^{70\text{Hz}}$ in figures 4.2a and 4.2b, respectively. As we observe in Fig 4.2c, although the energy efficiency of the systems decreases with increase of Tx antennas, the spectral efficiency increases rapidly for the UW-based systems. Nonetheless, such condition is not necessarily true for a CP-based system, because, for $\mathfrak{I} = 64$ and short block duration T_b, the number of subcarriers are not sufficient for estimating all the channel parameters, and thus $\eta_{S,\text{Pilots}}$ for the chosen $(\frac{T_b}{T_c})^{\text{target}}$ and $\mathfrak{I} = 64$ becomes zero. We also note that in Fig. 4.2c, although $(\frac{T_b}{T_c})^{\text{target}}$ in EVA channel is chosen to be smaller than that of ETU channel, the spectral efficiency still increases due to the larger modulation order μ.

4.2 Sequences for UW and OOB Radiation

In order to select an optimal choice of UW sequences for MIMO wireless transmission, the following aspects must be taken into account:

(1) An accurate MIMO synchronization can be achieved if the periodic auto-correlation functions of the *modulatable* sequences are Kronecker's Delta function $\delta[n]$, and also, their periodic cross-correlation functions—of the sequences being transmitted from different antennas—are constant values which yield to minimized maximum absolute value [MS01],[SH88].

(2) From channel estimation perspective, the UW sequences must cover the entire band where the payload is being transmitted, and also they must have constant-modulus DFT (i.e. optimum in estimator's MSE sense [CSBM06]).

(3) From PAPR point of view, the sequences must be of constant magnitude in time domain.

(4) Near optimum sequences from OOB emission perspective should not have abrupt changes in their *real* and *imaginary* components of the time domain signal. In addition, when the UW is transmitted along with a payload signal, the transition between UW and the payload should not have abrupt changes either.

4.2.1 Orthogonal Polyphase Sequences

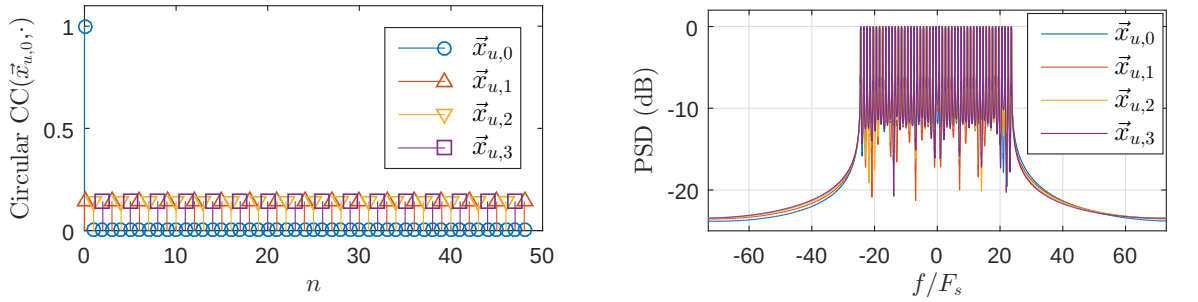
Similar to [MS01], we use the polyphase sequences proposed in [SH88] to generate the unique words for different Tx antennas. Thus, the UW sequence $x_{\mathrm{u},i}[n]$ from Tx antenna i is given by

$$x_{\mathrm{u},i}[n] = \exp(\frac{j2\pi(i+1)n_0 n_1}{\sqrt{N_{\mathrm{u}}}}), \tag{4.9}$$

where, $n_0 = \lfloor \frac{n}{\sqrt{N_{\mathrm{u}}}} \rfloor$ and $n_1 = (n \mod \sqrt{N_{\mathrm{u}}})$. Collecting the samples of $x_{\mathrm{u},i}[n]$ in form of a vector notation $\vec{x}_{\mathrm{u},i}$, N_{u} is the length of $\vec{x}_{\mathrm{u},i}$ and the square root $\sqrt{N_{\mathrm{u}}}$ must be a prime number [SH88]. We note that, the periodic auto-correlation function of $\vec{x}_{\mathrm{u},i}$ is a Kronecker's Delta function $\delta[n]$, while the periodic cross-correlation of $\vec{x}_{\mathrm{u},i}$ and $\vec{x}_{\mathrm{u},i' \neq i}$ is a constant value equal to $1/\sqrt{N_{\mathrm{u}}}$ [SH88]. The magnitude of $\vec{x}_{\mathrm{u},i}$ in both time and DFT domains is constant and The power spectral density (PSD) of the sequences for $i = \{0, 1, 2, 3\}$ with $\sqrt{N_{\mathrm{u}}} = 7$ is depicted in Fig. 4.3b. Nevertheless, Fig. 4.3b shows the OOB radiation of the signal if the UW sequence is solely being transmitted, without presence of any multi-carrier modulation. In the following, we show that integrating the sequence with a low OOB emitting multi-carrier can still preserve the desired signal characteristics.

4.2.2 Waveform Engineering for UW Sequences combined with GFDM

Let $\vec{x}_i = [\vec{x}_{d,i}^T \ \vec{x}_{\mathrm{u},i}^T]^T$ be the transmit signal from Tx antenna i, where the payload vector $\vec{x}_{d,i}$ is linearly modulated by a multi-carrier system. In case of a flexible and low OOB



(a) Circular CC of $\vec{x}_{\mathrm{u},i}$ **(b)** PSD of $\vec{x}_{\mathrm{u},i}$ with 1/3 subcarrier allocation.

Fig. 4.3: Signal characteristics of the orthogonal Polyphase sequences [SH88] with $\sqrt{N_{\mathrm{u}}} = 7$.

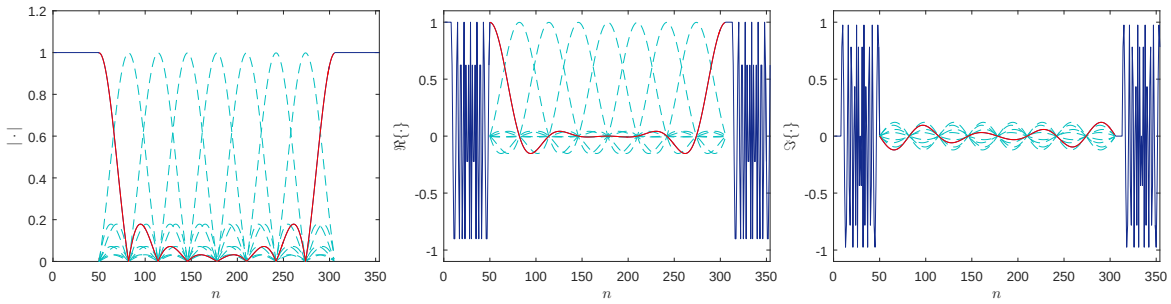


Fig. 4.4: The complex valued UW-GFDM signal in time domain (dark blue: UW, red: first GFDM subsymbol, light blue: rest of GFDM subsymbols) at zeroth subcarrier. If (4.11) holds, the red curve remains constant and does not get influenced by random data.

emitting waveform (e.g. GFDM), $\vec{x}_{d,i} \in \mathbb{C}^{N_d}$ is given by [MMG+14]

$$\vec{x}_{d,i} = \boldsymbol{A}\vec{d}_i, \tag{4.10}$$

where $\vec{d}_i = \mathtt{vec}(\mathbf{D}_i)$, in which $\mathbf{D}_i \in \mathbb{C}^{K\times M}$ is the allocation matrix for K subcarriers and M subsymbols, and its elements comprise data mapped to a modulation alphabet.

A UW-GFDM block would maintain GFDM's low OOB emission, provided that the transition between the UW sequence and the GFDM signal is a smooth transition without any abrupt change of the signal. Such condition shall be achieved by setting the first GFDM subsymbol as a guard symbol (GS). Thus, we have

$$d_{0,i}[k] = \delta[k - k_p], \tag{4.11}$$

where $(\vec{d}_{0,i})_{k=0:K-1}$ is the first column of $\mathbf{D}_i$ associated to the first subsymbol. The constant k_p can be any subcarrier within the band. Here, the Kronecker's Delta function $\delta[\cdot]$, generates the shape of the time domain prototype filter as the time domain signal. If its energy is normalized to its maximum value, its time domain signal begins with $x_{d,i}[0] = 1$ and also ends with $x_{d,i}[N_d - 1] = 1$. Having $x_{u,i}[0] = 1$ and $x_{u,i}[N_{\mathrm{u}} - 1] \approx 1$, a smooth

transition between UW and GFDM signal is achieved. Fig. 4.4 shows an example of the signal in time domain, in which the GFDM signal is surrounded by two UW sequences. As one can see, the first GFDM subsymbol that wraps around the block has the key impact on the transition between the UW and the GFDM block. Thus, having the first subsymbol as a GS, a smooth transition is achieved at the boundary of UW and GFDM blocks. Nevertheless, the abrupt changes of the *real* and *imaginary* components of the UW itself, would have negative impacts on the overall OOB radiation of the signal.

Assuming the energy of $\delta[k-k_p]$ to be $\sqrt{K}$, (i.e. the normalization factor of the prototype filter $\vec{g}_{k,m}$), the GS insertion of the above approach yields an increase of overhead by the factor $\frac{1}{M\sqrt{K}}$.

4.2.3 Simulation Results for OOB Emission of UW-GFDM

The performance of the UW based GFDM in terms of OOB radiation is depicted in Fig. 4.5. Here, the GFDM waveform is configured to have $K_{\text{on}} = 22$ active subcarriers out of $K = 64$ subcarriers, and a raised-cosine filter [NMZF17] with roll-off factor $\alpha = 0.3$ has been chosen for it. The total number of subsymbols is set to $M = 16$ and all M subsymbols are active for the basic GFDM as well as UW-GFDM. On the other hand, UW-GS-GFDM and GS-GFDM use $M_{\text{on}} = 15$ active subsymbols while their first subsymbol is configured as a guard symbol according to the description of Sec. 4.2.2. Moreover, the UW sequence is initially obtained via (4.9) by letting $N_{\text{u}} = K_{\text{on}}^{\text{UW}} = 25$ and afterwards, it is zero-padded to the total number of $K^{\text{UW}} = 72$ subcarriers. The PSD of a CP-OFDM signal with the same number of subcarriers and UW-DFT-spread OFDM (UW-DFT-s-OFDM) based on the approach of [SYGO15] are also plotted as benchmarks. Here, the UW sequences in DFT-s-OFDM are initially generated in form of PN-sequence that are mapped to QPSK symbols and then circularly filtered by a sinc function. Note that the comparison with UW-DFT-s-OFDM remains valid for only SISO systems[4].

As can be seen in Fig. 4.5, UW-GFDM without a guard symbol increases the OOB emission of basic GFDM which is due to the abrupt changes of the signal at the boundary of UW sequence and the GFDM signal. Setting the first GFDM subsymbol as a GS, we observe that OOB radiation of UW-GS-GFDM reduces by an order of magnitude. However, UW-GS-GFDM still suffers from higher OOB emission when it is compared to GS-GFDM, which is due to the abrupt changes of the signal within the UW sequence. We also note that at frequency of around $\frac{f}{F_s} \approx 31$, there is a peak of energy for UW-GS-GFDM and GS-GFDM which is due to the $\delta[k-k_p]$ function with energy $\sqrt{K}$. Comparing the UW-GS-GFDM with UW-DFT-s-OFDM, we observe that DFT-s-OFDM achieves a slightly smaller OOB radiation due to the frequency domain filtering of the UW sequence,

[4] As the UW-DFT-s-OFDM signal is jointly generated by precoding the data and filtering deterministic QPSK symbols, the resulting UW sequence does not provide the optimum cross-correlation properties for a high quality MIMO synchronization.

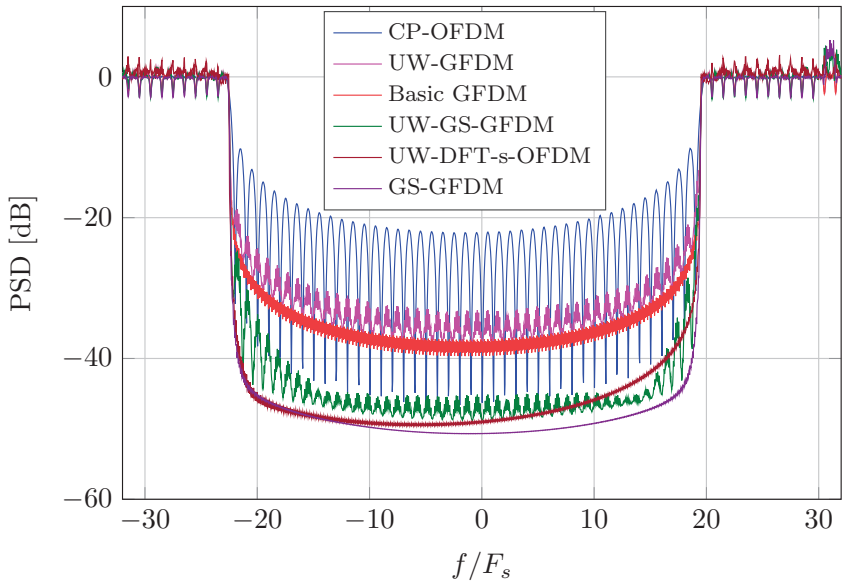


Fig. 4.5: Out of band radiation.

whereas the UW-GS-GFDM employs the sequence (4.9) directly. We also note that the PSD curve of UW-DFT-s-OFDM has an asymmetrical shape, because of the employed PN-sequence.

4.3 Synchronization

In this section, describing a centralized $\mathfrak{I} \times \mathfrak{Q}$ MIMO channel, we derive the corresponding time and frequency synchronization metrics for a robust transmission over doubly-dispersive channels.

4.3.1 Transmission over a Centralized MIMO Wireless Channel

We consider that in a MIMO system, the *Payload-UW* blocks are structured according to e.g. Fig. 4.6, and they are simultaneously transmitted from $\mathfrak{I}$ Tx antennas. In addition to the basic UW transmission, i.e. $U'_{i,\text{basic}} = U_i$, we also consider an alternative case, in which the UW after the payload block circularly rotates by $\mathfrak{I}/2$ for each antenna index i, i.e.

$$U'_{i,\text{circ}} = \Big((i + \frac{\mathfrak{I}}{2}) \mod \mathfrak{I}\Big). \tag{4.12}$$

The transmission is initiated by sending a double UW sequence in form of a preamble, which assures a primary time-frequency synchronization and channel estimation. Afterwards blocks of *Payload-UW* with length N are transmitted at a block-rate F_b.

We assume that all Tx antennas are collocated and they use a single oscillator clock source. Likewise, all the Rx antennas are collocated and they use a shared clock source, too. We

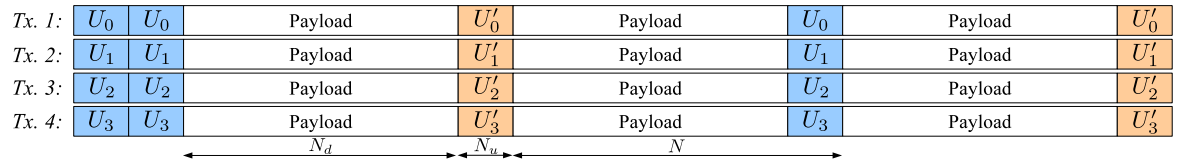


Fig. 4.6: Example of the frame structure for $\mathfrak{I} = 4$ Tx antennas. In case of basic UW transmission we set $U'_i = U_i$. In case of Circ.-UW frames (4.12) should hold.

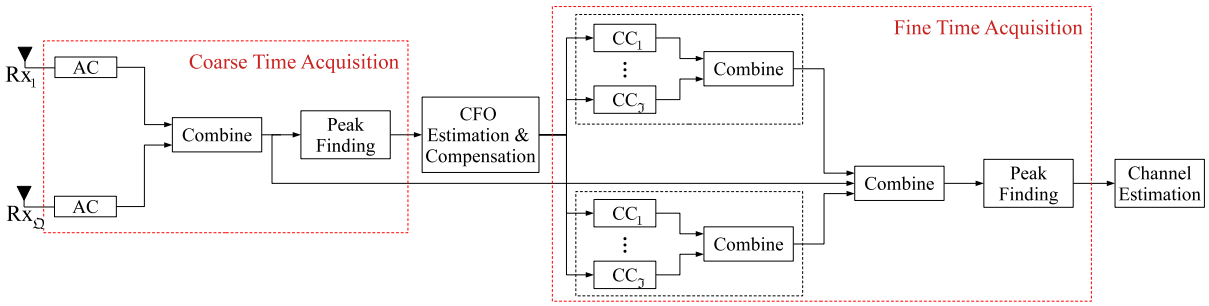


Fig. 4.7: Block diagram of the MIMO synchronization unit.

also assume that the transmit signals from $\mathfrak{I}$ Tx antennas pass through $\mathfrak{I} \times \mathfrak{Q}$ *multi-path* channels with identical PDP (due to the collocation of the antennas). With a further assumption of ideal front-ends, the receive signal $r_q[n]$ at antenna q is characterized by the following expression:

$$r_q[n] = \exp(j2\pi n \frac{\varepsilon}{N}) \sum_{i=0}^{\mathfrak{I}-1} \sum_{\ell=0}^{L-1} h_{\ell,i,q}[n] x_i[n - \ell - \vartheta] + w_q[n], \tag{4.13}$$

where, ϑ and ε are the parameters describing symbol time offset (STO) and CFO, respectively. Herein, CFO is normalized with respect to the block-rate F_b with periodicity $1/N$. $h_{\ell,i,q}[n]$ is the ℓ-th tap of the channel impulse response's complex envelop at time sample n between antennas i and q. The zero-mean AWGN process at Rx antenna q is denoted by $w_q[n]$. Note that in (4.13), we only consider single parameters of ϑ and ε for all Tx-Rx antenna pairs, since a centralized MIMO system is being assumed. In case of a distributed MIMO system with different clock sources, the STO and CFO parameters need to be considered for each Tx-Rx antenna pair.

4.3.2 Coarse Time Acquisition

Fig. 4.7 shows an overview of the synchronization unit. Here, an auto-correlation (AC) based approach is used to find an approximate starting position of the frames. Given the coarse starting sample of the UWs, we estimate and partially remove the CFO from the raw received signal. Afterwards, a fine time acquisition approach is employed, in order to find the exact starting sample of the frames.

Exploiting the repetition of the UW sequences $\vec{x}_{\mathrm{u},i}$ as well as their constant cross-correlation (CC) properties between the sequences transmitted from different Tx antennas,

one can apply the AC based approach [SC97] to obtain an approximate starting position of the preamble. Thus, auto-correlating the received signal from antenna q over a window of $\mathfrak{W}_{\mathrm{AC}}$ samples that are N_{AC} samples apart from each other, the AC metric is given by

$$\Phi_q[n] = \frac{|\phi_q[n]|^2}{|P_{r,q}[n]|^2}, \tag{4.14}$$

where,

$$\phi_q[n] = \sum_{i=n}^{n+\mathfrak{W}_{\mathrm{AC}}-1} r_q'^{*}[i]\; r_q'[i+N_{\mathrm{AC}}], \tag{4.15}$$

$$P_{r,q}[n] = \sum_{i=n}^{n+\mathfrak{W}_{\mathrm{AC}}-1} |r_q'[i+N_{\mathrm{AC}}]|^2. \tag{4.16}$$

Herein, the AC window size is set to $\mathfrak{W}_{\mathrm{AC}} = N_{\mathrm{u}}$ for both primary as well as UW specific synchronizations. Moreover, for primary synchronization (i.e. $\Phi_q^{\mathrm{Primary}}[n]$) where we use the preamble, we set $N_{\mathrm{AC}}^{\mathrm{Primary}} = N_{\mathrm{u}}$ and $r_q'^{\mathrm{Primary}}[n] = r_q[n]$, whereas for UW specific synchronization in which, the repetition of UW sequences after the payload blocks are exploited for $\Phi_q^{\mathrm{UW}}[n]$, we set $N_{\mathrm{AC}}^{\mathrm{UW}} = N$ and $N_{\mathrm{AC}}^{\mathrm{Circ.\text{-}UW}} = 2N$ for basic UW and Circ.-UW frame designs, respectively. Moreover, for both UW frame designs we have $r_q'^{\mathrm{UW}}[n] = \hat{r}_q[n]$, where the signal $\hat{r}_q[n]$ contains only residual CFO values that has not been estimated via preamble, and it will be further discussed in Sec. 4.3.3.

The AC metrics $\Phi_q[n]$ have local peaks that reveal an approximate starting position of the first transmitted UW (over two repetitive UW sequences). Moreover, since only a single parameter ϑ needs to be estimated, we combine the AC metrics $\Phi_q[n]$ over all Rx antennas, and thus, we obtain the metrics

$$\Phi_{\mathrm{sum}}[n] = \sum_{q=0}^{\mathfrak{Q}-1} \Phi_q[n], \tag{4.17}$$

$$\Phi_{\mathrm{prod}}[n] = \prod_{q=0}^{\mathfrak{Q}-1} \Phi_q[n]. \tag{4.18}$$

Here, $\Phi_{\mathrm{sum}}[n]$ is the conventional approach for combining the AC metrics [ZZZW05], whereas $\Phi_{\mathrm{prod}}[n]$ is an alternative approach that we consider in this work. The main reason for choosing $\Phi_{\mathrm{prod}}[n]$ is that in a centralized MIMO systems where the peaks of individual AC metrics occur at almost the same location, taking their product makes the combined metric sharper than taking their average. In addition, we will also further justify such an alternative metric selection in section 4.3.5.

Searching the local peaks over the metric $\Phi[n]$, the coarse STO acquisition is given by

$$\hat{\vartheta}_{\mathrm{AC}} = \arg\max_n \Phi[n] \text{ for } n \in \{bN, bN+1, \cdots, bN+N\}. \tag{4.19}$$

where $b = 0, 1, \cdots, B-1$ is the block index.

4.3.3 CFO Estimation and Removal

Once the approximate starting position of the UW sequence has been obtained, CFO value ε needs to be estimated. For the primary synchronization where we use the repetition of the UW sequences within the preamble, the CFO is estimated as:

$$\hat{\varepsilon}_p = \frac{1}{\mathfrak{Q}} \sum_{q=0}^{\mathfrak{Q}-1} \frac{N}{N_\mathrm{u}} \frac{\angle \Phi_q^{\mathrm{Primary}}[\hat{\vartheta}_{\mathrm{AC}}^{\mathrm{Primary}}]}{2\pi}. \tag{4.20}$$

Herein, we have the fractional term $\frac{N}{N_\mathrm{u}}$ because the UW sequences within the preamble have a periodicity of $1/N_\mathrm{u}$, while the CFO is normalized with respect to the block-rate with periodicity $1/N$. Hence, the primary synchronization can estimate the CFO values of up to $\frac{N}{2N_\mathrm{u}}$ (i.e. fractional as well as integer CFO values smaller than $\frac{N}{2N_\mathrm{u}}$).

Once an initial CFO value has been estimated, we remove it from the Rx signal via an inverse phase shift, i.e.

$$\hat{r}_q[n] = \exp(-j2\pi n \hat{\varepsilon}_p / N) r_q[n]. \tag{4.21}$$

Note that the signal $\hat{r}_q[n]$ still contains some residual CFO which has not been estimated via (4.20). Thus, exploiting the repetition of the UW sequences before and after the payload blocks, one can estimate the residual CFO via the secondary UW specific synchronization. In this case, after calculating the AC function $\Phi_q^{\mathrm{UW}}[n]$ and finding the corresponding coarse STO $\hat{\vartheta}_{\mathrm{AC}}^{\mathrm{UW}}$, we calculate the residual CFO via

$$\hat{\varepsilon}_{\mathrm{res}} = \frac{1}{\mathfrak{Q}} \sum_{q=0}^{\mathfrak{Q}-1} \frac{\angle \Phi_q^{\mathrm{UW}}[\hat{\vartheta}_{\mathrm{AC}}^{\mathrm{UW}}]}{2\pi}. \tag{4.22}$$

We note that the estimation range for residual CFO via (4.22) is $\frac{1}{2}$ and $\frac{1}{4}$ of the block-rate F_b for basic UW and Circ.-UW frames, respectively. Intuitively, considering the fact that the same energy of the UW sequences are focusing on a smaller range of CFO values, one can expect the expression (4.22) to have a higher accuracy with respect to (4.20). In addition, if the CFO ε remains constant over B number of payload-UW blocks, the estimations $\hat{\varepsilon}_{b,\mathrm{res}}$ associated to blocks b can be accumulated and thus, a more accurate estimation

$$\hat{\varepsilon}_{\mathrm{res}}^{B} = \frac{1}{B} \sum_{b=0}^{B-1} \hat{\varepsilon}_{b,\mathrm{res}}, \tag{4.23}$$

shall be obtained.

In the following, we analyze the accuracy of the above estimators by calculating their CRLB:

Let $z_q[n] = \exp(j2\pi n\varepsilon/N)\mathfrak{y}_q[n] + w_q[n]$ be a sufficient statistic for estimating ε, where $\mathfrak{y}_q[n] = \sum_{i=0}^{\mathfrak{I}-1} \sum_{\ell=0}^{L-1} h_{\ell,i,q} x_i[n-\ell]$ is a perfectly time-and-frequency synchronized sample of noiseless received sequence at antenna q for stationary channel $h_{\ell,i,q}$ from PDP $\vec{P}$ that is normalized to unity. Stacking the observed samples $z_q[n]$ into a vector form $\vec{z} =$

$([\vec{z}_0^T \ \cdots \ \vec{z}_{\mathfrak{Q}-1}^T])_{n\in\mathcal{N}_u}^T$, the complex pdf of $\vec{z}$ conditioned by the unknown parameters ε and $h_{\ell,i,q}[n]$ is given by

$$p(\vec{z}|\varepsilon, h_{\ell,i,q}) = \frac{1}{2\pi\sigma_w^2}\exp\left\{-\frac{1}{2\sigma_w^2}\sum_{q=0}^{\mathfrak{Q}-1}\sum_{n\in\mathcal{N}_u}\left|z_q[n]-\exp(j2\pi n\varepsilon/N)\mathfrak{y}_q[n]\right|^2\right\}, \quad (4.24)$$

where $\mathcal{N}_u$ is the set of indexes n for all observed sequences. In case of primary synchronization, $\mathcal{N}_u^{\text{Primary}} = \{0, 1, \cdots, 2N_\text{u}-1\}$. If UW specific synchronization is of interest, $\mathcal{N}_u^{\text{UW}} = \{0, 1, \cdots, N_\text{u}-1, \varrho N, \varrho N+1, \cdots, \varrho N+N_\text{u}-1\}$. If $B > 2$ number of UW sequences are being observed, $\mathcal{N}_u^{B\cdot\text{UW}} = \{0, 1, \cdots, N_\text{u}-1, \varrho N, \varrho N+1, \cdots, (B-1)\varrho N+N_\text{u}-1\}$. Here, $\varrho = 1$ and $\varrho = 2$ for basic UW and Circ.-UW frames, respectively.

Consequently, after some mathematical manipulations (see Appendix A.6), the *Fisher information* [Kay93] for $\vec{z}$ follows:

$$-\mathbb{E}\left[\frac{\partial^2 \ln p(\vec{z}|\varepsilon, h_{\ell,i,q})}{\partial\varepsilon^2}\right] = \frac{1}{2\sigma_w^2}\sum_{n\in\mathcal{N}_u} 8\pi^2\frac{n^2}{N^2}\mathbb{E}_h\left[\sum_{q=0}^{\mathfrak{Q}-1}|\mathfrak{y}_q[n]|^2\right], \quad (4.25)$$

Assuming $h_{\ell,i,q}[n]$ defined in (4.13) to be independent of ε, we have

$$\mathbb{E}_h\left[\sum_{q=0}^{\mathfrak{Q}-1}|\mathfrak{y}_q[n]|^2\right] = \mathfrak{Q}\mathfrak{I}, \quad (4.26)$$

where $\mathbb{E}_h[\cdot]$ takes the expected value with respect to $h_{\ell,i,q}$. Substituting the elements of $\mathcal{N}_u^{\text{Primary}}$ and $\mathcal{N}_u^{\text{UW}}$ into (4.25), one may note that the Fisher Information of UW specific synchronization is always larger for ε estimation than that of the primary synchronization. Although, the CFO estimation approach (4.22) requires smaller CFO range with respect to the (4.20). We also note that since the pdf (4.24) does not satisfy the *regularity condition* [Kay93, Sec. 3.4], an unbiased estimator of ε that attains the CRLB does not exist.

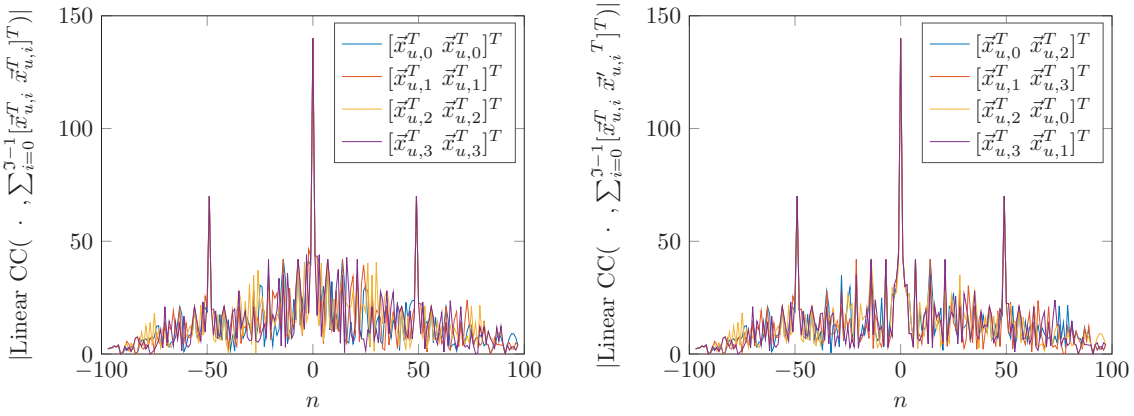
4.3.4 Fine Time Acquisition

In order to find the exact starting position of the preamble or a UW-payload-UW frame, we use a CC based (i.e. matched filter) approach of [TEF99] for the MIMO applications. Taking the cross-correlation of the CFO-compensated observed signal $\hat{r}_q[n]$ with $\vec{x}_{u,i}$, the metrics for primary, UW specific and Circ.-UW synchronizations follow

$$\Psi_{i,q}[n]^{\text{Primary}} = \frac{1}{2N_\text{u}}\sum_{\iota=0}^{2N_\text{u}-1}\hat{r}_q^*[n+\iota]x_{u,i}[\iota \mod N_\text{u}], \quad (4.27)$$

$$\Psi_{i,q}[n]^{\text{UW}} = \frac{1}{N_{u,\text{cc}}N_\text{u}}\Big(\sum_{n_\text{cc}=0}^{N_{u,\text{cc}}-1}\sum_{\iota=n_\text{cc}N}^{n_\text{cc}N+N_\text{u}-1}\hat{r}_q^*[n+\iota]x_{u,i}[-n_\text{cc}N+\iota]\Big), \quad (4.28)$$

$$\Psi_{i,q}[n]^{\text{Circ.-UW}} = \frac{1}{N_{u,\text{cc}}N_\text{u}}\Big(\sum_{n_\text{cc}=0}^{N_{u,\text{cc}}-1}\sum_{\iota=n_\text{cc}N}^{n_\text{cc}N+N_\text{u}-1}\hat{r}_q^*[n+\iota]x'_{u,i}[-n_\text{cc}N+\iota]\Big), \quad (4.29)$$



(a) basic UW **(b)** Circ.-UW

Fig. 4.8: CC metrics in noise-less flat 4×4 MIMO channel

respectively. Here in (4.28) and (4.29), $N_{u,\mathrm{cc}}$ denotes the number of UW sequences that would be used for calculation of the CC metric. For instance, setting $N_{u,\mathrm{cc}} = 2$ takes the cross-correlation over two UW sequences within the UW-Payload-UW block. Furthermore, in (4.29), we have $x'_{u,i}[n] = x_{u,(i+\frac{\mathfrak{I}}{2} \mod \mathfrak{I})}[n]$. Note that in basic UW synchronization i.e. (4.28), the reference signal comprises a repetition of $\vec{x}_{u,i}$, but in Circ.-UW synchronization, i.e. (4.29), we apply the CC metric over $[\vec{x}_{u,i}^T\ (\vec{x}'_{u,i})^T]^T$. An example of the two metrics in a noise-less flat 4×4 MIMO Channel is depicted in Fig. 4.8. As one can see, the metrics are not entirely equal. Indeed, the main peak of basic UW is slightly sharper than the main peak of Circ.-UW.

Moreover, given the MIMO system with $\mathfrak{I} \times \mathfrak{Q}$ antennas, we would have $\mathfrak{I}\mathfrak{Q}$ CC metrics. In conventional centralized MIMO systems, the CC metrics of different antenna pairs are summed together and a single combined metric is obtained as [MS01],[CT08] $\Psi_{\mathrm{sum}}[n] = \sum_{q=0}^{\mathfrak{Q}-1} \sum_{i=0}^{\mathfrak{I}-1} |\Psi_{i,q}[n]|$, where $\Psi_{i,q}[n]$ can be either from the primary or UW specific CC metrics. In this work however, we additionally combine the CC metrics by the AC metrics, i.e.

$$\Psi_{\mathrm{prod}}[n] = \Phi_{\mathrm{prod}}[n]\Psi_{\mathrm{sum}}[n]. \tag{4.30}$$

The approach (4.30) is motivated by the following two main reasons:

- Authors in [AKE08] and [GMMF14] showed that in a stationary channel condition, multiplying the AC metric with the CC metric—in SISO systems—filters the side peaks that occur at half preamble length. Indeed, the AC metric has a peak near to the main peak of CC metrics, but at half preamble length where CC metrics have considerable side peaks, the AC metric is significantly low and thus, multiplying them, filters out such side peaks. For further detailed information, we refer the interested readers to [AKE08] and [GMMF14].

- The metric (4.30) is further beneficial for low SNR time-varying channel conditions where the noise power as well as CFO due to the Doppler shifts may cause a complete

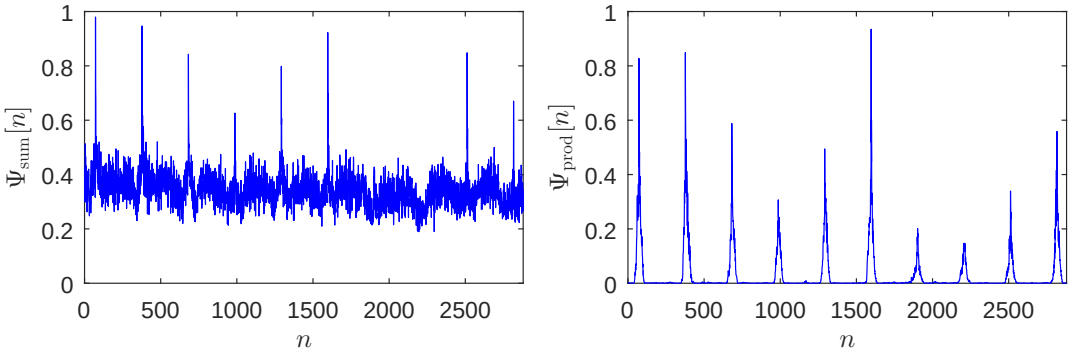


Fig. 4.9: Example of the CC metrics for 10 blocks in -5 dB SNR, $L = 9$ channel taps and 4.77% normalized Doppler frequency.

loss of peaks in CC metrics. For instance, Fig. 4.9 compares the $\Psi_{\text{prod}}[n]$ metric with the conventional approach [MS01],[CT08] $\Psi_{\text{sum}}[n]$ in a 4×4 MIMO system, in which 10 blocks (and thus 10 peaks) must be detected. As can be seen, relying solely on $\Psi_{\text{sum}}[n]$, the peaks of the 7th and the 8th blocks are completely lost inside the noise floor, which is due to the low SNR as well as time- and frequency-selectivity. On the other hand, observing the $\Psi_{\text{prod}}[n]$ metric, one may note that, not only the noise floor is being suppressed to near zero values, but also the two peaks corresponding to the 7th and 8th blocks are being recovered.

The exact starting position of the frame (i.e. preamble or *UW-payload-UW*) is obtained by

$$\hat{\vartheta} = \arg\max_{n} \Psi[n] \text{ for } n \in \{bN, bN+1, \cdots, bN+N\}. \tag{4.31}$$

where $\Psi[n]$ can be either $\Psi_{\text{sum}}[n]$ or $\Psi_{\text{prod}}[n]$.

4.3.5 Simulation Results

In this section, we validate the proposed synchronization techniques for the MIMO UW-based transmission systems via Monte-Carlo simulations and we compare their performances with the conventional approach.

Consider a 4×4 MIMO system in a high mobility urban scenario similar to LTE ETU 300, where the multi-path wireless channels are characterized by Rayleigh distribution. The PDP is exponentially decaying from 0 to -20 dB with the longest delay tap of 4.7µs (i.e. $L = 9$) and it is normalized to unity. The sampling rate is set to $F_s = 1.92$MHz. The UW length is set to $N_{\text{u}} = 49$ samples, whereas the payload slots have a length of $N_d = 256$ samples. Thus, a block contains $N = 305$ samples leading to a block rate of $F_b = 6.3$kHz. In our simulations, we consider time-varying channel conditions with maximum Doppler shift $f_d = 300$Hz (i.e. $\nu_d = \frac{f_d}{F_b}$ =4.77% Normalized Doppler frequency) modeled according to [Cla68]. After the preamble, $B = 20$ blocks are being transmitted.

The STO is assumed to be integer valued, while the normalized CFO range is assumed to be uniformly distributed within the range $\mathcal{U}[2, 3]$, at the beginning of each transmission. Thus, for the given sampling frequency, CFO varies from 12.6kHz to 18.9kHz, although, during the transmission of B blocks, it is assumed to remain constant.

Time Synchronization. The performance comparison of the proposed time acquisition, i.e. (4.30) vs. the conventional approach, i.e. $\Psi_{\text{sum}}[n]$ is depicted in Fig. 4.10, where we show the probability of a successful synchronization as well as the probability of perfect synchronization. In a successful synchronization, we assume that $|\vartheta - \hat{\vartheta}| < 5$ samples, whereas in a perfect synchronization $|\vartheta - \hat{\vartheta}| = 0$. Here, a perfect synchronization is specifically of our interest, because it has a significant impact on the channel estimation. In other words, an imperfect synchronization causes energy leakage in channel estimation [CT08], whereas a perfect synchronization would have the minimum number of channel taps that need to be estimated. In Fig. 4.10, the probability of successful synchronization for SNR values of > 5 dB remains *one* for both approaches. Although, at low SNR values, we note that the conventional approach loses around less than 20% of the blocks (i.e. 0.82 success rate at -4 dB SNR), which is due to the fact that, time- and frequency-selectivity of the channel yields to variable power of the peaks, and in low SNR conditions, if a peak power becomes equal or smaller than the noise floor, the synchronization unit would miss the corresponding peak detection. On the other hand, via the proposed approach, i.e. Ψ_{prod}, where the AC and CC metrics are combined by means of taking their product, the probability of successful synchronization always remains close to *one*, even at low SNR regions. Similar behavior has also been observed for the probability of perfect synchronization (i.e. the exact starting position of the block associated to the first channel tap). However, one may note that, even at very high SNR values, neither approaches achieve the full certainty in perfect synchronization, because in a multi-path environment, the second tap might occasionally have higher power. Such condition is commonly handled by taking the DFT window few samples earlier than what has been detected by the synchronization unit. More details concerning a safe DFT window vs. ISI-limited DFT window can be found in [CT08, Chap. 5].

Fig. 4.11 compares the time synchronization performance of basic UW design vs. Circ.-UW, as well as two cases where we use two UW sequences over the UW-Payload-UW block vs. four UW sequences over three blocks for calculation of the CC metrics. Here, the curves have been plotted for low SNR regions, because at SNR values of > 10 dB, the performance of all metrics saturate similar to Fig. 4.10. Comparing the synchronization performance of basic UW frames vs. Circ-UW frames, we observe that Circ.-UW frames have marginally lower probability of successful and perfect synchronization, which is due to the fact that the basic UW frames have slightly sharper peaks than Circ.-UW frames (see Fig. 4.8). On the other hand, if we compare the performances of 2 UW sequences vs. 4 UW sequences, one may note that using 4 UW sequences—with slight increase of computational efforts—

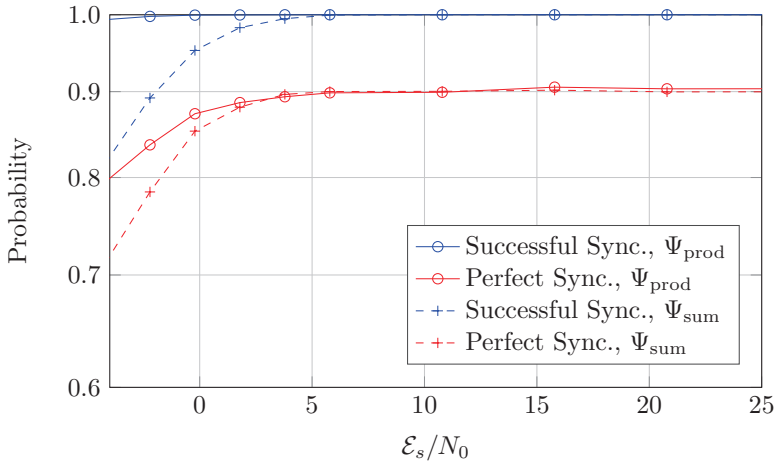


Fig. 4.10: Probability of Successful and Perfect synchronization of the proposed approach for basic UW vs. conventional approach at a frequency selective ($L = 9$ taps) and time-varying channel condition ($f_d = 300$ Hz).

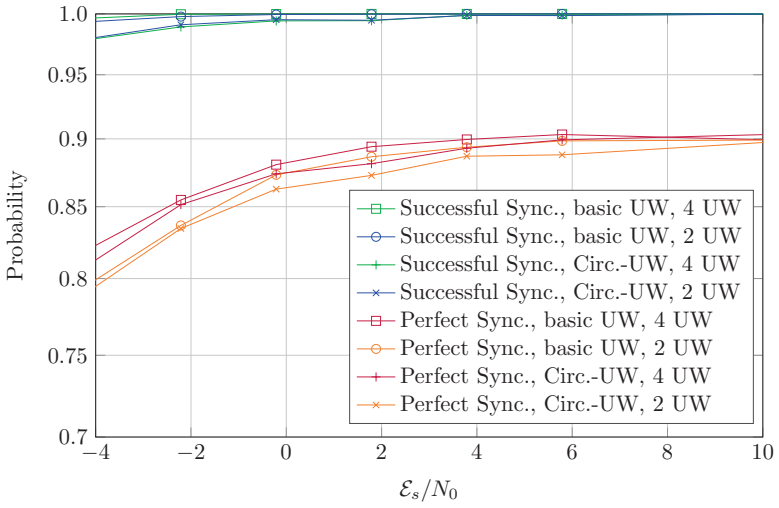


Fig. 4.11: Probability of Successful and Perfect synchronization for basic UW vs. Circ.-UW designs, as well as 2 UW sequences vs. 4 UW sequences.

for the CC metric calculation, improves the probability of perfect synchronization at very low SNR values. Nevertheless, the probability of successful synchronization is almost identical for both 2 UW vs. 4 UW cases.

Frequency Synchronization. The performance of CFO estimations in terms of MSE are provided in Fig. 4.12 and Fig. 4.13. The dotted lines in Fig. 4.12 are the CRLB for their corresponding estimation techniques that are plotted with the same color. The blue lines are associated to the primary CFO estimation $\hat{\varepsilon}_p$, the red lines belong to residual CFO estimation $\hat{\varepsilon}_{\text{res}}$ in which $\hat{\varepsilon}_p$ has also been used for their initial CFO estimation. The green lines are the MSE for residual CFO estimation $\hat{\varepsilon}_{\text{res}}^B$ that has been averaged over $B = 20$ blocks. Note that for the CRLB calculations, we have assumed perfect timing

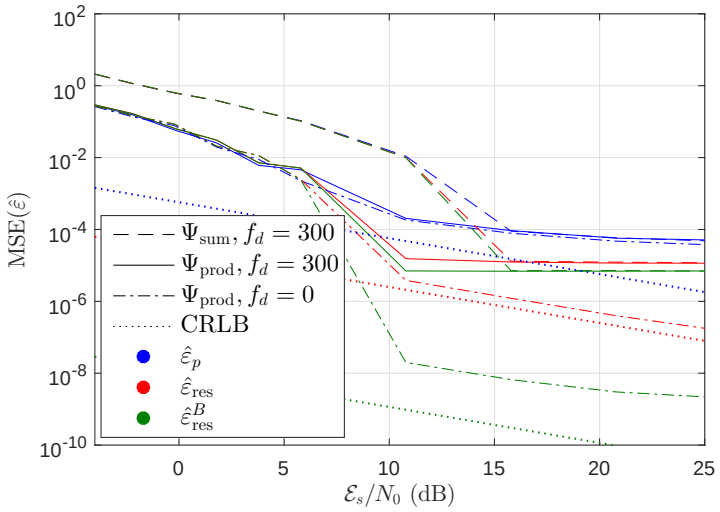


Fig. 4.12: CFO estimation performance for basic UW vs. conventional approach.

synchronization, whilst for the CFO estimations, we have used the peaks of AC metrics $\Phi[n]$. Therefore, even in a time-invariant scenario (i.e. $f_d = 0$, dash-dotted lines) an error floor is added up to the gap between the $\hat{\varepsilon}^B_{\text{res}}$ and its CRLB. Comparing the performances in time-variant situations, one can see that at MSE($\hat{\varepsilon}$) $= 10^{-2}$, there exists almost 7dB performance gain between the approaches with Φ_{prod} vs. Φ_{sum} (i.e. conventional). In addition, the CFO estimation from Φ_{prod} metric reaches the error floor (due to channel variations) at an SNR value of around 11dB, while the Φ_{sum} metric achieves its best estimation performance at 16dB SNR. We justify this performance gap, by considering the fact that taking the product of the AC metrics in Φ_{prod} yields to sharper peaks and thus, higher certainty in coarse time acquisition than Φ_{sum}. We also note that the performance of conventional CP based systems would always be limited to the performance of preamble based estimations i.e. blue curves.

Comparing the performances of basic UW vs. Circ.-UW frame designs (i.e. Fig. 4.13), one can see that at static channels where $f_d = 0$, an improvement of CFO estimation is obtained via Circ.-UW at high SNR regions, because first of all, the accuracy of primary CFO estimation lets the residual CFO to remain within $\frac{1}{4}$ of the block-rate (i.e. estimation range of Circ.-UW), and secondly, the CRLB of Circ-UW cases are at lower levels. Nevertheless, neither of the approaches achieve CRLB for three main reasons, i.e. *a*) the CFO estimators use the peaks of AC metrics. *b*) the channels are frequency selective for $f_d = 0$ and doubly-dispersive for $f_d = 300$. *c*) the pdf (4.24) does not satisfy the regularity condition. We also note that, Circ.-UW frame design does not achieve a considerable gain with respect to basic UW frames in high mobility conditions, because at low SNR regions, the uncertainties of peak detection dominate the performance, whilst, at high SNR regions, the error floor due to the time varying channel dominates the CFO estimation for both cases.

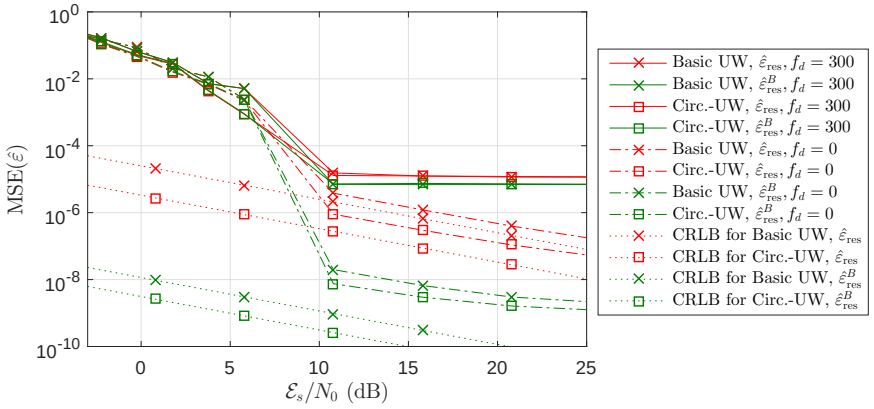


Fig. 4.13: CFO estimation performance for basic UW vs. Circ.-UW designs.

4.4 Channel Estimation

In the following, we elaborate on the problem of channel estimation for MIMO UW sequences under the assumption of perfect time and frequency synchronization.

4.4.1 MIMO UW-based LMMSE CE

In a block-fading situation, the receive signal associated to the last IBI-free $(N_{\text{u}} - L)$ samples of the UW $\vec{x}_{\text{u},i}$—defined in (4.9)—at antenna q and block b is given by

$$\vec{y}_{p,q,b} = [\boldsymbol{h}_{\text{Tpl},q,0,b}, \cdots, \boldsymbol{h}_{\text{Tpl},q,\mathfrak{I}-1,b}][\vec{x}^T_{p,0}, \cdots, \vec{x}^T_{p,\mathfrak{I}-1}]^T + \vec{w}_q, \tag{4.32}$$

where, $\vec{x}_{p,i} = (\vec{x}_{\text{u},i})_{L:N_{\text{u}}-1}$, and $\boldsymbol{h}_{\text{Tpl},q,i,b} \in \mathbb{C}^{(N_{\text{u}}-L)\times(N_{\text{u}}-L)}$ is a lower triangular Toeplitz channel matrix with $\vec{h}_{q,i,b}$—denoting the parameters vector of L-length CIR—on its first column. Rearranging the matrix-vector notations, we get

$$\vec{y}_{p,q,b} = \boldsymbol{X}_p \vec{h}_{q,b} + \vec{w}_q, \tag{4.33}$$

where $\boldsymbol{X}_p = [\boldsymbol{X}_{p,1}, \cdots, \boldsymbol{X}_{p,\mathfrak{I}}] \in \mathbb{C}^{(N_u-L)\times \mathfrak{I}L}$ is the observation matrix with $\boldsymbol{X}_p \in \mathbb{C}^{(N_{\text{u}}-L)\times L}$ being the last $(N_{\text{u}} - L)$ rows and first L columns of $\boldsymbol{X}_{u,i}$. Here, $\boldsymbol{X}_{u,i}$ is a lower triangular Toeplitz matrix with $\vec{x}_{\text{u},i}$—defined in (4.9)—on its first column. $\vec{h}_{q,b} = [\vec{h}^T_{q,i,b}]^T_{i=1:\mathfrak{I}} \in \mathbb{C}^{\mathfrak{I}L}$ denotes the parameters vector of $\mathfrak{I}$ channel impulse responses of length L. $\vec{w}_q \in \mathbb{C}^{N_{\text{u}}-L}$ denotes the AWGN process of variance σ^2_w. The LMMSE estimation of $\vec{h}_{q,b}$ and its covariance follow [Kay93]

$$\hat{\vec{h}}_{q,b} = (\sigma^2_w \boldsymbol{\Sigma}^{-1}_{hh} + \boldsymbol{X}^H_p \boldsymbol{X}_p)^{-1} \boldsymbol{X}^H_p \vec{y}_{p,q,b}, \tag{4.34}$$

$$\hat{\boldsymbol{\Sigma}}_{hh} = \boldsymbol{\Sigma}_{hh} - (\sigma^2_w \boldsymbol{\Sigma}^{-1}_{hh} + \boldsymbol{X}^H_p \boldsymbol{X}_p)^{-1} \boldsymbol{X}^H_p \boldsymbol{X}_p \boldsymbol{\Sigma}_{hh}, \tag{4.35}$$

respectively. Here, $\boldsymbol{\Sigma}_{hh} = \text{diag}(\vec{1}_{\mathfrak{I}} \otimes \vec{\mathcal{P}})$ wherein, $\vec{\mathcal{P}} \in \mathbb{C}^L$ denotes the single PDP between all Tx-Rx antenna pairs in a centralized MIMO system, and it is assumed to be perfectly

known at the receiver. Once, an estimate $\hat{\vec{h}}_{q,b} = [\hat{\vec{h}}^T_{q,i,b}]^T_{i=1:\mathfrak{I}}$ for each of both UW sequences over the *UW-Payload-UW* block is obtained, the two estimations are being averaged and fed to the equalization unit, i.e.

$$\hat{\vec{h}}_{q,i,b,\mathrm{B}} = \frac{1}{2}(\hat{\vec{h}}_{q,i,b-1} + \hat{\vec{h}}_{q,i,b}). \tag{4.36}$$

Note that the linear model (4.33) is only used to derive the LMMSE CE under a block-fading channel assumption. The simulations of a realistic time-varying scenario will adopt (4.13) with $\varepsilon = 0$, $\vartheta = 0$ (i.e. perfect synchronization assumption).

An advantage of the above LMMSE estimation is that the observation signal $\vec{x}_{p,i}$ is interference-free from IBI perspective. However, it also has a drawback due to the partial selection of $\vec{x}_{\mathrm{u},i}$. The sequence $\vec{x}_{\mathrm{u},i}$ has a constant frequency magnitude which is an optimal sequence for channel estimation from MSE sense. By selecting the last $N_\mathrm{u} - L$ samples of $\vec{x}_{\mathrm{u},i}$, the signal magnitude in frequency domain decreases at near edge subcarriers. Therefore, in case of a full subcarrier allocation, it would be recommended to use the whole $\vec{x}_{\mathrm{u},i}$ sequence for channel estimation and apply iterative interference cancellation in analogous way as in [EMCF19], although, applying such an approach goes beyond the scope of this thesis.

4.4.2 Adaptive Filtering

Considering that $b \leq B$ number of *Payload-UW* blocks have been transmitted after the preamble, and thus, $b+1$ channel estimations are available, we apply a Wiener-Hopf filter analogously as in Sec. 3.5 to improve the channel estimation at block b, although, with the main difference that the Wiener filter is *causal* and it is being applied solely to use the already estimated CIRs to obtain an improved estimation of the channel at the center[5] of the payload block b.

Fig. 4.14 illustrates the filtering approach, in which, the $b+1$ channel estimations

$$\boldsymbol{h}_{q,i} = [\hat{\vec{h}}_{q,i,0}, \hat{\vec{h}}_{q,i,1}, \cdots, \hat{\vec{h}}_{q,i,b}] \tag{4.37}$$

are used to calculate the CIR at time sample n_d in the center of the b-th payload block. Once the estimations have been filtered, the resulting vector $(\hat{\vec{h}}_{q,i,b,\mathrm{W}})_{\ell=0:L-1}$ is being used for the equalization of the corresponding block. The adaptive filtering based on *Wiener-Hopf* Approach [Kay93],[6] follows:

Let $\mathbf{R}_h \in \mathbb{C}^{(b+2)\times(b+2)}$ be the temporal auto-correlation matrix of the channel tap ℓ at the respective time samples, i.e.,

$$[\mathbf{R}_h]_{n.n'} = \mathbb{E}(h_n[\ell]^* h_n[\ell]) \text{ for } n, n' \in \{n_{u,0}, \cdots, n_{u,b-1}, n_d, n_{u,b}\}, \tag{4.38}$$

[5] With the assumption that the channel varies within the block duration.

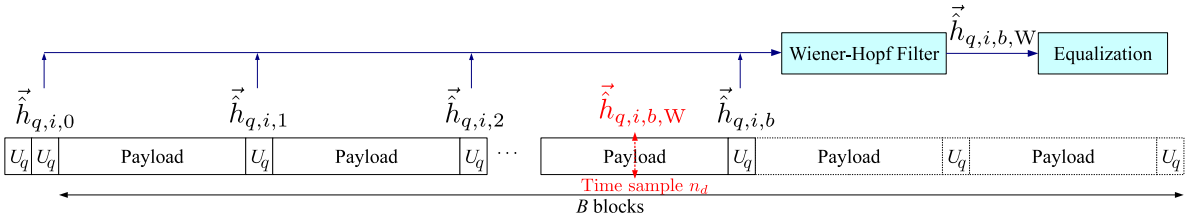


Fig. 4.14: Adaptive filtering of the estimated CIRs for $b \leq B$ blocks.

where the sample indexes n_d and $n_{u,b}$ are the sample time indexes associated to the center of the b-th payload and UW, respectively. The filtered CIR's ℓ-th tap at block b is given by

$$\hat{h}_{q,i,b,\mathrm{W}}[\ell] = [\vec{R}_h]_{\mathcal{U},b}^H (\mathbf{R}_h^{(\mathcal{U})} + \hat{\sigma}_\ell^2 \boldsymbol{I}_{(b+1)})^{-1} \vec{\hat{h}}_{q,i,\ell}, \tag{4.39}$$

where $\vec{\hat{h}}_{q,i,\ell}$ is the transposed ℓ-th row of $\boldsymbol{h}_{q,i}$, $[\vec{R}_h]_{\mathcal{U},b} \subset \mathbf{R}_h^{(\mathcal{U},d)}$ is the b-th column of $\mathbf{R}_h^{(\mathcal{U},d)}$. The sub-matrix $\mathbf{R}_h^{(\mathcal{U},d)} \subset \mathbf{R}_h$ contains the rows of $\mathbf{R}_h$ associated to the UW slots (i.e. without the row index associated to n_d). $\mathbf{R}_h^{(\mathcal{U})} \subset \mathbf{R}_h^{(\mathcal{U},d)}$ contains the columns of $\mathbf{R}_h^{(\mathcal{U},d)}$ associated to the UW slots. $\hat{\sigma}_\ell^2$ denotes the ℓ-th diagonal element of $\widehat{\boldsymbol{\Sigma}}_{hh}$. Moreover, the covariance of the filtered CE associated to tap ℓ is given by

$$(\widehat{\boldsymbol{\Sigma}}_{hh,\ell,\mathrm{W}})_{q,i} = \left(\mathbf{R}_h - \mathbf{R}_h^{(\mathcal{U},d)^H} (\mathbf{R}_h^{(\mathcal{U})} + \hat{\sigma}_\ell^2 \boldsymbol{I}_{(b+1)})^{-1} \mathbf{R}_h^{(\mathcal{U},d)} \right)_{q,i}. \tag{4.40}$$

Therefore, at block b, the CE covariance becomes

$$(\widehat{\boldsymbol{\Sigma}}_{hh,\mathrm{W}}[b])_{q,i} = \left(\mathrm{diag}([(\widehat{\boldsymbol{\Sigma}}_{hh,\ell,\mathrm{W}}[b,b])_{\ell=0:L-1}]^T) \right)_{q,i}. \tag{4.41}$$

4.4.3 Circular UW Transmission

Recall the Circ.-UW frames that were discussed in Sec. 4.3.1, in which, via (4.12) the UW sequences transmitted from each Tx antenna are circularly rotated before and after the payload blocks. For instance, Fig. 4.15 illustrates the frame structure for $\mathfrak{I} = 4$ Tx antennas. As in Fig. 4.15, for each Tx antenna, the UW sequences around the payload blocks are circularly rotated by two antenna indexes.

In this case, the LMMSE estimate of the channels shall be computed jointly for the UW sequences before and after the payload blocks. Thus,

$$\vec{\hat{h}}_q^{\mathrm{Circ.-UW}} = (\sigma_w^2 \boldsymbol{\Sigma}_{hh}^{-1} + \boldsymbol{X}_{p,2b}^H \boldsymbol{X}_{p,2b})^{-1} \boldsymbol{X}_{p,2b}^H \, \vec{y}_{p,q,2b}, \tag{4.42}$$

$$\widehat{\boldsymbol{\Sigma}}_{hh}^{\mathrm{Circ.-UW}} = \boldsymbol{\Sigma}_{hh} - (\sigma_w^2 \boldsymbol{\Sigma}_{hh}^{-1} + \boldsymbol{X}_{p,2b}^H \boldsymbol{X}_{p,2b})^{-1} \boldsymbol{X}_{p,2b}^H \boldsymbol{X}_{p,2b} \boldsymbol{\Sigma}_{hh}, \tag{4.43}$$

with

$$\boldsymbol{X}_{p,2b} = \begin{bmatrix} \boldsymbol{X}_{p,1}, \cdots, \boldsymbol{X}_{p,\frac{\mathfrak{I}}{2}}, \boldsymbol{X}_{p,\frac{\mathfrak{I}}{2}+1}, \cdots, \boldsymbol{X}_{p,\mathfrak{I}} \\ \boldsymbol{X}_{p,\frac{\mathfrak{I}}{2}+1}, \cdots, \boldsymbol{X}_{p,\mathfrak{I}}, \boldsymbol{X}_{p,1}, \cdots, \boldsymbol{X}_{p,\frac{\mathfrak{I}}{2}} \end{bmatrix}, \tag{4.44}$$

and $\vec{y}_{p,q,2b} = [(\vec{y}_u)_{L:N-1,b-1}^T, (\vec{y}_u)_{L:N-1,b}^T]^T$ being the concatenation of two IBI-free UW sequences surrounding the payload block.

Tx. 1:	U_0	U_0	Payload	U_2	Payload	U_0	Payload	U_2
Tx. 2:	U_1	U_1	Payload	U_3	Payload	U_1	Payload	U_3
Tx. 3:	U_2	U_2	Payload	U_0	Payload	U_2	Payload	U_0
Tx. 4:	U_3	U_3	Payload	U_1	Payload	U_3	Payload	U_1

Fig. 4.15: Example of the Circ.-UW frames for $\mathfrak{I} = 4$ Tx antennas.

4.4.4 Simulation Results

Consider the same simulation scenario of Sec. 4.3.5 (i.e. 4×4 MIMO, exponentially decaying PDP with $L = 9$ from 0 dB to -20 dB, $F_s = 1.92$MHz, $N_u = 49$, $N_d = 256$) with the only difference that here, in order to solely focus on channel estimation performance, we assume perfect time and frequency synchronization. We also consider $B = 50$ to further emphasize the influence of Wiener-Hopf filters. The time-varying channel estimation is simulated by considering that each tap of the CIR being temporally correlated according to [Cla68] $[\mathbf{R}_h]_{n,n'} = J_0(2\pi(n-n')\nu_d)$, where $\nu_d = f_d/F_s$ and $J_0(\cdot)$ denotes the zeroth order Bessel function of the first kind. The number of observations of the UW sequences would be $N_u - L = 40$ time domain samples (i.e. $\vec{y}_{p,q} \in \mathbb{C}^{40}$). The CP-OFDM system is configured to adopt $N_{\text{ofdm}}^{\text{ON}} = 240$ active subcarriers out of $N_{\text{ofdm}} = 256$, in which 40 equispaced of them are pilot subcarriers. The OFDM pilot subcarriers are being generated via PN-sequences mapped into QPSK constellations. The MSE of the channel estimation is calculated by comparing the DFT domain of the estimated channel responses with respect to the true channel realizations at the center of each payload block.

Fig. 4.16 shows the MSE of the channel estimation for the proposed UW-based frames vs. CP-OFDM system. Comparing the pilot-aided CE of CP-OFDM with blockwise UW-based CE, i.e. (4.36), at $f_d = 300$Hz, we observe that at low SNR the performances are the same, however, at high SNR values CP-OFDM pilots suffer from nearly one order of magnitude larger error floor which is due to the channel variations within the OFDM block duration. On the other hand, Blockwise UW-based CE at $f_d = 300$ follows its $f_d = 0$ bound with a very close gap for SNR values < 25 dB, and at high SNR values, it has a smaller error floor than the CP-OFDM pilots, because the number of channel realizations within the UW sequence is smaller than that of CP-OFDM block duration (i.e. $N_u - L = 40$ channel realizations vs. $N_{\text{ofdm}} = 256$ for CP-OFDM). In other words, the energy of the UW sequences are concentrated into $(N_u - L)$-length time slot (i.e. 3.5% of the channel coherence time), whilst in CP-OFDM, the energy of the pilots are distributed over $N_{\text{ofdm}} = 256$ samples (i.e. 22.3% of the channel coherence time). We also note that the blockwise UW-based CE at $f_d = 0$ that averages the CE of two estimates surrounding a payload block follows its theoretical bound of joint LMMSE of two UW sequences with marginal degradation. Comparing the blockwise UW-based channel estimations with their corresponding Wiener-filtered cases, i.e. (4.39), one may note that at low mobility, i.e. $f_d = 0$, the Wiener filter provides nearly 5 dB gain with respect to its blockwise channel estimation at MSE=10^{-3}, although, at high mobility scenario, i.e. $f_d = 300$, the gain

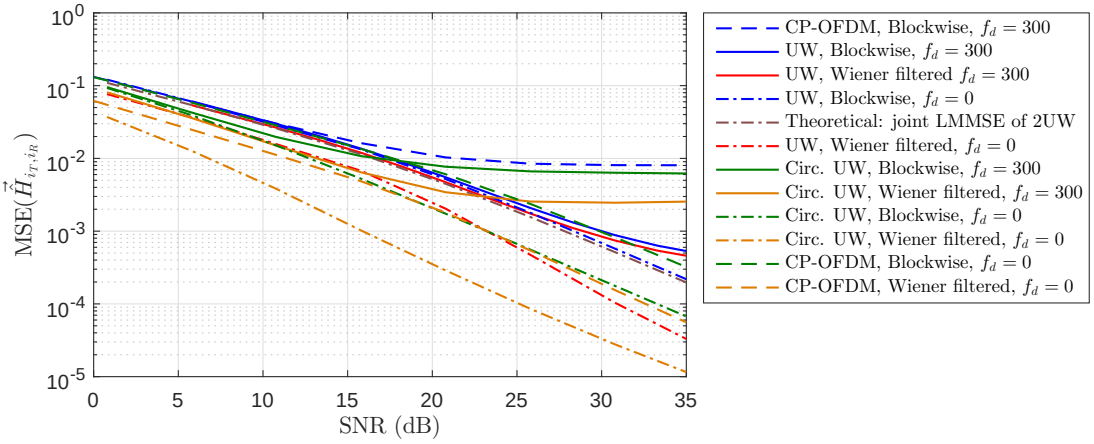


Fig. 4.16: Channel estimation performance of MIMO UW based systems.

is marginal, because the matrix $\mathbf{R}_h$ defined in (4.38) considers less temporal correlation of the channel estimations. Moving on to the channel estimation performances of Circ.-UW sequences, we observe that at low mobility cases, the blockwise Circ.-UW CE outperforms the normal UW-based CE as well as blockwise CP-OFDM with nearly 5 and 6.6 dB gain, respectively, whereas, the Wiener-filtered Circ.-UW gains more than 7.3 dB with respect to its normal Wiener-filtered UW-based CE at $f_d = 0$. The reason for such performance gain shall be explained by comparing the condition number of the observation matrices $\text{Cond}(\boldsymbol{X}_{p,2b}) = 4.54 < \text{Cond}(\boldsymbol{X}_p) = 13.4 < \text{Cond}(\boldsymbol{X}_p^{\text{OFDM pilots}}) = 33.78$. Observing the performance of Circ.-UW-based channel estimation at high mobility $f_d = 300$, the error floor at high SNR values increases with respect to the normal UW-based CE because of the same reason that happens to CP-OFDM, i.e. the two UW sequences observe $2(N_u - L) = 80$ channel realizations. The Wiener-filtered Circ.-UW-based CE at $f_d = 300$ does not provide significant gain at low SNR values (due to small temporal correlation of the estimates), whereas, at high SNR values, it reduces the channel estimation error floor.

4.5 Equalization with Imperfect Channel Knowledge

The state-of-the-art of EQ techniques for UW based multi-carrier systems consider using the circularity of the channel over the *Payload-UW* block and equalizes the payload together with the UW sequence [DdL17], [HOH11]. Since our frame design in Sec. 4.1 does not consider the UW sequence within the FFT size, the payload block itself does not experience a circular channel condition. On the other hand, considering an odd size FFT not only increases the complexity, but also collects more noise samples in DFT domain i.e. N time-domain noise samples instead of N_d. In the following, we derive a UW-free joint channel-equalization-and-demodulation for a generalized flexible waveform that can emulate any orthogonal or non-orthogonal single/multi-carrier.

4.5.1 UW-Free Equalization

While in a block-fading system, the Tx signal experiences a linear convolution with the CIR, different techniques—as illustrated in Fig. 4.17—facilitate a cyclic signal observation at the receiver. In Fig. 4.17b, one can see that in CP-based systems, since $\vec{x}_{cp_1}$ is identical to the last L samples of $\vec{x}_d$, the receiver can deduce that $\vec{x}_d$ experiences a circular channel matrix. Similarly, in conventional UW-based systems (see 4.17c), the receiver can also emulate a cyclic signal condition for the transmitted $\vec{x} = [\vec{x}_d^T, \vec{x}_u^T]^T$ sequence. For a UW-free equalization, i.e. 4.17d, we reconstruct a circular channel condition by adding the first L samples of the UW to the first L samples of the payload block as in Fig. 4.18. In this case, while the last L samples of the payload has leaked into its following UW sequence, i.e. due to the linear convolution of the CIR with the *UW-Payload-UW* block, taking the first L samples of UW and adding it to the beginning of the payload sequence makes $\vec{y}_{\text{cyclic},q} \in \mathbb{C}^{N_d}$ to experience a circular convolution of the transmitted payload with the CIR. This approach also yields some interference due to the UW sequence. However, since the estimated CIR and the UW sequence are known at this stage, one can simply remove the UW influence. Therefore, we define the UW-free cyclic observed signal as

$$\vec{y}_{\text{cyclic},q} \triangleq \vec{y}_{d,q} + [(\vec{y}_{u,q})_{[0:L-1]}^T \ \vec{0}_{N_d-L}^T]^T, \tag{4.45}$$

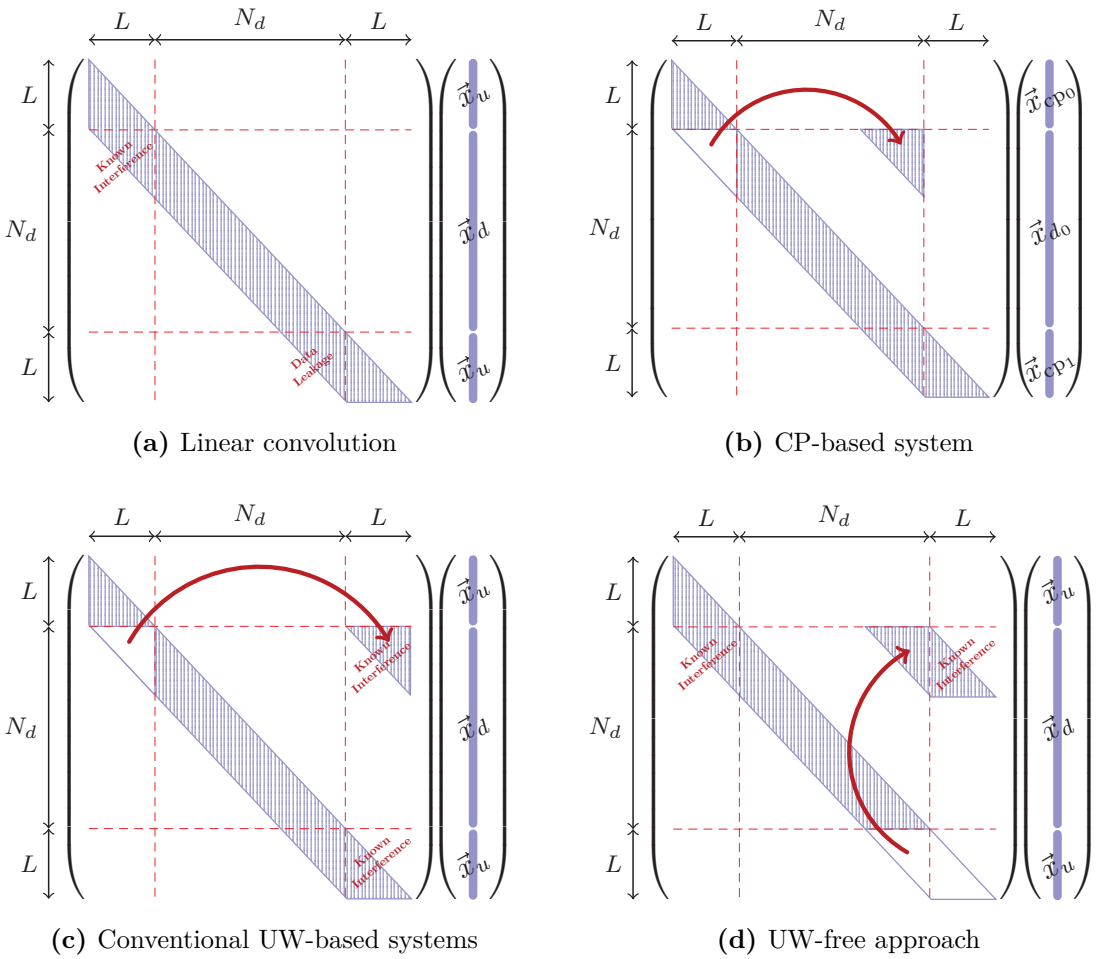
where $(\vec{y}_{u,q})_{[0:L-1]}$ denotes the first L samples of the received UW sequence after the payload, and $\vec{y}_{d,q} \in \mathbb{C}^{N_d}$ denotes the received payload sequence at block b. Further details concerning the observed signal $\vec{y}_{\text{cyclic},q}$ can be found in Appendix A.7 where we provide an algebraic proof that $\vec{y}_{\text{cyclic},q}$ experiences a circular channel matrix. In (4.45) and the rest of expressions in this section, the subscript b is ignored for brevity. Also note that in (4.45), the amount of noise collection would be identical to the conventional approaches, i.e. N time-domain noise samples, but it provides the advantage of adapting radix-2 FFT-size for the payload size. Additionally, in the above approach since the UW slot is no longer part of the FFT-size, a low complexity equalization approach that relies on the structure of the modulation matrix $\boldsymbol{\mathcal{M}}$ can be directly applied. For instance, an OFDM signal (i.e. $\boldsymbol{\mathcal{M}}_{\text{ofdm}} = \boldsymbol{F}_{N_d}^H$) would have an almost diagonal effective channel, whereas, the circular pulse-shaping GFDM signal with $\boldsymbol{\mathcal{M}} = \boldsymbol{A}$ would have a banded matrix.

Once a cyclic receive signal via (4.45) has been obtained, we remove the UW influence and thus, the receive payload sequence in DFT domain yields

$$\vec{Y}_{\text{cyclic},d,q} = \boldsymbol{F}_{N_d}\vec{y}_{\text{cyclic},q} - \boldsymbol{F}_{N_d}\sum_{i=0}^{\mathfrak{I}}[(\boldsymbol{F}_{N_\text{u}}^H\hat{\boldsymbol{H}}_{q,i,N_\text{u}}\boldsymbol{F}_{N_\text{u}}\vec{x}_{u,i})_{[0:L-1]}^T \ \vec{0}_{N_d-L}^T]^T, \tag{4.46}$$

where $\hat{\boldsymbol{H}}_{q,i,N_\text{u}} = \text{diag}(\sqrt{N_\text{u}}\boldsymbol{F}_{N_\text{u},L}\hat{\vec{h}}_{q,i})$. In this case, the receive signal $\vec{Y}_{\text{cyclic},d} = [(\vec{Y}_{\text{cyclic},d,q})_{q=1:\mathfrak{Q}}^T]^T$ can be decomposed into the following linear model

$$\begin{aligned}\vec{Y}_{\text{cyclic},d} = {} & \hat{\boldsymbol{H}}_{N_d}\overline{\boldsymbol{F}}_{N_d}\overline{\boldsymbol{\mathcal{M}}}\vec{d} + \tilde{\boldsymbol{H}}_{N_d}\overline{\boldsymbol{F}}_{N_d}\overline{\boldsymbol{\mathcal{M}}}\vec{d} \\ & + \overline{\overline{\boldsymbol{F}}}_{N_d}\left[\begin{pmatrix}(\overline{\boldsymbol{F}}_{N_\text{u}}^H\tilde{\boldsymbol{H}}_{N_\text{u}}\overline{\boldsymbol{F}}_{N_\text{u}}\vec{x}_u)_{0:L-1} \\ \vec{0}_{N_d-L}\end{pmatrix}_{q=1:\mathfrak{Q}}^T\right]^T + \overline{\overline{\boldsymbol{F}}}_{N_d}\left[\begin{pmatrix}(\vec{w}_u)_{0:L-1} \\ \vec{0}_{N_d-L}\end{pmatrix}_{q=1:\mathfrak{Q}}^T\right]^T + \overline{\overline{\boldsymbol{F}}}_{N_d}\vec{w}_d.\end{aligned} \tag{4.47}$$



(a) Linear convolution

(b) CP-based system

(c) Conventional UW-based systems

(d) UW-free approach

Fig. 4.17: The channel matrix $\boldsymbol{\mathcal{H}}$ that different receiver types would experience in the linear model $\vec{y} = \boldsymbol{\mathcal{H}}\vec{x} + \vec{w}$.

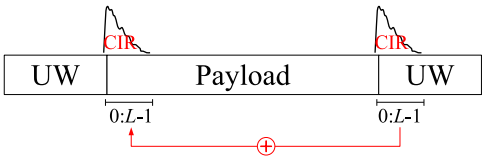


Fig. 4.18: Emulating a circular CIR for the N_d FFT-size payload sequence at the receiver.

Here $\overline{\boldsymbol{\mathcal{M}}} = \boldsymbol{I}_{\mathfrak{I}} \otimes \boldsymbol{\mathcal{M}}$, $\overline{\boldsymbol{F}}_{N_d} = \boldsymbol{I}_{\mathfrak{I}} \otimes \boldsymbol{F}_{N_d}$, $\overline{\overline{\boldsymbol{F}}}_{N_d} = \boldsymbol{I}_{\mathfrak{Q}} \otimes \boldsymbol{F}_{N_d}$, $\overline{\boldsymbol{F}}_{N_\mathrm{u}} = \boldsymbol{I}_{\mathfrak{I}} \otimes \boldsymbol{F}_{N_\mathrm{u}}$, $\vec{x}_u = [(\vec{x}_{u,i})_{i=1:\mathfrak{I}}^T]^T$, $\vec{w}_u \in \mathbb{C}^{N_\mathrm{u}}$ denotes the AWGN process within the UW sequence, whereas $\vec{w}_d \in \mathbb{C}^{\mathfrak{Q}N_d}$ denotes the noise samples associated to the payload sequences of $\mathfrak{Q}$ antennas. Moreover, $\tilde{\boldsymbol{H}}_{N_d} = \boldsymbol{H}_{N_d} - \hat{\boldsymbol{H}}_{N_d}$ denotes the error of the channel estimation that is uncorrelated to $\vec{d}$. The matrix $\hat{\boldsymbol{H}}_{N_d} \in \mathbb{C}^{\mathfrak{Q}N_d \times \mathfrak{I}N_d}$ is MIMO version of $\mathfrak{Q} \times \mathfrak{I}$ matrices $\hat{\boldsymbol{H}}_{q,i,N_d} = \mathrm{diag}(\sqrt{N_d}\boldsymbol{F}_{N_d,L}\hat{\vec{h}}_{q,i})$, (Likewise $\boldsymbol{H}_{N_d}, \tilde{\boldsymbol{H}}_{N_d}$ and $\tilde{\boldsymbol{H}}_{N_\mathrm{u}}$ with respective dimensions). Considering the observation matrix $\hat{\boldsymbol{H}}_{N_d}\overline{\boldsymbol{F}}_{N_d}\overline{\boldsymbol{\mathcal{M}}}$ in (4.47), the CWCU LMMSE [HL14] equalizer of d with imperfect channel knowledge with covariance $(\hat{\boldsymbol{\Sigma}}_{hh,\mathrm{W}})_{q,i}$—

defined in (4.41)—is given by

$$\vec{\hat{d}} = \frac{\boldsymbol{\Sigma}_{dY}(\boldsymbol{\Sigma}_{d\hat{H}} + \boldsymbol{\Sigma}_{d\tilde{H}} + \boldsymbol{\Sigma}_{x_u\tilde{H}} + \boldsymbol{\Sigma}_{w_u} + \sigma_w^2\boldsymbol{I}_{\mathfrak{Q}N_d})^{-1}\vec{Y}_d}{\mathrm{diag}^{-1}(\widetilde{\boldsymbol{\Sigma}}_{dd})}, \tag{4.48}$$

$$\widehat{\boldsymbol{\Sigma}}_{dd} = \boldsymbol{\Sigma}_{dd} - \widetilde{\boldsymbol{\Sigma}}_{dd}, \tag{4.49}$$

herein, we calculate each term as

$$\boldsymbol{\Sigma}_{d\hat{H}} = \hat{\boldsymbol{H}}\overline{\boldsymbol{F}}_{N_d}\overline{\boldsymbol{\mathcal{M}}}\boldsymbol{\Sigma}_{dd}\overline{\boldsymbol{\mathcal{M}}}^H\overline{\boldsymbol{F}}_{N_d}^H\hat{\boldsymbol{H}}^H, \tag{4.49a}$$

$$\boldsymbol{\Sigma}_{d\tilde{H}} = \boldsymbol{I}_{\mathfrak{Q}} \otimes \sum_{i=1}^{\mathfrak{I}} \boldsymbol{F}_{N_d}\boldsymbol{\mathcal{M}}\boldsymbol{\Sigma}_{dd}\boldsymbol{\mathcal{M}}^H\boldsymbol{F}_{N_d}^H \circ \widehat{\boldsymbol{\Sigma}}_{\mathrm{HH},i}, \tag{4.49b}$$

$$\boldsymbol{\Sigma}_{x_u\tilde{H}} = \boldsymbol{I}_{\mathfrak{Q}} \otimes \boldsymbol{F}_{N_d}\begin{bmatrix}\left[\boldsymbol{F}_{N_\mathrm{u}}^H\left(\sum_{i=1}^{\mathfrak{I}}\boldsymbol{\Sigma}_{X_uX_u,i}\circ\widehat{\boldsymbol{\Sigma}}_{\mathrm{HH},i}\right)\boldsymbol{F}_{N_\mathrm{u}}\right]_{\substack{0:L-1,\\0:L-1}} & \boldsymbol{0}_{L\times N_d-L}\\ \boldsymbol{0}_{N_d-L\times L} & \boldsymbol{0}_{N_d-L\times N_d-L}\end{bmatrix}\boldsymbol{F}_{N_d}^H \tag{4.49c}$$

$$\boldsymbol{\Sigma}_{w_u} = \boldsymbol{I}_{\mathfrak{Q}} \otimes \boldsymbol{F}_{N_d}\begin{bmatrix}\sigma_w^2\boldsymbol{I}_L & \boldsymbol{0}_{L\times N_d-L}\\ \boldsymbol{0}_{N_d-L\times L} & \boldsymbol{0}_{N_d-L\times N_d-L}\end{bmatrix}\boldsymbol{F}_{N_d}^H, \tag{4.49d}$$

$$\boldsymbol{\Sigma}_{dY} = \boldsymbol{\Sigma}_{dd}\overline{\boldsymbol{\mathcal{M}}}^H\overline{\boldsymbol{F}}_{N_d}^H\hat{\boldsymbol{H}}^H, \tag{4.49e}$$

$$\widetilde{\boldsymbol{\Sigma}}_{dd} = \boldsymbol{\Sigma}_{dY}(\boldsymbol{\Sigma}_{d\hat{H}} + \boldsymbol{\Sigma}_{d\tilde{H}} + \boldsymbol{\Sigma}_{x_u\tilde{H}} + \boldsymbol{\Sigma}_{w_u} + \sigma_w^2\boldsymbol{I}_{\mathfrak{Q}N_d})^{-1}\boldsymbol{\Sigma}_{dY}^H, \tag{4.49f}$$

where $\boldsymbol{\Sigma}_{dd} = \mathbb{E}[\vec{d}\vec{d}^H] = \boldsymbol{I}_{\mathfrak{I}N_d}$ provided that the elements of $\vec{d}$ are i.i.d. standard Gaussian random variables, $\widehat{\boldsymbol{\Sigma}}_{\mathrm{HH},i} = \boldsymbol{F}_{N_\mathrm{u},L}(\widehat{\boldsymbol{\Sigma}}_{hh,W})_{q,i}\boldsymbol{F}_{N_\mathrm{u},L}^H$ and for a centralized MIMO with identical PDP between Tx-Rx antenna pairs, $\widehat{\boldsymbol{\Sigma}}_{\mathrm{HH},i}$ is identical for all $q \in \{1,\cdots,\mathfrak{Q}\}$. Further, $\boldsymbol{\Sigma}_{X_uX_u,i} = \boldsymbol{F}_{N_\mathrm{u}}\vec{x}_{u,i}\vec{x}_{u,i}^H\boldsymbol{F}_{N_\mathrm{u}}^H$. From a complexity point of view, the matrices $\boldsymbol{\Sigma}_{d\hat{H}}$ and $\boldsymbol{\Sigma}_{d\tilde{H}}$ become band diagonal matrices for GFDM with proper permutations. The lower and upper bandwidth of the two matrices is given by $B_l = B_u = \mathfrak{I}(M + N_\alpha)$, where $N_\alpha < M$ is the number of overlapping frequency bins with the neighboring subcarriers and it depends on the filter roll-off factor α. The matrix $\boldsymbol{\Sigma}_{x_u\tilde{H}}$ is a full $\mathfrak{Q}N_d \times \mathfrak{Q}N_d$ matrix which can be ignored by setting $\boldsymbol{\Sigma}_{x_u\tilde{H}} = \boldsymbol{0}_{\mathfrak{Q}N_d\times\mathfrak{Q}N_d}$. Moreover, $\boldsymbol{\Sigma}_{w_u}$ defined in (4.49d) shows that the noise becomes correlated in DFT domain and thus, for the sake of saving complexity, we consider taking its diagonal elements.

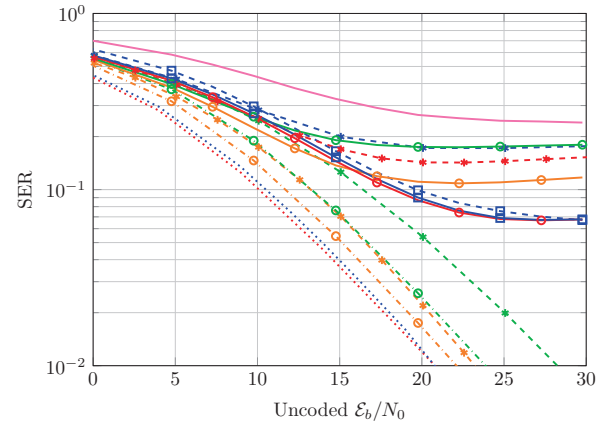
4.5.2 Simulation Results

Given the imperfect channel estimation techniques proposed in Sec. 4.4, we evaluate the transceiver performances in terms of uncoded SER, mutual information (MI), and coded FER via Monte Carlo simulations. Here, we consider again the same simulation parameters of Sec. 4.4.4 (i.e. 4×4 MIMO, exponentially decaying PDP with $L = 9$ from 0 dB to -20 dB, $F_s = 1.92$MHz, $N_\mathrm{u} = 49$, $N_d = 256$) and on top, we consider that the transmitted bits are encoded via *Parallel Concatenated Convolutional Codes* with code-rate $r_\mathrm{code} = 2/3$ and octal generator polynomial (1,15/13), and they are mapped into 16-QAM symbols, i.e. modulation order $\mu = 4$. The receiver employs the CWCU LMMSE

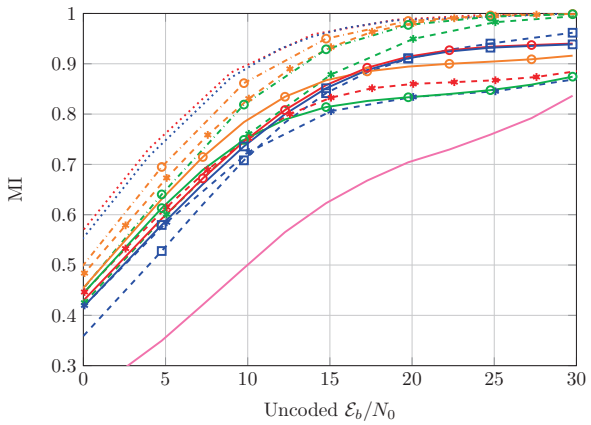
equalization technique and transforms the equalized signal constellations into ML log-likelihoods by taking into account the fact that each element of the received constellations $\vec{\tilde{d}} = \vec{d} + \vec{\tilde{w}}$ subjects to an effective noise enhancement approximately distributed as $\vec{\tilde{w}} \sim \mathcal{N}_C(\vec{0}_{N_d}, \mathrm{diag}(\frac{1}{\mathrm{diag}^{-1}(\widetilde{\mathbf{\Sigma}}_{dd})}) - \sigma_w^2 \vec{1}_{N_d})$ [MZF18]. The SNR per bit is denoted by E_b/N_0 and it includes the gain of modulation and coding (i.e. $\mathcal{E}_b/N_0 = \mathcal{E}_s/N_0 - 10\log_{10}\mu r_{\mathrm{code}}$, where $\mathcal{E}_s/N_0$ denotes SNR per symbol). As benchmarks, in addition to the CP-OFDM pilot transmission (i.e. identically configured as in Sec. 4.4.4), we also consider a conventional approach of UW based transmission for GFDM similar to a MIMO extended version of [DdL17], i.e. the channel estimation of UW slot $b-1$ is used for its following payload block b, and also the estimated channel matrix $\hat{H}$ and the GFDM modulation matrix A are being separately equalized via LMMSE approach and without considering the imperfect CE statistics (i.e. setting $\mathbf{\Sigma}_{d\tilde{H}} = \mathbf{0}_{\mathfrak{Q}N_d \times \mathfrak{Q}N_d}$ and $\mathbf{\Sigma}_{x_u\tilde{H}} = \mathbf{0}_{\mathfrak{Q}N_d \times \mathfrak{Q}N_d}$). Moreover, the conventional approach uses the circularity of the channel over the *Payload-UW* time slot and thus, calculation of $\mathbf{\Sigma}_{w_u}$ is also unnecessary.

Fig. 4.19a compares the SER performance of the proposed approaches with the state-of-the-art equalization techniques at high mobility scenarios. One may notice that the conventional UW-based EQ for GFDM suffers from a large error floor in SER performance due to three major aspects: a) The GFDM signal suffers from self-interference caused by the circular filtering, b) the conventional approach uses the CE of a single UW before the payload sequence, and c) it does not consider the imperfect CE statistics. The CP-OFDM also suffers from large error floor of SER, because at high mobility its channel estimation has the largest error floor, and also the orthogonality of the symbols are being lost due to the ICI caused by Doppler effects. Comparing the performance of UW-GFDM with Circ.-UW-GFDM, one may notice that at high mobility of $f_d = 300$, Circ.-UW-GFDM with blockwise CE approach has larger error floor than normal UW-GFDM with blockwise CE, because its observation considers a longer time duration with respect to the channel coherence time than the normal UW-GFDM. But on the other hand, the Wiener filtering that has a marginal gain in normal UW-GFDM, achieves a better performance gain in Circ.-UW-GFDM. Comparing the performances of CP-OFDM vs. Circ.-UW-GFDM at stationary channel conditions, i.e. $f_d = 0$, we note that Circ.-UW-GFDM with blockwise CE outperforms the blockwise CP-OFDM by almost 3 dB SNR, which is mainly due to the better channel estimation of Circ.-UW-GFDM. Applying Wiener filter on top of them, the gap becomes smaller, although, the Wiener filtered Circ.-UW-GFDM achieves a SER performance that is closest to its Genie-Aided receiver[6]. Note that here the gap between SER of Circ.-UW-GFDM and Genie-Aided receiver is ∼1.19 dB, in which ∼0.79 dB is due to the UW overhead. Comparing the GFDM and OFDM systems with equal channel state information, we observe that Genie-aided receiver of OFDM has a marginal gain compared to its equivalent GFDM system, however, for the receivers with imperfect channel knowledge, i.e. UW-GFDM and UW-OFDM, the frequency diversity is better

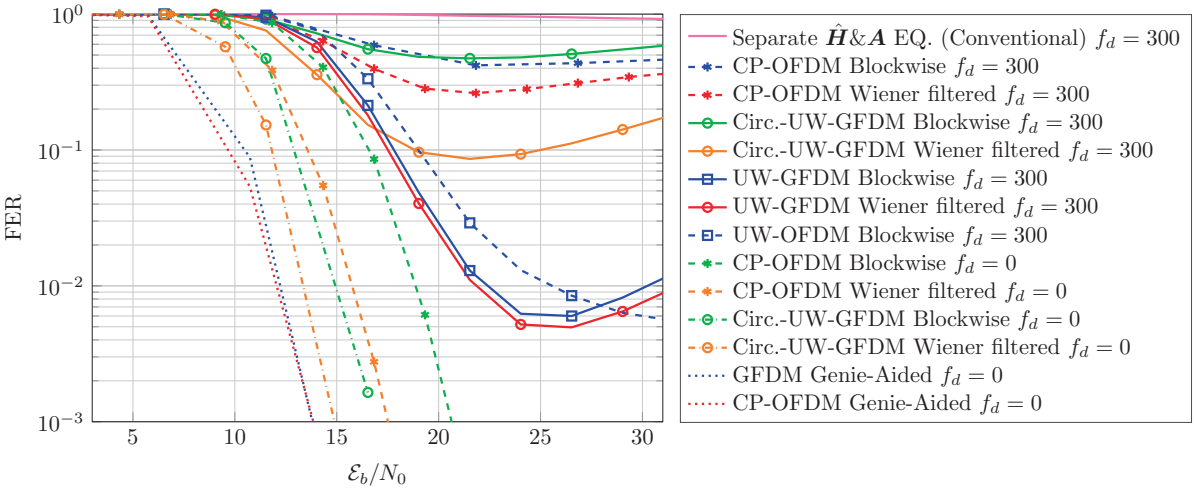
[6] Genie-Aided receiver considers the perfect channel knowledge without the need of UW transmission.



(a) Uncoded SER for 16-QAM symbols



(b) MI of Tx bits and Rx bit LLRs



(c) Coded performance of 16-QAM 2/3 turbo codes with soft demapping.

Fig. 4.19: Equalization and Detection Performance of MIMO UW based systems.

exploited in GFDM with its wider subcarriers. For further details of GFDM vs. OFDM systems with identical perfect and imperfect CSI knowledge, we refer the interested readers to [MZF18].

The demapper performances in terms of mutual information of transmitted bits (after encoder) and received bit LLRs (before decoder) are compared in Fig. 4.19b. As could be expected, the MI curves of different systems follow a similar behavior of their SER performnaces. For instance, the Wiener-filtered Circ.-UW-GFDM at blockfading situation achieves a higher MI than the Wiener-filtered CP-OFDM for the same reason that occurred to its SER performance. In addition, an interesting observation from Fig. 4.19b is the behavior of the curves for interference-limited systems. Here, we note that the MI curves of conventional UW-GFDM, CP-OFDM, blockwise and Wiener-filtered Circ.-UW-GFDM tend to saturate over a nearly horizontal line between SNR values of 15 and 25 dB. However, at very high SNR values, the MI curves start to increase again. We explain this behavior by considering the fact that the interference-statistics occurring due to the

time-varying channels have not been considered in computation of (4.49f). Therefore, as the input SNR of the demapper increases, the output bit LLRs have higher confidence and thus, they result in higher mutual information. Nevertheless, as we see in the following, such high confident LLRs do not necessarily yield correct decoding performance.

Observing the detection performances in terms of FER, i.e. Fig. 4.19c, one can see that the FER performances of different EQ techniques follow a similar behavior that already has been seen in SER and MI performances. However, at very high $\mathcal{E}_b/N_0$ values, we observe that all high mobility curves tend an upward trend, which happens due to the ICI (caused by Doppler and GFDM's self-interference) that has not been considered by the CWCU-LMMSE EQ. and the decoder. As also explained in the MI performances, by ignoring the off-diagonal elements of $\widetilde{\mathbf{\Sigma}}_{dd}$ and also the interference-statistics due to time-varying channels, the bit LLRs become too much confident, which results to have a worse FER at very high $\mathcal{E}_b/N_0$ values. Such an upward trend of the curves shall be solved by employing a time-varying equalizer, and/or an iterative MMSE *parallel-interference-cancellation* detection technique, e.g. [MZF18].

4.6 Summary

In this chapter, we have proposed a new UW-based MIMO frame design for multi-carrier transmission over frequency selective and time-variant channels. Through overhead and efficiency analysis, we have showed that the proposed UW-based system achieves a higher transmission efficiency than conventional pilot-aided CP-OFDM systems. Employing orthogonal UW sequences for a non-orthogonal MIMO system, we have shown that the proposed UW-GFDM can maintain the spectral properties of CP-GFDM with slight degradation. We have also derived the time and frequency synchronization metrics for the UW based MIMO systems and we have validated the robustness of our approach in doubly-dispersive channels. Simulation results show that the proposed approach achieves a larger probability of successful timing synchronization with respect to the conventional MIMO synchronization techniques, while it also achieves nearly 7 dB performance gain in CFO estimation. Deriving the LMMSE estimation of the channels and their corresponding Wiener-Hopf filters, we have showed that the proposed UW-based frame design achieves nearly one order of magnitude higher estimation accuracy when compared to the pilot-aided CP-OFDM system. Proposing a frame design that circularly rotates the transmit UW sequences over the Tx. antennas, we have showed that in low mobility conditions, the Circ.-UW-based system, achieves the highest estimation accuracy. Deriving a UW-free equalization technique, we have showed that the UW sequences does not necessarily need to be placed within the FFT-size. In addition, we have observed that the proposed UW-based transmission scheme achieves a good detection performance in highly doubly-dispersive channel conditions, that CP-OFDM pilots are unable to address.

Chapter 5

Conclusions and Perspectives

5.1 Main Outcomes in Short

In conclusion, this thesis made a thorough analysis of various channel estimation techniques of MIMO multi-carriers in frequency selective and time varying channels. Starting from time invariant conditions for MIMO non-orthogonal waveforms in Sec. 3.1, we comprehensively studied the interference-limited pilot-aided channel estimation. The major outcomes of the first three sections of chapter 3 is that,

- ▷ At low SNR regions where the noise floor is dominant with respect to the interference floor, non-orthogonal pilot-aided CE achieves almost the same performance of channel estimation in OFDM. At high SNR regions, the self-interference terms dominate the error due to the noise enhancement and thus, it necessitates further considerations on the pilots pattern. In Sec. 3.1, we observed that one approach for suppressing the interference statistics is to inject more pilot subcarriers at the cost of increasing the pilots overhead.
- ▷ Keeping the pilot overhead to remain identical, in Sec. 3.2 we showed that rectangular grid-pilot pattern better suppresses the interference error floor than the block-type pilots. However, due to having fewer number of pilot-subcarriers, fewer number of channel taps can be estimated with such pilot pattern.
- ▷ Via modifications on frequency domain modulation of the signal or proper precoding techniques, and for the sake of clear pilot observation, we proposed an approach to insert orthogonal pilots into a non-orthogonal waveform. Analyzing the resulting signal characteristics, we observed that interference-free pilot insertion (IFPI) GFDM achieves almost the same signal characteristics of basic GFDM (see Sec. 3.3). Overall, the complexity of data transmission for IFPI GFDM is smaller than that of basic GFDM, because the channel estimation procedure in IFPI GFDM does not need to handle interference statistics. Moreover, the CE in IFPI GFDM achieves identical estimation quality as in OFDM, because of clear pilot observation.

Extending the non-orthogonal CE techniques to an iterative approach based on LMMSE-PIC for close-to-optimal joint channel estimation and equalization of signal constellations, and digging into its complexity analysis, we conclude that:

- ▷ In case of employing a robust MCS, e.g. QPSK 1/3, the channel estimation based on CP-SLMMSE is most suitable, because by using the pilots' information from CP it achieves the sufficient estimation quality for the decoder while it also has a lower complexity compared to LMMSE and CP-LMMSE approaches.
- ▷ If an intermediate MCS, e.g. 16-QAM is being used, channel estimation based on CP-LMMSE fits best, because it handles the interference terms more accurately compared to sequential LMMSE approaches, and it also has a lower complexity with respect to iterative LMMSE-PIC approaches.
- ▷ For high-throughput data transmission where complexity is less noteworthy, iterative CE based on CP-LMMSE-PIC approach achieves the sufficient estimation quality for employing high MCS orders such as 64-QAM 3/4. At the cost of higher complexity, the channel estimation via CP-LMMSE-PIC of GFDM achieves 2.4 dB higher quality with respect to OFDM.
- ▷ In time-varying channel conditions, the information of CE over multiple blocks can be linearly combined via the Wiener filtering approach presented in Sec. 3.5, in order to suppress the error floor caused by the ICI due to Doppler effects.

In Chapter 4, by considering a scenario that the wireless channel is extremely doubly-dispersive, we extended the SISO unique word design concept to MIMO applications. There,

- ▷ We showed that in extreme channel conditions where CP-OFDM resources become congested by pilots, the transmission can be handled more efficiently via MIMO UW frame designs. If the channel is frequency-selective (e.g. 4.7µs delay length, 4×4 MIMO) and the coherence time is too small (e.g. 2 ms), CP-OFDM needs to dedicate nearly 60% of the resources to pilots and CP (and thus, only 40% of the resources for data transmission), whereas the UW approach uses only 40% of the resources as overhead and it can dedicate nearly 60% of the resources to data transmission.
- ▷ Proposing sequences for the MIMO unique words, and integrating them in the context of GFDM framework, we showed that by inserting a guard symbol (and thus slightly increasing the overhead), the GFDM's spectral advantages with respect to OFDM can be maintained. Nevertheless, the UW-GFDM has slightly worse OOB radiation with respect to basic GFDM, which is due to the abrupt changes of the signal within the UW sequence.

- ▷ Proposing a new technique to combine the time synchronization metrics in a centralized MIMO system, we showed that our approach achieves a more robust synchronization in low SNR doubly-dispersive channels than the conventional MIMO synchronization techniques. In addition, since the CP based systems are limited to preamble-based synchronization, we showed that the CFO estimation in UW based systems has smaller MSE bounds than the primary CFO estimation (which is solely based on preamble.).

- ▷ In Sec. 4.4, we derived the LMMSE based CE for basic and Circ.-UW sequences, and we also derived the adaptive Wiener-Hopf filters for a more robust channel estimation. There, we observed that if the channel is stationary, the Wiener filtered Circ.-UW sequences achieves the most robust channel estimation with smallest MSE, which is due to the strong temporal correlation of UW sequences, and also smaller condition number of its observation matrix. On the other hand, in an extreme doubly-dispersive channel, we observed that the CE of basic UW sequences has smaller error floor, because their observation matrix considers a shorter time with respect to the channel coherence time.

- ▷ Unlike the conventional UW-based equalization techniques that considers the FFT size over *Payloud-UW* slot, we proposed the UW-free equalization that takes the FFT only over the payload block. Deriving the CWCU-LMMSE for joint channel-equalization-and-demodulation with imperfect channel knowledge, we showed that the UW-based MIMO transmission with a non-orthogonal multi-carrier over a doubly-dispersive channel has nearly two orders of magnitude smaller coded FER than the conventional pilot-aided CP-OFDM. This performance improvement is justified by higher quality channel estimation of UW-based system than the pilot-aided CP based system in time-varying conditions. Furthermore, we have also observed that if the channel remains stationary, the Wiener filtered Circ.-UW GFDM achieves the smallest gap with respect to its Genie-aided receiver, which is due to the fact that i) Circ-UW achieves higher CE quality than basic UW, and ii) B number of channel estimations are being averaged via the Wiener-hopf filter.

5.2 Open Challenges

Below, we outline further open topics and aspects that can extend what has been covered in this thesis.

- ▷ The algorithm developments in this thesis were mostly relying on the assumptions of centralized MIMO systems. In distributed MIMO systems, the scenario becomes more challenging as it influences every module from synchronization to channel estimation and to equalization. In distributed MIMO systems, synchronization must

find the starting position of the frame that has been transmitted from each Tx antenna independently (likewise the CFO estimation). Channel estimation not only needs to consider different PDP between each Tx-Rx antenna pair, but also the timing and frequency misalignment of the frames. Similarly, the equalization must consider the timing and frequency mismatches.

- ▷ The iterative LMMSE-PIC approach of joint channel-estimation-and-equalization that has been derived in Sec. 3.4 was using the remodulated equalized signal constellations as feedback. This iterative approach can be further improved by taking into account the bit LLRs after the channel decoder for interference-cancellation of the channel estimation unit. Making a trade-off between complexity and estimation-detection quality, it would make more room for employing higher MCS values.

- ▷ In synchronization algorithms of Chapter 4, we assumed to have integer STO and we also further assumed that CFO remains constant over multiple blocks. The synchronization can be further extended to the case where fractional STO terms exist and they can be handled by means of early-late synchronizer. Moreover, if the CFO varies slightly during multiple blocks, an adaptive filter can be designed to better track the CFO in UW-based systems.

- ▷ The UW-free equalization approach presented in Sec. 4.5 employed a non-iterative detection approach based on joint channel-equalization-and-demodulation. The approach can be further extended to iterative detection based on LMMSE-PIC, particularly, in an extreme doubly-dispersive channel scenario, where a time-varying channel needs to be equalized. By employing the LMMSE-PIC equalizer, the spectral efficiency of higher MCS e.g. 64-QAM can be exploited in such extreme channel conditions.

- ▷ In our UW-based MIMO frame design, we have considered skipping the first IBI limited L samples of the UW sequence. By doing so, the peak of the timing metric of synchronization unit in AWGN condition depends on the length of UW sequence. One may also consider employing a CP for the UW sequence to improve the channel estimation quality. However, if the UW block length remains identical, the peak of the timing metric may reduce with respect to SNR. Thus, the trade-off between CP-aided UW and CP-less UW for synchronization and channel estimation can be further studied.

Appendix A

Complementary Materials

A.1 Linear Algebra Identities

We have the following well-known identities from linear algebra (available, e.g., from [Gen07],[RC91]):

- Given the quadratic form $\vec{\mathsf{X}}^H \mathbf{A} \vec{\mathsf{X}}$, the following identity holds:

$$\vec{\mathsf{X}}^H \mathbf{A} \vec{\mathsf{X}} = \mathrm{Tr}(\vec{\mathsf{X}}^H \mathbf{A} \vec{\mathsf{X}}). \tag{A.1}$$

- Given the matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}$ such that their dimensions are compatible, the following identity holds:

$$\mathrm{Tr}(\mathbf{ABC}) = \mathrm{Tr}(\mathbf{CAB}). \tag{A.2}$$

- Given a deterministic matrix $\mathbf{A}$ and a random vector $\vec{\mathsf{X}}$, the following identity holds:

$$\mathbb{E}[\mathrm{Tr}(\mathbf{A}\vec{\mathsf{X}}\vec{\mathsf{X}}^H)] = \mathrm{Tr}(\mathbf{A}\mathbb{E}[\vec{\mathsf{X}}\vec{\mathsf{X}}^H]). \tag{A.3}$$

- Given the matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}$ such that their dimensions are compatible, the following identity holds:

$$\mathrm{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\mathrm{vec}(\mathbf{B}). \tag{A.4}$$

- Given the matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}$ such that their dimensions are compatible, the following identity holds:

$$\mathbf{AB}(\boldsymbol{I} + \mathbf{C}^H \mathbf{AB})^{-1} = (\mathbf{A}^{-1} + \mathbf{BC}^H)^{-1}\mathbf{B}. \tag{A.5}$$

- Given the matrix $\mathbf{A} \in \mathbb{C}^{N \times N}$ and the vector $\vec{\mathsf{X}} \in \mathbb{C}^N$, the following identity holds:

$$\mathrm{diag}(\vec{\mathsf{X}})\mathbf{A}\mathrm{diag}(\vec{\mathsf{X}})^H = \vec{\mathsf{X}}\vec{\mathsf{X}}^H \circ \mathbf{A}. \tag{A.6}$$

Proof: Through element-wise computation, one obtains:

$$\begin{aligned}\left[\mathrm{diag}(\vec{\mathsf{X}})\mathbf{A}\mathrm{diag}(\vec{\mathsf{X}})^H\right]_{i,j} &= \vec{\mathsf{X}}_i\mathbf{A}_{i,j}\vec{\mathsf{X}}_j^* \\ &= \vec{\mathsf{X}}_i\vec{\mathsf{X}}_j^*\mathbf{A}_{i,j} \\ &= (\vec{\mathsf{X}}\vec{\mathsf{X}}^H)_{i,j}\mathbf{A}_{i,j} \\ &= \left[\vec{\mathsf{X}}\vec{\mathsf{X}}^H \circ \mathbf{A}\right]_{i,j}.\end{aligned}$$

A.2 Proof of lower triangular Toeplitz channel matrix being defective

We prove by contradiction that the lower triangular Toeplitz channel matrix $\boldsymbol{\mathcal{T}}_{N''}$—defined in (3.67)—is not diagonalizable. Suppose $\boldsymbol{\mathcal{T}}_{N''}$ is diagonalizable. Hence, there exists an invertible matrix $\boldsymbol{P}$ such that

$$\boldsymbol{\mathcal{T}}_{N''} = \boldsymbol{P}\boldsymbol{\mathfrak{D}}_{\mathcal{H}}\boldsymbol{P}^{-1}, \tag{A.7}$$

where $\boldsymbol{\mathfrak{D}}_{\mathcal{H}}$ is a diagonal matrix containing the eigenvalues of $\boldsymbol{\mathcal{T}}_{N''}$ on its diagonal. Since the eigenvalues of any lower triangular matrix is its diagonal elements; and the diagonal elements of any Toeplitz matrix is repetition of its first diagonal element,

$$\boldsymbol{\mathfrak{D}}_{\mathcal{H}} = \lambda\boldsymbol{I}_{N''} \tag{A.8}$$

where λ is any diagonal element of $\boldsymbol{\mathcal{T}}_{N''}$. Therefore, substituting (A.8) into (A.7), we have

$$\begin{aligned}\boldsymbol{\mathcal{T}}_{N''} &= \boldsymbol{P}\lambda\boldsymbol{I}_{N''}\boldsymbol{P}^{-1} \\ &= \lambda\boldsymbol{I}_{N''}\boldsymbol{P}\boldsymbol{P}^{-1} \\ \Rightarrow\Leftarrow\ &\lambda\boldsymbol{I}_{N''},\end{aligned} \tag{A.9}$$

which contradicts the initial assumption on structure of $\boldsymbol{\mathcal{T}}_{N''}$. Since we assumed $\boldsymbol{\mathcal{T}}_{N''}$ is a non-diagonal lower triangular Toeplitz matrix, it is obvious that $\boldsymbol{\mathcal{T}}_{N''} \neq \lambda\boldsymbol{I}_{N''}$. Therefore, the assumption of $\boldsymbol{\mathcal{T}}_{N''}$ being diagonalizable i.e. expression (A.7) is not true.

A.3 Calculation of noise-plus-interference covariance matrix for Pilot- and CP-aided CE

The noise-plus-interference covariance matrix $\boldsymbol{\Sigma}_{EE,q}^{\boldsymbol{J}}$ defined in (3.72) follows:

$$\begin{aligned}\boldsymbol{\Sigma}_{EE,q}^{\boldsymbol{J}} &= \boldsymbol{J}\boldsymbol{F}_{N''}^H\mathbb{E}_{X_d}\left[\boldsymbol{X}_d\bar{\boldsymbol{F}}_{N'',L}\mathbb{E}_h[\vec{h}_q\vec{h}_q^H|\boldsymbol{X}_d]\bar{\boldsymbol{F}}_{N'',L}^H\boldsymbol{X}_d^H\right]\boldsymbol{F}_{N''}\boldsymbol{J}^H + \sigma_w^2\boldsymbol{I}_{N''} \\ &= \boldsymbol{J}\boldsymbol{F}_{N''}^H\left(\sum_{i=1}^{\mathfrak{I}}\boldsymbol{\Sigma}_{\mathrm{XX},N''}^i \circ \boldsymbol{\Sigma}_{\mathrm{HH},q}^i\right)\boldsymbol{F}_{N''}\boldsymbol{J}^H + \sigma_w^2\boldsymbol{I}_{N''}\end{aligned} \tag{A.10}$$

where $\boldsymbol{\Sigma}^{i}_{\mathrm{HH},q} = \boldsymbol{F}_{N'',L}\boldsymbol{\Sigma}_{hh,iq}\boldsymbol{F}^{H}_{N'',L}$ and $\boldsymbol{\Sigma}^{i}_{\mathrm{XX},N''} = \boldsymbol{F}_{N''}\boldsymbol{\mathcal{M}}_{N''}\boldsymbol{\Sigma}^{i}_{dd,ab}\boldsymbol{\mathcal{M}}^{H}_{N''}\boldsymbol{F}^{H}_{N''}$. The covariance matrix $\boldsymbol{\Sigma}^{i}_{dd,ab} \in \mathbb{C}^{2N\times 2N}$ is associated to the transmitted data symbols over two blocks a and b from antenna i, and assuming the symbols are i.i.d. it follows:

$$\boldsymbol{\Sigma}^{i}_{dd,ab} = \mathrm{diag}\left([(\vec{\sigma}^{2}_{d,a})^{T}_{i}\ (\vec{\sigma}^{2}_{d,b})^{T}_{i}]^{T}\right). \tag{A.11}$$

The vectors $(\vec{\sigma}^{2}_{d,a})_i$ and $(\vec{\sigma}^{2}_{d,b})_i$ contain the data variances in blocks a and b at Tx antenna i, respectively. At pilot indexes, $(\vec{\sigma}^{2}_{d,a})_i$ and $(\vec{\sigma}^{2}_{d,b})_i$ have zeros.

A.4 Bock diagonalization of the effective channel for GFDM

Recall the initial system model (2.18) and substitute the modulation matrix $\overline{\boldsymbol{\mathcal{M}}}$ by the GFDM matrix $\bar{\boldsymbol{A}} = \boldsymbol{I}_{\mathfrak{I}} \otimes \boldsymbol{A}$. Thus, we have:

$$\vec{y}_{\mathfrak{Q}N} = \boldsymbol{\mathcal{C}}_{N}\bar{\boldsymbol{A}}\vec{d} + \vec{w}_{\mathfrak{Q}N}, \tag{A.12}$$

in which, for each Rx antenna q:

$$\vec{y}_{q,N} = \sum_{i=1}^{\mathfrak{I}} \boldsymbol{\mathcal{C}}_{N,qi}\boldsymbol{A}\vec{d}_i + \vec{w}_{q,N}. \tag{A.13}$$

Since the individual channels $\boldsymbol{\mathcal{C}}_{N,qi}$ are circulant matrices and $\boldsymbol{A}$ is a block-circulant matrix, the product of $\boldsymbol{\mathcal{C}}_{N,qi}\boldsymbol{A}$ is block-circulant as well. Hence, a DZT of the Rx signal $\vec{y}_{q,N}$ via the matrix $\boldsymbol{Z} = \boldsymbol{F}_{M} \otimes \boldsymbol{I}_{K}$ allows block-diagonalization of $\boldsymbol{\mathcal{C}}_{N,qi}\boldsymbol{A}$ [Qiu95], i.e.

$$(\boldsymbol{I}_{\mathfrak{Q}} \otimes \boldsymbol{Z})\vec{y}_{\mathfrak{Q}N} = \boldsymbol{\mathfrak{B}}\bar{\boldsymbol{Z}}\vec{d} + (\boldsymbol{I}_{\mathfrak{Q}} \otimes \boldsymbol{Z})\vec{w}_{\mathfrak{Q}N}. \tag{A.14}$$

Thus, for the Rx antenna q, we have

$$\boldsymbol{Z}\vec{y}_{q,N} = \sum_{i=1}^{\mathfrak{I}} \boldsymbol{\mathfrak{B}}_{q,i}\boldsymbol{Z}\vec{d}_i + \boldsymbol{Z}\vec{w}_{q,N}, \tag{A.15}$$

with $\boldsymbol{\mathfrak{B}} \in \mathbb{C}^{\mathfrak{Q}N\times\mathfrak{I}N}$ being $\mathfrak{Q}\times\mathfrak{I}$ blocks of $\boldsymbol{\mathfrak{B}}_{q,i} \in \mathbb{C}^{N\times N}$ and $\boldsymbol{\mathfrak{B}}_{q,i}{=}\boldsymbol{Z}\boldsymbol{\mathcal{C}}_{N,qi}\boldsymbol{A}\boldsymbol{Z}^{H}$ being a block-diagonal matrix with M blocks of size $K\times K$. As a consequence, by treating each block individually, we can split the DZT domain Rx signal $\boldsymbol{Z}\vec{y}_{q,N}$ into M sequences of length K samples.

A.5 Detailed complexity analysis of Sec. 3.4

In this appendix, we analyze the complexity of each algorithm of Sec. 3.4 in terms of floating point operations (FLOPs). Table A.1 is considered as a reference for complexity analysis.

Table A.1: Reference for complexity analysis

Algorithm/Operation	Expression	Dimensions	Number of FLOPs
Cholesky [19] decomposition, [forward, back substitution]	$\boldsymbol{AX} = \boldsymbol{B}$	$\boldsymbol{A} \in \mathbb{C}^{N\times N}, \boldsymbol{X} \in \mathbb{C}^{N\times M}$	$N^3/3$ $+[M(2N^2 - N)]$
Banded solver [20, functions *ZGBTRF, ZGBTRS*]	$\boldsymbol{A}\vec{x} = \vec{b}$	$\boldsymbol{A} \in \mathbb{C}^{N\times N}, \vec{x} \in \mathbb{C}^{N}$	$2N\,(B_l(B_l + B_u) + 3B_l + B_u)$
Matrix multiplication	$\boldsymbol{AB}$	$\boldsymbol{A} \in \mathbb{C}^{N\times L}, \boldsymbol{B} \in \mathbb{C}^{L\times M}$	$NM(2L - 1)$
Diagonal to full matrix multiplication	$\mathrm{diag}(\vec{b})\boldsymbol{A}$	$\vec{b} \in \mathbb{C}^{N}$, $\boldsymbol{A} \in \mathbb{C}^{N\times M}$	NM
Matrix addition	$\boldsymbol{A} + \boldsymbol{B}$	$\boldsymbol{A} \in \mathbb{C}^{N\times M}, \boldsymbol{B} \in \mathbb{C}^{N\times M}$	NM

B_l : Lower bandwidth of the banded matrix, B_u : Upper bandwidth of the banded matrix, $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{X}, \vec{x}$ and $\vec{b}$ are generic.

Assuming the receiver is capable to process the Rx signals in parallel for each receive antenna, the complexity of (3.78) in terms of number of FLOPs is

$$\begin{aligned} C_{(3.78)} = \mathfrak{Q}\Big([\mathfrak{I}LN'] + [(N')^2(2\mathfrak{I}L - 1)] + [(N')^2] \\ +[\frac{(N')^3}{3} + 2(N')^2 - N'] + [\mathfrak{I}L(2N' - 1)]\Big), \end{aligned} \tag{A.16}$$

where each successive term corresponds to calculations of $(\boldsymbol{\Sigma}^q_{hh}\boldsymbol{\mathcal{Q}}^H_p)$, $\boldsymbol{\mathcal{Q}}_p(\boldsymbol{\Sigma}^q_{hh}\boldsymbol{\mathcal{Q}}^H_p)$, $\boldsymbol{\mathcal{Q}}_p\boldsymbol{\Sigma}^q_{hh}\boldsymbol{\mathcal{Q}}^H_p + \boldsymbol{\Sigma}^{\boldsymbol{J}}_{\Psi\Psi,q}$, $(\boldsymbol{\mathcal{Q}}_p\boldsymbol{\Sigma}^q_{hh}\boldsymbol{\mathcal{Q}}^H_p + \boldsymbol{\Sigma}^{\boldsymbol{J}}_{\Psi\Psi,q})^{-1}\vec{y}_{q,N'}$, $\hat{\vec{h}}_{\mathrm{CPLMMSE},q}$, respectively.

In a reduced complexity fashion where we ignore the interference term as in Sec. 3.4.4, the complexity becomes

$$C_{\mathrm{SLMMSE}} = \mathfrak{Q}\Big(C_Z + C_{(3.85)} + (M-1)(C_{(3.92)} + C_{(3.93)})\Big), \tag{A.17}$$

where $C_Z = KC_{\boldsymbol{F}_M}$, is the effort to take the time domain signal $\vec{y}_{q,N}$ into M sequences of DZT domain observations $\vec{Y}_{m,q,K}$. $C_{\boldsymbol{F}_M}$ is the required number of FLOPs for M-point DFT operation and we approximate it by $C_{\boldsymbol{F}_M} \approx M\log_2 M$. Next, the computational effort of initial estimation (3.85) is

$$\begin{aligned} C_{(3.85)} \;=\; [K] + [\mathfrak{I}L(2K-1)] + [\mathfrak{I}^2L^2(2K-1)] \\ +[\mathfrak{I}^2L^2] + [\mathfrak{I}L] + [\frac{\mathfrak{I}^3L^3}{3} + 2\mathfrak{I}^2L^2 - \mathfrak{I}L], \end{aligned} \tag{A.18}$$

where we calculate $\frac{\vec{Y}_{0,q,K}}{\sigma^2_w}$, $\frac{\tilde{\boldsymbol{X}}^H_{p,0}}{\sigma^2_w}\vec{Y}_{0,q,K}$, $\tilde{\boldsymbol{X}}^H_{p,0}\tilde{\boldsymbol{X}}_{p,0}$, $\frac{\tilde{\boldsymbol{X}}^H_{p,0}\tilde{\boldsymbol{X}}_{p,0}}{\sigma^2_w}$, $\left((\boldsymbol{\Sigma}^q_{hh})^{-1} + \frac{\tilde{\boldsymbol{X}}^H_{p,0}\tilde{\boldsymbol{X}}_{p,0}}{\sigma^2_w}\right)$ and $\hat{\vec{h}}_{0,q}$ in consecutive order. Then, for (3.92) we have

$$\begin{aligned} C_{(3.92)} = [K(2\mathfrak{I}^2L^2 - \mathfrak{I}L)] + [K^2(2\mathfrak{I}L - 1)] + [K] \\ +[\frac{K^3}{3} + 2K^2 - K] + [\mathfrak{I}L(2K-1)] + [\mathfrak{I}L], \end{aligned} \tag{A.19}$$

here, the first term is the effort of K times forward and back substitution for $\widehat{\boldsymbol{\Sigma}}^q_{hh,m-1}\tilde{\boldsymbol{X}}^H_{p,m}$ in which we omitted $\frac{\mathfrak{I}^3L^3}{3}$ FLOPs, because Cholesky decomposition [Gen07] has already been performed in (3.85). The rest of operations correspond to $\tilde{\boldsymbol{X}}_{p,m}\boldsymbol{\Sigma}_{h\tilde{Y},m}$,

$(\tilde{\boldsymbol{X}}_{p,m}\boldsymbol{\Sigma}_{h\tilde{Y},m} + \boldsymbol{I}_K\sigma_w^2)$, $\boldsymbol{\Sigma}^{-1}_{\tilde{Y}\tilde{Y},m}\vec{\tilde{Y}}_{m,q,K}$, $\Delta\hat{\vec{h}}_{m,q}$ and $\hat{\vec{h}}_{m,q}$, respectively. Next, for computation of the covariance matrix $\widehat{\boldsymbol{\Sigma}}^q_{hh,m}$ we need

$$\begin{aligned} C_{(3.93)} \;=\; & [\mathfrak{I}L(2K^2-K)] + [\mathfrak{I}^2L^2(2K-1)] \\ & + [\mathfrak{I}L] + [2\mathfrak{I}^3L^3 - \mathfrak{I}^2L^2] \end{aligned} \tag{A.20}$$

FLOPs, where we calculate $\boldsymbol{\Sigma}^{-1}_{\tilde{Y}\tilde{Y},m}\tilde{\boldsymbol{X}}_{p,m}$, $\boldsymbol{\Sigma}_{h\tilde{Y},m}\boldsymbol{\Sigma}^{-1}_{\tilde{Y}\tilde{Y},m}\tilde{\boldsymbol{X}}_{p,m}$, $\Big(\boldsymbol{I}_{\mathfrak{I}L} - \boldsymbol{\Sigma}_{h\tilde{Y},m}\boldsymbol{\Sigma}^{-1}_{\tilde{Y}\tilde{Y},m}\tilde{\boldsymbol{X}}_{p,m}\Big)$ and $\widehat{\boldsymbol{\Sigma}}^q_{hh,m}$, consecutively. Note that here again, the computational effort for Cholesky decomposition of $\boldsymbol{\Sigma}_{\tilde{Y}\tilde{Y},m}$ and $\widehat{\boldsymbol{\Sigma}}^q_{hh,m-1}$ are saved due to prior calculations, though, for the first and the last steps, we had to perform forward and back substitutions for $\mathfrak{I}L$ columns.

In case, we are interested to improve the channel estimation quality via the additional pilots energy in CP, we need to repeat the SLMMSE for another M sequences $\vec{Y}'_{m,q,K}$. Therefore,

$$C_{\text{CP-SLMMSE}} = C_{\text{SLMMSE}} + \mathfrak{Q}\Big(C_Z + M(C_{(3.92)} + C_{(3.93)})\Big), \tag{A.21}$$

is the computational burden of what we described in CP-SLMMSE estimation part of Sec. 3.4.4.

Should the channel estimation be done at high SNR regions where interference is dominant and thus, it becomes a bottleneck for utilization of high MCS, we employ parallel interference cancellation at the following expenses:

$$\begin{aligned} C_{\text{CP-LMMSE-PIC}} = N_{\text{PIC}}(& C_{(3.99)} + C_{(3.100)} + C_{(3.102)} \\ & + C_{(3.104)} + C_{(3.105)}) + (N_{\text{PIC}}+1)C_{(3.78)}, \end{aligned} \tag{A.22}$$

where N_{PIC} is the number of PIC iterations. Herein, $C_{(3.99)}$ and $C_{(3.101)}$ are the required number of FLOPs for equalization (3.99) and its covariance (3.101), and they are given by

$$\begin{aligned} C_{(3.99)} \;=\; & C_{(3.101a)} + C_{(3.101b)} + 2\mathfrak{Q}N \\ & + 2\mathfrak{Q}N(B_l(B_l+B_u) + 3B_l + B_u) \\ & + \mathfrak{Q}N(2B_l + 2B_u - 1) \\ & + \mathfrak{I}N(2B_l + 2B_u - 1), \end{aligned} \tag{A.23}$$

$$C_{(3.100)} \;=\; \mathfrak{Q}N + C_{(3.101)}, \tag{A.24}$$

$$\begin{aligned} C_{(3.101)} \;=\; & 2\mathfrak{I}\mathfrak{Q}N^2(2B_l + B_u) \\ & + \mathfrak{I}\mathfrak{Q}N^2(2B_l + 2B_u - 1), \end{aligned} \tag{A.25}$$

respectively, where

$$C_{(3.101a)} \;=\; \mathfrak{Q}\mathfrak{I}N^2 + \mathfrak{Q}^2(2\mathfrak{I}-1)N^2, \tag{A.26}$$

$$C_{(3.101b)} \;=\; (2\mathfrak{I}-1)N^2 \tag{A.27}$$

wherein for $C_{(3.101a)}$, we assumed, the product of $\bar{\boldsymbol{F}}\overline{\boldsymbol{\mathcal{M}}}\bar{\boldsymbol{\Sigma}}_{dd}\overline{\boldsymbol{\mathcal{M}}}^H\bar{\boldsymbol{F}}^H$ is precalculated and only the multiplication of its product with the diagonal channels $\hat{\boldsymbol{H}}_{qi}$ is necessary to be

performed for each block. A similar rule also holds for $C_{(3.101b)}$ with the assumption of $\boldsymbol{F}_N \boldsymbol{\mathcal{M}} \boldsymbol{\Sigma}_{dd,i} \boldsymbol{\mathcal{M}}^H \boldsymbol{F}_N^H$ being precalculated.

Next, we calculate the complexity of interference cancellation (3.102) as

$$\begin{aligned} C_{(3.102)} \;=\; & 2\mathfrak{I} C_{\boldsymbol{\mathcal{M}}} + \mathfrak{I} C_{\boldsymbol{F}_{N''}} \\ & + \mathfrak{Q}(2\mathfrak{I}-1)N'' + \mathfrak{Q} C_{\boldsymbol{F}_{N''}} + \mathfrak{Q} N' \end{aligned} \tag{A.28}$$

for one iteration, which includes the complexity of modulation $C_{\boldsymbol{\mathcal{M}}}$ for two consecutive blocks, N''-point DFT $C_{\boldsymbol{F}_{N''}} \approx N'' \log_2 N''$, diagonal channels multiplication, and N''-point IDFT. In the special case of GFDM, $C_{\boldsymbol{\mathcal{M}}}$ becomes [MMG+16]

$$C_{\boldsymbol{\mathcal{M}}_{\mathrm{GFDM}}} = M C_{\boldsymbol{F}_K} + K(2M-1). \tag{A.29}$$

In (A.28), the complexity of windowing operation, i.e. multiplication with $\boldsymbol{J}$, and Kronecker product with $\boldsymbol{I}_{\mathfrak{Q}}$ has not been taken into account, because they can be performed as manipulations of registers. Moreover, since for each iteration of PIC, the interference covariance matrix needs to be updated, we have,

$$\begin{aligned} C_{(3.104)} = {} & C_{\hat{\boldsymbol{H}}_{\mathrm{eff}}} + \mathfrak{Q} N'(4\mathfrak{I} N - 1) + \\ & (\mathfrak{Q} N')^2 (4\mathfrak{I} N - 1), \end{aligned} \tag{A.30}$$

$$C_{(3.105)} = N''^2(2\mathfrak{I}-1) + 2N'' C_{\boldsymbol{F}_{N''}}, \tag{A.31}$$

with

$$C_{\hat{\boldsymbol{H}}_{\mathrm{eff}}} = \mathfrak{Q} N''(2\mathfrak{I} N'' - 1) + \mathfrak{Q} N'' C_{\boldsymbol{F}_{N''}}, \tag{A.32}$$

where we assume $\bar{\boldsymbol{F}}_{N''} \overline{\boldsymbol{\mathcal{M}}}_{N''}$ is precalculated.

In order to reduce the complexity (as discussed in Sec. 3.4.4) by ignoring the interference covariance matrix and performing SLMMSE, the computational efforts of SLMMSE and CP-SLMMSE combined with PIC reduce respectively to

$$\begin{aligned} C_{\text{SLMMSE-PIC}} \quad = {} & N_{\mathrm{PIC}}(C_{(3.99)} + C_{(3.100)} \\ & + C'_{\mathrm{PIC}}) + (N_{\mathrm{PIC}} + 1) C_{\mathrm{SLMMSE}}, \end{aligned} \tag{A.33}$$

$$\begin{aligned} C_{\text{CP-SLMMSE-PIC}} = {} & N_{\mathrm{PIC}}(C_{(3.99)} + C_{(3.100)} \\ & + C_{(3.102)}) + (N_{\mathrm{PIC}} + 1) C_{\text{CP-SLMMSE}}. \end{aligned} \tag{A.34}$$

Herein,

$$C'_{\mathrm{PIC}} = \mathfrak{I} C_{\boldsymbol{\mathcal{M}}} + \mathfrak{I} N \log_2 N + \mathfrak{Q}(2\mathfrak{I}-1)N + \mathfrak{Q}(N \log_2 N) \tag{A.35}$$

is the computational effort of interference cancellation without consideration of pilots energy inside CP.

A.6 CRLB derivations for the pdf (4.24)

The measurement vector $\vec{z}$ depends on two independent random variables ε and $w_q[n]$. In the following, we consider the conditional pdf $p(\vec{z}|\varepsilon, h_{\ell,i,q})$, conditioned by the unknown parameters ε and $h_{\ell,i,q}[n]$ is given by

$$p(\vec{z}|\varepsilon, h_{\ell,i,q}) = \frac{1}{2\pi\sigma_w^2} \exp\left\{ -\frac{1}{2\sigma_w^2} \sum_{q=0}^{\mathfrak{Q}-1} \sum_{n\in\mathcal{N}_u} \left| z_q[n] - \exp(j2\pi n\varepsilon/N)\mathfrak{y}_q[n] \right|^2 \right\}, \quad \text{(A.36)}$$

which can alternatively be written as

$$\begin{aligned} p(\vec{z}|\varepsilon, h_{\ell,i,q}) &= \frac{1}{2\pi\sigma_w^2} \exp\left\{ -\frac{1}{2\sigma_w^2} \sum_{q=0}^{\mathfrak{Q}-1} \sum_{n\in\mathcal{N}_u} \right. \quad \text{(A.37)} \\ &\qquad \left. \Big(z_q[n] - \exp(j2\pi n\varepsilon/N)\mathfrak{y}_q[n]\Big)^* \Big(z_q[n] - \exp(j2\pi n\varepsilon/N)\mathfrak{y}_q[n]\Big) \right\}, \\ &= \frac{1}{2\pi\sigma_w^2} \exp\left\{ -\frac{1}{2\sigma_w^2} \sum_{q=0}^{\mathfrak{Q}-1} \sum_{n\in\mathcal{N}_u} \Big(z_q^*[n]z_q[n] - z_q^*[n]\exp(j2\pi n\varepsilon/N)\mathfrak{y}_q[n] \right. \\ &\qquad \left. - z_q[n]\exp(-j2\pi n\varepsilon/N)\mathfrak{y}_q^*[n] + \mathfrak{y}_q^*[n]\mathfrak{y}_q[n] \Big) \right\}. \end{aligned}$$

Taking the derivative of the log-likelihood function $\ln(p(\vec{z}|\varepsilon, h_{\ell,i,q}))$ with respect to ε yields

$$\begin{aligned} \frac{\partial \ln(p(\vec{z}|\varepsilon, h_{\ell,i,q}))}{\partial\varepsilon} = -\frac{1}{2\sigma_w^2} \sum_{q=0}^{\mathfrak{Q}-1} \sum_{n\in\mathcal{N}_u} \Big(& - z_q^*[n]\mathfrak{y}_q[n](j2\pi\frac{n}{N})\exp(j2\pi n\varepsilon/N) \quad \text{(A.38)} \\ & - z_q[n]\mathfrak{y}_q^*[n](-j2\pi\frac{n}{N})\exp(-j2\pi n\varepsilon/N)\Big). \end{aligned}$$

Subsequently, the second derivative of the log-likelihood function yields

$$\begin{aligned} \frac{\partial^2 \ln(p(\vec{z}|\varepsilon, h_{\ell,i,q}))}{\partial\varepsilon^2} = -\frac{1}{2\sigma_w^2} \sum_{q=0}^{\mathfrak{Q}-1} \sum_{n\in\mathcal{N}_u} \Big(& - z_q^*[n]\mathfrak{y}_q[n](-4\pi^2\frac{n^2}{N^2})\exp(j2\pi n\varepsilon/N) \quad \text{(A.39)} \\ & - z_q[n]\mathfrak{y}_q^*[n](-4\pi^2\frac{n^2}{N^2})\exp(-j2\pi n\varepsilon/N)\Big). \end{aligned}$$

Then, we take the negative expected value and thus, we have

$$-\mathbb{E}\left[\frac{\partial^2 \ln p(\vec{z}|\varepsilon, h_{\ell,i,q})}{\partial\varepsilon^2}\right] = \frac{1}{2\sigma_w^2} \sum_{n\in\mathcal{N}_u} 8\pi^2 \frac{n^2}{N^2} \mathbb{E}_h\left[\sum_{q=0}^{\mathfrak{Q}-1} |\mathfrak{y}_q[n]|^2\right], \quad \text{(A.40)}$$

where we set $\mathbb{E}\Big[z_q[n]\Big] = \exp(j2\pi n\varepsilon/N)\mathfrak{y}_q[n]$ and $|\mathfrak{y}_q[n]|^2 = \mathfrak{y}_q^*[n]\mathfrak{y}_q[n]$. Additionally, since $h_{\ell,i,q}$ is independent of ε and $w_q[n]$, taking the expected value $\mathbb{E}_h[\cdot]$ with respect to $h_{\ell,i,q}$, we obtain

$$\begin{aligned} \mathbb{E}_h\left[\sum_{q=0}^{\mathfrak{Q}-1} |\mathfrak{y}_q[n]|^2\right] &= \sum_{q=0}^{\mathfrak{Q}-1} \sum_{i=0}^{\mathfrak{I}-1} \sum_{\ell=0}^{L-1} P_\ell |x_i[n-\ell]|^2 \quad \text{(A.41)} \\ &= \mathfrak{Q}\mathfrak{I}, \quad \text{(A.42)} \end{aligned}$$

where we set $|x_i[n]| = |x_{\mathrm{u},i}[n]| = 1$, i.e. the UW sequence defined in (4.9).

A.7 Proof that (4.45) emulates a circular CIR at the receiver

In this Appendix, we mathematically prove that the definition of $\vec{y}_{\text{cyclic},q}$ in (4.45) emulates a circular channel matrix for the receive signal. Without loss of generality, as the extension of SISO to MIMO in this proof is a straightforward approach, we ignore the antenna indexes i and q for brevity. Additionally, we set the UW length to L samples (i.e. $\vec{x}_u \in \mathbb{C}^L$) for simplicity of the expressions. By extending the vector-matrix dimensions in a respectful manner, the proof also remains valid for the cases where the length of $\vec{x}_u$ is larger than L.

Consider a stationary channel condition where the channel matrix based on linear convolution is a lower triangular Toeplitz matrix decomposed as follows:

$$\boldsymbol{\mathcal{T}} \triangleq \begin{bmatrix} \boldsymbol{\mathcal{T}}_{\mathbf{11}} & \mathbf{0}_{L\times N_d} & \mathbf{0}_{L\times L} \\ \boldsymbol{\mathcal{T}}_{\mathbf{21}} & \boldsymbol{\mathcal{T}}_{\mathbf{22}} & \mathbf{0}_{N_d\times L} \\ \mathbf{0}_{L\times L} & \boldsymbol{\mathcal{T}}_{\mathbf{32}} & \boldsymbol{\mathcal{T}}_{\mathbf{33}} \end{bmatrix} \tag{A.43}$$

where $\boldsymbol{\mathcal{T}}_{\mathbf{11}} \in \mathbb{C}^{L\times L}, \boldsymbol{\mathcal{T}}_{\mathbf{22}} = \begin{bmatrix} \boldsymbol{\mathcal{T}}_{\mathbf{22,12}} & \mathbf{0}_{L\times L} \\ \boldsymbol{\mathcal{T}}_{\mathbf{22,21}} & \boldsymbol{\mathcal{T}}_{\mathbf{22,22}} \end{bmatrix} \in \mathbb{C}^{N_d\times N_d}$ and $\boldsymbol{\mathcal{T}}_{\mathbf{33}} \in \mathbb{C}^{L\times L}$ are also lower triangular Toeplitz matrices with the CIR $\vec{h} \in \mathbb{C}^L$ on their first column. Moreover, the matrix $\boldsymbol{\mathcal{T}}_{\mathbf{22,12}}$ takes only the first L rows and the first $N_d - L$ columns of $\boldsymbol{\mathcal{T}}_{\mathbf{22}}$; $\boldsymbol{\mathcal{T}}_{\mathbf{22,21}}$ contains the last $N_d - L$ rows and the first $N_d - L$ columns of $\boldsymbol{\mathcal{T}}_{\mathbf{22}}$ and $\boldsymbol{\mathcal{T}}_{\mathbf{22,22}}$ includes the last $N_d - L$ rows and the last L columns of $\boldsymbol{\mathcal{T}}_{\mathbf{22}}$. For the receive signal

$$\vec{y}_{\text{UW-P-UW}} = \boldsymbol{\mathcal{T}}[\vec{x}_u^T,\ \vec{x}_d^T,\ \vec{x}_u^T]^T + \vec{w}_{\text{UW-P-UW}}, \tag{A.44}$$

the matrices $\boldsymbol{\mathcal{T}}_{\mathbf{21}} = [\boldsymbol{\mathcal{T}}_{\mathbf{21,1}}^T,\ \mathbf{0}_{(N_d-L)\times L}^T]^T$ and $\boldsymbol{\mathcal{T}}_{\mathbf{32}} = [\mathbf{0}_{L\times(N_d-L)},\ \boldsymbol{\mathcal{T}}_{\mathbf{32,2}}]$ cause the leakage of the first UW sequence into the receive counterpart of $\vec{x}_d$ and the leakage of the payload into the receive counterpart of second UW sequence, respectively. Here, $\boldsymbol{\mathcal{T}}_{\mathbf{21,1}} = \boldsymbol{\mathcal{T}}_{\mathbf{32,2}} \in \mathbb{C}^{L\times L}$ are identical upper triangular Toeplitz matrices with $[\vec{h}_{(1:L-1)}^T,\ 0]^T$ on their last column.

We note that in a CP based system where the left $\vec{x}_u$ in (A.44) is substituted by $\vec{x}_{\text{cp}}$ (i.e. the last L samples of $\vec{x}_d$), the receive payload $\vec{y}_{d,\text{CP-based}} = \boldsymbol{\mathcal{C}}_{\text{CP-based}}\vec{x}_d + \vec{w}_d$ experiences a circular channel $\boldsymbol{\mathcal{C}}_{\text{CP-based}} = \begin{bmatrix} \boldsymbol{\mathcal{T}}_{\mathbf{22,12}} & \boldsymbol{\mathcal{T}}_{\mathbf{21,1}} \\ \boldsymbol{\mathcal{T}}_{\mathbf{22,21}} & \boldsymbol{\mathcal{T}}_{\mathbf{22,22}} \end{bmatrix}$. In fact, having the matrix $\boldsymbol{\mathcal{T}}_{\mathbf{21,1}}$ on top right corner of $\boldsymbol{\mathcal{T}}_{\mathbf{22}}$ makes the matrix $\boldsymbol{\mathcal{C}}_{\text{CP-based}}$ circulant. In conventional UW-based systems where the circularity of the channel is facilitated by taking an N-point FFT over *Payload-UW* block, we have $\vec{y}_{\text{P-UW}} = \boldsymbol{\mathcal{C}}_{\text{conv-UW}}[\vec{x}_d^T,\ \vec{x}_u^T]^T + \vec{w}_{\text{P-UW}}$, where $\boldsymbol{\mathcal{C}}_{\text{conv-UW}} = \begin{bmatrix} \boldsymbol{\mathcal{T}}_{\mathbf{22}} & \boldsymbol{\mathcal{T}}_{\mathbf{21}} \\ \boldsymbol{\mathcal{T}}_{\mathbf{32}} & \boldsymbol{\mathcal{T}}_{\mathbf{33}} \end{bmatrix}$ and thus, $\boldsymbol{\mathcal{C}}_{\text{conv-UW}}$ is also ciruclant. In (4.45) however, we aim to generate a circulant channel $\boldsymbol{\mathcal{C}}$ over only the payload sequence, and thus obtain the receive signal

$$\vec{y}_d = \boldsymbol{\mathcal{H}}\vec{x}_d + \vec{w}'_d. \tag{A.45}$$

In order to prove that $\boldsymbol{\mathcal{H}} = \boldsymbol{\mathcal{C}}$, we decompose the receive *Payload-UW* signal as

$$\begin{bmatrix} \vec{y}_d \\ \vec{y}_u \end{bmatrix} = \boldsymbol{\mathfrak{C}} \left(\boldsymbol{\mathcal{T}}\boldsymbol{\mathfrak{R}}[\vec{x}_d^T,\ \vec{x}_u^T]^T + \begin{bmatrix} \vec{w}_d \\ \vec{w}_u \end{bmatrix} \right), \tag{A.46}$$

where each individual matrix is defined as

$$\boldsymbol{\mathfrak{R}} \triangleq \begin{bmatrix} \mathbf{0}_{L\times N_d} & \boldsymbol{I}_L \\ \boldsymbol{I}_{N_d} & \mathbf{0}_{N_d\times L} \\ \mathbf{0}_{L\times N_d} & \boldsymbol{I}_L \end{bmatrix}, \tag{A.46a}$$

$$\boldsymbol{\mathfrak{C}} \triangleq \begin{bmatrix} \mathbf{0}_{L\times L} & \boldsymbol{I}_L & \mathbf{0}_{L\times(N_d-L)} & \boldsymbol{I}_L \\ \mathbf{0}_{(N_d-L)\times L} & \mathbf{0}_{(N_d-L)\times L} & \boldsymbol{I}_{N_d-L} & \mathbf{0}_{(N_d-L)\times L} \\ \mathbf{0}_{L\times L} & \mathbf{0}_{L\times L} & \mathbf{0}_{(N_d-L)\times(N_d-L)} & \boldsymbol{I}_L \end{bmatrix}. \tag{A.46b}$$

Here, the repetition matrix $\boldsymbol{\mathfrak{R}}$ is located at the transmitter side and it copies the UW sequence before the payload, i.e. $\boldsymbol{\mathfrak{R}}[\vec{x}_d^T,\ \vec{x}_u^T]^T = [\vec{x}_u^T,\ \vec{x}_d^T,\ \vec{x}_u^T]^T$. On the other hand, the matrix $\boldsymbol{\mathfrak{C}}$, which is located at the receiver side performs the operation defined in (4.45).

Evaluating the effective channel of $[\vec{x}_d^T,\ \vec{x}_u^T]^T$ in (A.46), we have

$$\boldsymbol{\mathcal{T}}\boldsymbol{\mathfrak{R}} = \begin{bmatrix} \mathbf{0}_{L\times N_d} & \boldsymbol{\mathcal{T}}_{\mathbf{11}} \\ \boldsymbol{\mathcal{T}}_{\mathbf{22}} & \boldsymbol{\mathcal{T}}_{\mathbf{21}} \\ \boldsymbol{\mathcal{T}}_{\mathbf{32}} & \boldsymbol{\mathcal{T}}_{\mathbf{33}} \end{bmatrix}, \tag{A.47}$$

$$\boldsymbol{\mathfrak{C}}\boldsymbol{\mathcal{T}}\boldsymbol{\mathfrak{R}} = \begin{bmatrix} [\boldsymbol{\mathcal{T}}_{\mathbf{22,12}} \ \ \mathbf{0}_{L\times L}] + \boldsymbol{\mathcal{T}}_{\mathbf{32}} & \boldsymbol{\mathcal{T}}_{\mathbf{21,1}} + \boldsymbol{\mathcal{T}}_{\mathbf{33}} \\ [\boldsymbol{\mathcal{T}}_{\mathbf{22,21}} \ \ \boldsymbol{\mathcal{T}}_{\mathbf{22,22}}] & \mathbf{0}_{N_d-L\times L} \\ \boldsymbol{\mathcal{T}}_{\mathbf{32}} & \boldsymbol{\mathcal{T}}_{\mathbf{33}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mathcal{C}} & \begin{bmatrix} \boldsymbol{\mathcal{T}}_{\mathbf{21,1}} + \boldsymbol{\mathcal{T}}_{\mathbf{33}} \\ \mathbf{0}_{N_d-L\times L} \end{bmatrix} \\ \boldsymbol{\mathcal{T}}_{\mathbf{32}} & \boldsymbol{\mathcal{T}}_{\mathbf{33}} \end{bmatrix}, \tag{A.48}$$

where

$$\boldsymbol{\mathcal{C}} = \begin{bmatrix} \boldsymbol{\mathcal{T}}_{\mathbf{22,12}} & \boldsymbol{\mathcal{T}}_{\mathbf{32,2}} \\ \boldsymbol{\mathcal{T}}_{\mathbf{22,21}} & \boldsymbol{\mathcal{T}}_{\mathbf{22,22}} \end{bmatrix} = \boldsymbol{\mathcal{H}}. \tag{A.49}$$

Here, having the matrix $\boldsymbol{\mathcal{T}}_{\mathbf{32,2}}$ on top right corner of $\boldsymbol{\mathcal{H}}$ makes it circulant. In addition, the matrix $\boldsymbol{\mathcal{C}}_{\text{ISI}} = \boldsymbol{\mathcal{T}}_{\mathbf{21,1}} + \boldsymbol{\mathcal{T}}_{\mathbf{33}}$—which also emulates a circulant channel on the first L samples of $\vec{x}_u$—causes ISI into the first L samples of $\vec{y}_d$. However, since both $\vec{x}_u$ and $\hat{\vec{h}}$ are known, one can simply remove the known interference, i.e. (4.46). We also note that, since the matrix $\boldsymbol{\mathfrak{C}}$ is located at the receiver side, the first L samples of $\vec{w}_u$ would also be added to the first L samples of $\vec{w}_d$. If $N_d >> L$, the noise correlation would be negligible. Nonetheless, the operation (4.45) provides the significant advantage of designing payload and UW sequences separately, and thus performing N_d-point FFT only on the received payload signal.

Publications of the Author

Journal Publications & Book Chapters

[1] **Shahab Ehsanfar**, Marwa Chafii, and Gerhard Fettweis. On UW-based Transmission for Massive MIMO Multi-Carriers with Spatial Multiplexing. In *IEEE Transactions on Wireless Communications*, 2019. [resubmitted]

[2] **Shahab Ehsanfar**, Maximilian Matthé, Marwa Chafii, and Gerhard Fettweis. Pilot- and CP-Aided Channel Estimation in MIMO Non-Orthogonal Multi-Carriers. In *IEEE Transactions on Wireless Communications*, 2018.

[3] Ahmad Nimr, **Shahab Ehsanfar**, Nicola Michailow, Martin Danneberg, Dan Zhang, Henry Douglas Rodrigues, Luciano Leonel Mendes, and Gerhard Fettweis. Multiple Access Techniques for 5G Wireless Networks and Beyond Chapter: *Generalized Frequency Division Multiplexing: A Flexible Multicarrier Waveform* Springer, 2019.

Conference Publications

[4] **Shahab Ehsanfar**, Marwa Chafii, and Gerhard Fettweis. Unique-Word based Synchronization for MIMO Systems over Time-Varying Channels. In *IEEE Wireless Communications and Networking Conference (WCNC)*, 2020. [accepted]

[5] **Shahab Ehsanfar**, Marwa Chafii, and Gerhard Fettweis. A Frame Design for MIMO UW based Systems: Overhead Analysis & Channel Estimation. In *IEEE 5G World Forum (WF-5G)*, 2019.

[6] **Shahab Ehsanfar**, Marwa Chafii, and Gerhard Fettweis. Time-Variant Pilot-and CP-Aided Channel Estimation for GFDM. In *IEEE International Conference on Communications (ICC)*, 2019.

[7] **Shahab Ehsanfar**, Maximillian Matthé, Dan Zhang, and Gerhard Fettweis. Interference-free pilots insertion for MIMO-GFDM channel estimation. In *IEEE Wireless Communications and Networking Conference (WCNC)*, 2017.

[8] **Shahab Ehsanfar**, Maximillian Matthé, Dan Zhang, and Gerhard Fettweis. Theoretical analysis and CRLB evaluation for pilot-aided channel estimation in GFDM. In *IEEE Global Communications Conference (GLOBECOM)*, 2016.

[9] **Shahab Ehsanfar**, Maximillian Matthé, Dan Zhang, and Gerhard Fettweis. A study of pilot-aided channel estimation in MIMO-GFDM systems. In *20th International ITG Workshop on Smart Antennas (WSA)*, 2016.

[10] Martin Danneberg, Zhongju Li, Paul Kühne, Ahmad Nimr, **Shahab Ehsanfar**, Marwa Chafii, and Gerhard Fettweis. Real-Time Waveform Prototyping. In *IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2019.

[11] Martin Danneberg, Roberto Bomfin, **Shahab Ehsanfar**, Ahmad Nimr, Zhitao Lin, Marwa Chafii, and Gerhard Fettweis. Online Wireless Lab Testbed In *IEEE WCNC 2019 Workshop on Enabling Ultra Reliability and Low Latency Communications for 5G and Beyond (EURCOMM)*, 2019.

[12] Martin Danneberg, Ahmad Nimr, Nicola Michailow, **Shahab Ehsanfar**, Maximillian Matthé, Ana-Belen Martinez, Dan Zhang, and Gerhard Fettweis. Universal Waveforms Processor. In *European Conference on Networks and Communications (EuCNC)*, 2018.

Bibliography

[3GP18] 3GPP. NR; NR and NG-RAN Overall Description; Stage 2, Rel. 15. Technical Specification (TS) 38.300, 3rd Generation Partnership Project (3GPP); Tech. Spec. Group Radio Access Network, 06 2018. V15.2.0.

[Abo12] A. Aboltins. Carrier Frequency Offset Estimator Based on Unique Word Cross-correlation. In *2012 20th Telecommunications Forum (TELFOR)*, pages 486–489. IEEE, 2012.

[Acc09] E. U. T. R. Access. Physical channels and modulation. *3GPP TS*, 36:V8, 2009.

[AKE08] A. B. Awoseyila, C. Kasparis, and B. G. Evans. Improved preamble-aided timing estimation for OFDM systems. *IEEE Commun. Lett.*, 12(11):825–827, 2008.

[BLM03] I. Barhumi, G. Leus, and M. Moonen. Optimal Training Design for MIMO OFDM Systems in Mobile Wireless Channels. *IEEE Transactions on signal processing*, 51(6):1615–1624, 2003.

[Cla68] R. H. Clarke. A statistical theory of mobile-radio reception. *The Bell System Technical Journal*, 47(6):957–1000, July 1968.

[CSBM06] J. Coon, M. Sandell, M. Beach, and J. McGeehan. Channel and Noise Variance Estimation and Tracking Algorithms for Unique-Word based Single-Carrier Systems. *IEEE Transactions on Wireless Communications*, 5(6):1488–1496, 2006.

[CT08] T.-D. Chiueh and P.-Y. Tsai. *OFDM baseband receiver design for wireless communications*. John Wiley & Sons, 2008.

[CTL12] T.-D. Chiueh, P.-Y. Tsai, and I.-W. Lai. *Baseband receiver design for wireless MIMO-OFDM communications*. John Wiley & Sons, 2012.

[DdL17] J. T. Dias and R. C. de Lamare. Unique-Word GFDM Transmission Systems. *IEEE Wireless Communications Letters*, 6(6):746–749, 2017.

[DGE00] L. Deneire, B. Gyselinckx, and M. Engels. Training sequence vs. cyclic prefix a new look on single carrier communication. In *Globecom'00-IEEE. Global Telecommunications Conference. Conference Record (Cat. No. 00CH37137)*, volume 2, pages 1056–1060. IEEE, 2000.

[DNM+18] M. Danneberg, A. Nimr, N. Michailow, S. Ehsanfar, M. Matthé, A.-B. Martinez, D. Zhang, and G. Fettweis. Universal Waveforms Processor. In *2018 European Conference on Networks and Communications (EuCNC)*, pages 357–362. IEEE, 2018.

[EMCF19] S. Ehsanfar, M. Matthé, M. Chafii, and G. Fettweis. Pilot- and CP-Aided Channel Estimation in MIMO Non-Orthogonal Multi-Carriers. *IEEE Transactions on Wireless Communications*, 18(1):650–664, Jan 2019.

[Fet14] G. Fettweis. The Tactile Internet: Applications and Challenges. *IEEE Veh. Technol. Mag.*, 9(1):64–70, 2014.

[FRG+17] J. Ferreira, H. Rodrigues, A. Gonzalez, A. Nimr, M. Matthé, D. Zhang, L. Mendes, and G. Fettweis. GFDM Frame Design for 5G Application Scenarios. *Journal of Communication and Information Systems*, 32(1), Jul. 2017.

[Gen07] J. E. Gentle. Matrix Transformations and Factorizations. *Matrix Algebra: Theory, Computations, and Applications in Statistics*, pages 173–200, 2007.

[GMMF14] I. S. Gaspar, L. L. Mendes, N. Michailow, and G. Fettweis. A synchronization technique for generalized frequency division multiplexing. *EURASIP Journal on Advances in Signal Processing*, 2014(1):67, 2014.

[GMN+13] I. Gaspar, N. Michailow, A. Navarro, E. Ohlmer, S. Krone, and G. Fettweis. Low complexity GFDM receiver based on sparse frequency domain processing. In *Vehicular Technology Conference (VTC Spring), 2013 IEEE 77th*, pages 1–6. IEEE, 2013.

[HJ12] S. Hou and J. Jiang. Low Complexity Fast LMMSE-Based Channel Estimation for OFDM Systems in Frequency Selective Rayleigh Fading Channels. In *Proc. IEEE Vehicular Technology Conference (VTC Fall)*, pages 1–5, 2012.

[HL14] M. Huemer and O. Lang. On component-wise conditionally unbiased linear Bayesian estimation. In *Proc. of 48th Asilomar Conference on Signals, Systems and Computers*, 2014.

[HOH11] M. Huemer, A. Onic, and C. Hofbauer. Classical and bayesian linear data estimators for unique word ofdm. *IEEE Transactions on Signal Processing*, 59(12):6073–6085, 2011.

[HWH03] M. Huemer, H. Witschnig, and J. Hausner. Unique Word Based Phase Tracking Algorithms for SC/FDE-Systems. In *GLOBECOM'03. IEEE Global Telecommunications Conference (IEEE Cat. No. 03CH37489)*, volume 1, pages 70–74. IEEE, 2003.

[JBS00] M. Jeruchim, P. Balaban, and K. Shanmugan. *Simulation of Communication Systems, Second Edition*. New York, Kluwer Academic/Plenum, 2000.

[Kay93] S. M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*, volume 1. Upper Saddle River, NJ: Prentice Hall, 1993.

[KK16] A. K. Kohli and D. S. Kapoor. Adaptive Filtering Techniques Using Cyclic Prefix in OFDM Systems for Multipath Fading Channel Prediction. *Circuits Syst. Signal Process.*, 35(10):3595–3618, October 2016.

[KY02] K. J. Kim and J. Yue. Joint Channel Estimation and Data Detection Algorithms for MIMO-OFDM Systems. In *Conference Record of the Thirty-Sixth Asilomar Conference on Signals, Systems and Computers, 2002.*, volume 2, pages 1857–1861 vol.2, Nov 2002.

[MMG+14] N. Michailow, M. Matthé, I. S. Gaspar, A. N. Caldevilla, L. L. Mendes, A. Festag, and G. Fettweis. Generalized Frequency Division Multiplexing for 5th Generation Cellular Networks. *IEEE Trans. Commun.*, 62(9):3045–3061, 2014.

[MMG+16] M. Matthé, L. Mendes, I. Gaspar, N. Michailow, D. Zhang, and G. Fettweis. Precoded GFDM transceiver with low complexity time domain processing. *EURASIP Journal on Wireless Commun. Netw.*, 2016:138, 2016.

[MMGF14] M. Matthé, N. Michailow, I. Gaspar, and G. Fettweis. Influence of pulse shaping on bit error rate performance and out of band radiation of Generalized Frequency Division Multiplexing. In *Proc. IEEE Int. Conf. Commun. Workshops (ICC)*, pages 43–48, 2014.

[MP92] A. M. Mathai and S. B. Provost. Quadratic forms in random variables: theory and applications. *Journal of the American Statistical Association*, 1992.

[MS01] A. N. Mody and G. L. Stuber. Synchronization for MIMO OFDM systems. In *IEEE Global Telecommunications Conference (GLOBECOM'01)*, volume 1, pages 509–513. IEEE, 2001.

[MZF18] M. Matthé, D. Zhang, and G. Fettweis. Low-Complexity Iterative MMSE-PIC Detection for MIMO-GFDM. *IEEE Transactions on Communications*, 66(4):1467–1480, April 2018.

[NMZF17] A. Nimr, M. Matthé, D. Zhang, and G. Fettweis. Optimal Radix-2 FFT Compatible Filters for GFDM. *IEEE Communications Letters*, 21(7):1497–1500, 2017.

[Qiu95] S. Qiu. Block-circulant Gabor-matrix structure and discrete Gabor transforms. *Optical Engineering*, 34:2872–2872, 1995.

[QS10] A. A. Quadeer and M. S. Sohail. Enhanced channel estimation using Cyclic Prefix in MIMO STBC OFDM sytems. In *The 10th IEEE International Symp. on Signal Processing and Information Technology*, pages 277–282, Dec 2010.

[RC91] A. H. Roger and R. J. Charles. Topics in matrix analysis, 1991.

[SC97] T. M. Schmidl and D. C. Cox. Robust Frequency and Timing Synchronization for OFDM. *IEEE Trans. commun.*, 45(12):1613–1621, 1997.

[SH88] N. Suehiro and M. Hatori. Modulatable orthogonal sequences and their application to SSMA systems. *IEEE Trans. Inf. Theory*, 34(1):93–100, 1988.

[SYGO15] A. Sahin, R. Yang, M. Ghosh, and R. L. Olesen. An Improved Unique Word DFT-spread OFDM Scheme for 5G Systems. In *2015 IEEE Globecom Workshops (GC Wkshps)*, pages 1–6. IEEE, 2015.

[TC07] G. Taricco and G. Coluccia. Optimum Receiver Design for Correlated Rician Fading MIMO Channels with Pilot-Aided Detection. *IEEE J. Sel. Areas Commun.*, 25(7):1311–1321, 2007.

[TEF99] F. Tufvesson, O. Edfors, and M. Faulkner. Time and Frequency Synchronization for OFDM using PN-Sequence Preambles. In *Gateway to 21st Century Communications Village. VTC 1999-Fall. IEEE VTS 50th Vehicular Technology Conference (Cat. No. 99CH36324)*, volume 4, pages 2203–2207. IEEE, 1999.

[WMS+02] H. Witschnig, T. Mayer, A. Springer, A. Koppler, L. Maurer, M. Huemer, and R. Weigel. A different look on cyclic prefix for SC/FDE. In *The 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, volume 2, pages 824–828. IEEE, 2002.

[ZFF17] D. Zhang, A. Festag, and G. P. Fettweis. Performance of Generalized Frequency Division Multiplexing Based Physical Layer in Vehicular Communications. *IEEE Transactions on Vehicular Technology*, 66(11):9809–9824, 2017.

[ZMMF17] D. Zhang, M. Matthé, L. L. Mendes, and G. Fettweis. A Study on the Link Level Performance of Advanced Multicarrier Waveforms Under MIMO Wireless Communication Channels. *IEEE Transactions on Wireless Communications*, 16(4):2350–2365, 2017.

[ZZZW05] E. Zhou, X. Zhang, H. Zhao, and W. Wang. Synchronization algorithms for MIMO OFDM systems. In *IEEE Wireless Communications and Networking Conference, 2005*, volume 1, pages 18–22. IEEE, 2005.