Christian Hofbauer

Design and Analysis of Unique Word OFDM

DISSERTATION

zur Erlangung des akademischen Grades

Doktor der Technischen Wissenschaften

Alpen-Adria-Universität Klagenfurt
Fakultät für Technische Wissenschaften

Betreuer
Prof. Dr. techn. Mario Huemer
Johannes-Kepler-Universität Linz
Institut für Signalverarbeitung

Erstgutachter
Prof. Dr. techn. Mario Huemer
Johannes-Kepler-Universität Linz
Institut für Signalverarbeitung

Zweitgutachter
Prof. Dr.-Ing. Dr.-Ing. habil. Johannes Huber
Universität Erlangen-Nürnberg
Lehrstuhl für Informationsübertragung

Klagenfurt, 06/2016

Affidavit

I hereby declare in lieu of an oath that

- the submitted academic paper is entirely my own work and that no auxiliary materials have been used other than those indicated,
- I have fully disclosed all assistance received from third parties during the process of writing the paper, including any significant advice from supervisors,
- any contents taken from the works of third parties or my own works that have been included either literally or in spirit have been appropriately marked and the respective source of the information has been clearly identified with precise bibliographical references (e.g. in footnotes),
- to date, I have not submitted this paper to an examining authority either in Austria or abroad and that
- the digital version of the paper submitted for the purpose of plagiarism assessment is fully consistent with the printed version.

I am aware that a declaration contrary to the facts will have legal consequences.

(Signature) (Place, Date)

Abstract

Unique Word orthogonal frequency division multiplexing (OFDM) is a promising alternative to cyclic prefix based OFDM (CP-OFDM), which currently denotes the method of choice for many digital communication standards, with applications ranging from audio and video broadcasting, last mile internet access to modern cellular networks. In this signaling concept, the guard interval (GI) is filled with an arbitrary deterministic sequence — the so-called "unique word" (UW) — instead of the random CP. This sequence provides the same advantages as a CP (no intersymbol interference and diagonalization of the channel matrix), but can additionally be designed to optimally meet synchronization and estimation tasks. Furthermore, most important, and different to almost all signaling schemes of the OFDM family, the UW is already part of the discrete Fourier transform (DFT) interval. Ensuring such time domain properties entails the introduction of a certain redundancy in the frequency domain. This redundancy can be exploited beneficially to enhance range, reliability, capacity or battery lifespan. In this sense, UW-OFDM transforms the usually disregarded guard interval into a multipurpose sequence, thus tackling the well-known inefficiency problem of guard intervals in current communication systems. Moreover, adapting the UW and therefore the GI length to different channel conditions will not impact the DFT length and thus keeps relevant processing chain structures untouched. Hence, UW-OFDM allows supporting a wide range of communication scenarios while still ensuring high efficiency.

The implementation of the inherent redundancy — the primary root of the special UW-OFDM properties — is ambiguous, giving rise to a variety of different signal variants. Main topic of this work is the investigation of the signal generation process of UW-OFDM symbols and its impact on the performance.

The first principle approach of generating OFDM symbols with unique word is the concept of systematically encoded UW-OFDM. This concept is based on the idea of a systematic block code, leading to dedicated data and redundant subcarriers. This redundancy translates to beneficial properties regarding spectral behavior and bit error ratio (BER) performance. Systematically encoded UW-OFDM shows a superior sidelobe suppression over conventional CP-OFDM and outperforms it in terms of the BER performance in a multipath environment, for coded as well as uncoded transmission. Still, the required energy to load the redundant subcarriers is significantly higher than for the data subcarriers, showing further potential for improvement. Introducing additional redundant subcarriers or allowing systematic

noise in the guard interval reduce the OFDM symbol energy further. The resulting BER enhancement comes at the price of either reduced bandwidth efficiency or an inevitable error floor.

The second principle approach of generating OFDM symbols with unique word is denoted as non-systematically encoded UW-OFDM, which annihilates the flaws of the systematic approach. Inspired by a non-systematic code, the idea of dedicated data and redundant subcarriers is discarded and all-purpose subcarriers introduced instead. A non-systematic generation translates to an additional gain in terms of spectral properties and BER behavior. BER results are obtained for various setups, channel conditions, data symbol constellations and imperfect channel knowledge, all delivering results significantly in favor of UW-OFDM.

An extension of the UW-OFDM framework enables the inclusion of pilot tones into the frequency domain symbol, while still preserving all beneficial properties known from a pilotless UW-OFDM concept. A mean square error (MSE) analysis of an exemplary carrier frequency offset (CFO) estimation task reveals a significant better performance of pilot tone based estimation concepts in UW-OFDM than in CP-OFDM, a result inherited from the introduced redundancy.

The effects of a CFO on UW-OFDM are studied in detail and compared to those in single-carrier and OFDM signaling schemes. The CFO induced error after data estimation is on the one hand due to subtracting a disturbed UW and pilot offset, and on the other hand due to insufficient CFO compensation. Both error sources are investigated independently and alternative approaches with different computational complexity are presented. MSE results confirm UW-OFDM to achieve a higher robustness against CFO than conventional CP-OFDM. BER simulations additionally illuminate the effects of CFO impairments, again identifying UW-OFDM as the clearly better performing alternative.

Contents

1. Introduction **3**
- 1.1. State of the Art . . . 4
- 1.2. Scope of this Work . . . 6

2. The Unique Word OFDM Signaling Scheme **9**
- 2.1. Unique Word OFDM Symbol Generation . . . 16
- 2.2. System Model . . . 22
- 2.3. Linear Estimators . . . 23
  - 2.3.1. Classical Data Estimators - Zero Forcing Solutions . . . 24
  - 2.3.2. Linear Bayesian Data Estimators – LMMSE Solutions . . . 25
- 2.4. Simulation Setup . . . 26
  - 2.4.1. Schematic of Simulation Chain . . . 26
  - 2.4.2. Multipath Channel Model . . . 32
  - 2.4.3. Simulation . . . 33

3. Systematically Encoded Unique Word OFDM **37**
- 3.1. Signal Properties . . . 40
- 3.2. Addition of UW: Two-Step versus Direct Approach . . . 50
  - 3.2.1. Symbol Energy for Direct Approach . . . 51
  - 3.2.2. Symbol Energy for Two-Step Approach . . . 52
  - 3.2.3. Comparison of Both Approaches . . . 54
- 3.3. Optimization Algorithm for Redundant Subcarrier Distribution . . . 55
  - 3.3.1. Split Distribution . . . 56
  - 3.3.2. Quasi-Uniform Distribution . . . 57
  - 3.3.3. Heuristic Distribution . . . 58
- 3.4. Alternative Optimization Criteria . . . 60
- 3.5. Performance Evaluation . . . 63
  - 3.5.1. Principles of UW-OFDM . . . 63
  - 3.5.2. Power Spectral Density . . . 69
  - 3.5.3. Bit Error Ratio Simulations . . . 72
- 3.6. Introduction of Additional Redundant Subcarriers . . . 74
- 3.7. Systematic Noise in UW-OFDM . . . 79
  - 3.7.1. Systematic Noise in the Guard Interval . . . 82
  - 3.7.2. Systematic Noise on the Data Subcarriers . . . 86

4. Non-Systematically Encoded Unique Word OFDM **91**
4.1. Signal Properties . . . 93
4.2. Optimization Algorithm . . . 99
4.3. Performance Evaluation . . . 103
4.3.1. Power Spectral Density . . . 103
4.3.2. Bit Error Ratio Simulations . . . 104
4.3.3. Impact of Setup and UW . . . 110
4.3.4. Impact of Imperfect Channel Knowledge . . . 117

5. Inclusion of Pilot Tones in Unique Word OFDM **123**
5.1. Optimization of Data Generator Matrix . . . 124
5.2. Optimization of Pilot Generator Matrix . . . 125

6. Carrier Frequency Offset **129**
6.1. CFO Model . . . 131
6.2. Receiver Model . . . 143
6.3. Impact of Imperfect UW and Pilot Offset Subtraction . . . 151
6.3.1. Individual UW and Pilot Offset Compensation . . . 152
6.3.2. Common UW and Pilot Offset Compensation . . . 155
6.4. Impact of Imperfect CFO Compensation . . . 157
6.5. Estimation of Phase Rotation and Carrier Frequency Offset . . . 158
6.5.1. Phase Offset Compensation . . . 164
6.5.2. Performance Evaluation . . . 167
6.5.3. Estimation Setup . . . 170
6.6. Bit Error Ratio Simulations . . . 170

7. Conclusion **179**

A. LLR calculation **181**

B. Derivation of the Global Minimum of J_{BLUE} **187**

C. Derivation of the Global Minimum of J_{LMMSE} **189**

List of Abbreviations **194**

Bibliography **195**

1. Introduction

The recent past has shown an omnipresent demand in society of staying connected at any time and any place, while having access to any information without limitation. These observations pushed the increase of data rates to one of the main dominating factors in developing digital communication systems. Due to the expectation of full connectivity regardless of time and space, providing high data rates is crucial for any type of communication system, from wireline to wireless, from non-optical to optical, from fixed to mobile. As predicted by Moore's law, semiconductor industry has constantly increased the transistor density on integrated circuits and therefore laid the technical groundwork to computationally handle these rates. The biggest challenge in supporting high data rates with sufficient robustness originates from channel impairments, with the dispersive properties of a propagation channel determining a main source. In wireless communication scenarios, time dispersity emerges from a transmit signal travelling along multiple paths to the receiver due to reflections, diffusions, diffractions and scattering at buildings, trees, or any other obstacle. This causes signals spread in time, leading to interferences among initially independent data symbols. The effects of interference become more challenging in high rate communication systems, as the time interval to transmit one symbol is small. Real-time requirements on transmission even aggravate terms, posing computational challenges on the countermeasures.

A popular means to cope with channel dispersity induced interferences while meeting processing time constraints at reasonable complexity is the concept of block based transmission. The idea is predicated on grouping together several data symbols and modulating them in a defined way on a transmit block or transmit symbol. This operation is conducted for all data symbols. The transmit symbols are then separated by so-called guard intervals. Guard intervals collect the dispersive effects of the channel to prevent intersymbol interference (ISI), which in turn enables independent processing of the blocks.

The most popular scheme to modulate a transmit block is orthogonal frequency division multiplexing (OFDM) [1], which loads the information on narrowband orthogonal subcarriers within the available bandwidth; a success story mainly founded on its advantages of simple processing as well as efficient handling of severe multipath propagation and narrowband interferences (in combination with channel coding). First ideas of parallel data streams and frequency division multiplexing have already been published in the 1960s [2, 3]. However, it took till the early eighties [4, 5], until hardware became capable of handling OFDM, which

finally led to its breakthrough in the nineties [6]. Today, this technique is part in many modern digital communication standards. Its field of application spans from audio and video broadcasting DAB and DVB [7, 8], over wired [9, 10] as well as wireless [11] last mile internet access techniques, to modern cellular communication standards such as LTE [12].

Despite its advantages, OFDM – as well as many other block based transmission techniques – requires the insertion of guard intervals to eliminate intersymbol interference. Taking into account a typical guard interval length of up to 25% of the transmit symbol duration, it is a rather wasteful consumer of transmit time and thus reduces bandwidth efficiency significantly. Furthermore, depending on the sequence transmitted during the guard interval, it also wastes transmit energy, an aspect which gets more dominant for battery-driven devices.

Many ideas have been proposed addressing the inefficiency problem of guard intervals, the most prominent ones are cited in the following section. Some of these concepts fill the guard interval with a (pseudo)random, others with a deterministic sequence. Unique Word OFDM promises to surmount all concepts available so far by filling the guard interval also with a deterministic sequence, but in a completely different manner. This difference translates to many nice properties regarding system relevant aspects, such as improved bit error ratio (BER) behavior or beneficial spectral properties.

1.1. State of the Art

The introduction of guard intervals between consecutive OFDM symbols is an effective means to eliminate intersymbol interference among those symbols. Furthermore, implementing these guard intervals as cyclic prefixes (CPs) additionally eliminates intercarrier interference. This cyclic extension transforms the linear convolution of the transmit signal with a dispersive channel into a cyclic convolution, enabling a blockwise low complexity equalization in the frequency domain. CP-OFDM is currently the most popular multi-carrier signaling scheme and the method of choice in many communication standards, as such denoting the gold standard test for all other signaling schemes.

Besides equalization, the purpose of a random cyclic prefix is rather limited. Known symbol padding OFDM (KSP-OFDM) promises to overcome this flaw by filling the guard interval with a deterministic sequence instead. This sequence may be selected to optimally match system parameter estimation tasks, providing thus a second use for the guard interval. Removing the effects of the known sequence (taking into account the channel dispersion) at the receiver and adding the transient to the beginning of the OFDM block, cyclicity as in CP-OFDM is obtained, enabling simple frequency domain equalization again. KSP-OFDM and UW-OFDM share the opportunity of a deterministic sequence designed for synchronization and estimation

tasks, but beyond that differ completely. In UW-OFDM the guard interval is part of the discrete Fourier transform (DFT) interval, introducing therefore correlations in the frequency domain, which can be beneficially exploited to improve spectral properties or the bit error ratio performance. These correlations are not present in KSP-OFDM.

Setting the known sequence in KSP-OFDM to zero yields zero padded OFDM (ZP-OFDM) [13]. Consequently, ZP-OFDM does not incorporate any correlations and thus shows the same shortcomings over UW-OFDM as KSP-OFDM.

Various other approaches labeled time domain synchronous OFDM (TDS-OFDM) [14, 15, 16], pseudorandom prefix OFDM (PRP-OFDM) [17], OFDM with pseudo noise (PN) sequence [16] or even OFDM with Unique Word [18] implement deterministic sequences in the guard interval. Differing from each other in the specific instance of the sequence, all implement the guard interval outside of the DFT interval. Hence, no correlations as in UW-OFDM are present.

The generation of a UW in time domain within the DFT interval introduces redundancy in the frequency domain. From a coding theory point of view, UW-OFDM time domain symbols contain with the UW a block of fixed samples, hence the set of all corresponding vectors in discrete frequency domain forms in fact a coset to a Reed Solomon (RS) code. Usually RS codes of length n are defined for a finite field with a suitable discrete Fourier transform. The set of code words is specified by the fact, that the (inverse) DFT of all code words contains a block of $d_{\min} - 1$ successive zeros, where $d_{\min}$ is the minimum Hamming distance of the RS code. If this block of $d_{\min} - 1$ successive symbols differs from zero, but is also fixed for all code words, a coset code to an RS code is generated in the other domain w.r.t. this Fourier transform with the same minimum distance $d_{\min}$, cf. [19]. This RS property suggests to apply algebraic RS decoding methods. Investigations in [20, 21] identify this decoding approach as solving a very ill-conditioned system of equations, hence motivating methods based on classical and Bayesian estimation theory used in this work as a preferable alternative.

The idea of UW-OFDM is inspired by its pendant in single-carrier/frequency domain equalization (SC/FDE) based systems. Various publications point out the potential of UWs [22, 23, 24, 25, 26] instead of CPs [27, 28]. The introduction of UWs in SC/FDE systems does not impose any challenges, as the data symbols and the UW are simply placed consecutively to each other in time domain. In UW-OFDM, the UW is still defined in time domain, but the data symbols are arranged in frequency domain, preventing a straightforward implementation. Investigations in [29] compare the UW approach for OFDM and SC/FDE, [30, 31] identify the latter as a specific realization within the UW-OFDM framework.

First attempts of transferring the concept of unique words to OFDM have already been proposed in [32] termed discrete multi-tone - known symbol padding

(DMT-KSP). The presented method corresponds to the direct approach described in Sec. 3.2.1 and suffers from an unmanageable high transmit symbol energy, making it inapplicable for practical systems. Furthermore, the particular UW waveform influences the data estimation performance and thus loses its advantage of being solely selectable to perfectly match synchronization and parameter estimation tasks.

1.2. Scope of this Work

Currently in 2016, OFDM plays a major role in almost every digital communication standard. Despite its popularity in recent years, research on this topic has declined worldwide based on the premise of having said everything already. A well-known and excepted inefficiency of OFDM is the requirement for a guard interval. Unique Word OFDM promises to overcome this flaw by implementing a multipurpose guard interval providing various beneficial properties. This work is dedicated to the signal generation process of Unique Word OFDM symbols in order to optimally exploit the guard intervals needed in OFDM signaling schemes.

The principles of the signaling scheme Unique Word OFDM are introduced in the second chapter. A mathematical description is developed, yielding a linear model describing the processing chain from modulation to the data estimation at the receiver. Various estimators based on classical as well as Bayesian estimation theory are introduced as a preparatory step for subsequent chapters. Since the focus of this work is on the UW-OFDM framework itself rather than on details in the estimation process, considerations are limited to the class of linear estimators. The chapter closes with a detailed description of the simulation framework and the corresponding relevant system parameters utilized throughout this work.

Chapter 3 introduces the concept of systematically encoded UW-OFDM, the first of two principle approaches to implement OFDM with Unique Word. This concept shares similarities with a systematic code, leading to dedicated data and redundant subcarriers. A thorough analysis of the signal properties inherent to this concept serves as entry point. An optimum distribution of the redundant subcarriers among the available spectrum is identified as the key to obtain a reasonable OFDM symbol energy and consequently make UW-OFDM feasible for practical systems. The presented heuristic solver is able to deliver this distribution independent of the specific setup. Experiments prove a minimization of the symbol energy to be also optimum w.r.t. the whole transceiver performance of systematically encoded UW-OFDM. Furthermore, two variants of the UW generation and their impact on the bit error ratio are presented. Introducing additional redundant subcarriers or allowing systematic noise in the guard interval reduce the OFDM

symbol energy further. The resulting BER improvement comes at the price of reduced bandwidth efficiency for the first and an inevitable error floor for the second case.

Non-systematically encoded UW-OFDM detailed in chapter 4 annihilates the flaws of the systematic concept. Analogously to a non-systematic code, the idea of dedicated data and redundant subcarriers is discarded and all-purpose subcarriers introduced instead. Again, a summary of the signal properties form the entry. Analytical derivations prove this approach to be optimum w.r.t. the whole transceiver performance of non-systematically encoded UW-OFDM, a numerical optimization algorithm provided in this chapter delivers the corresponding generator matrices. The non-systematic approach outperforms both, the systematic approach as well as CP-OFDM, significantly in terms of spectral properties and BER behavior. Results are obtained for various setups. An analysis of imperfect channel knowledge at the receiver reveals approximately the same performance degradation for UW-OFDM as for CP-OFDM.

Chapter 5 deals with pilot tone insertion in UW-OFDM frequency domain symbols. Since this is straightforward in systematically encoded UW-OFDM, the main focus is laid on non-systematically encoded UW-OFDM. Reasonable optimization criteria are presented and possible optimization parameters identified. Furthermore, interactions between frequency pilots for system parameter estimation on the one hand and data estimation on the other hand are analyzed. The investigations confirm optimization possibilities towards both directions at the same time.

Carrier frequency offset (CFO) – known as one of the most critical impairments to multi-carrier systems – is investigated in chapter 6. A thorough mathematical description enables a detailed analysis of its effects on UW-OFDM, insights are provided in time and frequency domain. The effects experienced by UW-OFDM correspond to a combination of those for single-carrier and conventional multi-carrier based systems, the specific UW-OFDM realization determines a bias towards one or the other system. The derived model covers the processing chain from modulation till the input of the data/pilot estimator. The CFO induced error after data estimation is on the one hand due to subtracting the wrong UW and pilot offset, and on the other hand due to insufficient CFO compensation. Both error sources are investigated independently and alternatives are presented. A mean square error (MSE) analysis confirms UW-OFDM to achieve a higher robustness against CFO than conventional CP-OFDM. For CFO estimation, a frequency pilot tone based estimation method is presented and differences to CP-OFDM are highlighted. BER simulations for coded and uncoded transmission in a multipath environment finally conclude the CFO investigations.

The following mathematical notation applies throughout this work: Vectors and matrices are denoted in bold face lower case $\mathbf{a}$ and upper case letters $\mathbf{A}$, respectively. If a variable is represented in both time and frequency domain, a tilde is

used for the latter to clearly distinguish between both representations (e.g., $\mathbf{x}$ for the time and $\tilde{\mathbf{x}}$ for the frequency domain). This distinction is omitted, if a variable is defined in only one domain, consequently an interpretation is provided from the context. The kth element of a vector is named a_k. For reasons of simplified notation, the nomenclature $a[k]$ is used for the kth element in chapter 6 instead. The kth column of a matrix consisting of N columns $\mathbf{A} = [\mathbf{a}_0, \mathbf{a}_1, \ldots, \mathbf{a}_{N-1}]$ is given by $\mathbf{a}_k$. To address the element in column k and row l of the same matrix, the notation $[\mathbf{A}]_{k,l}$ applies. Following this notation, $[\mathbf{A}]_{k,*}$ represents all elements of row number k. The transpose operation is expressed as $(\cdot)^T$, the conjugate as $(\cdot)^*$ and the conjugate transpose or Hermitian as $(\cdot)^H$. Further, $(\cdot)^\dagger$ is used to denote the Moore-Penrose Pseudo-Inverse of a matrix, $\mathrm{tr}\,(\mathbf{A})$ to denote the trace operation, $\mathrm{diag}\,(\mathbf{A})$ to extract the main diagonal entries of a matrix $\mathbf{A}$ and $\mathrm{flip}\{\mathbf{a}\}$ to flip the elements of a vector in vertical direction. The identity matrix is given by $\mathbf{I}$ and a zero matrix as $\mathbf{0}$. $\mathrm{E}\{\cdot\}$ denotes expectation, $[\cdot]_R$ represents rounding to the nearest integer, $\mathrm{Re}\{\cdot\}$ returns the real and $\mathrm{Im}\{\cdot\}$ the imaginary part of a complex number. $\mathbb{N}$, $\mathbb{R}$ and $\mathbb{C}$ represent the set of natural, real, and complex numbers, respectively. The probability of an event is expressed as $\Pr(\cdot)$, and a probability density function (PDF) as $p(\cdot)$. Further, $\hat{\mathbf{a}}$ stands for an estimate of $\mathbf{a}$.

For all signals and systems in this work, the equivalent complex baseband representation according to [33] applies.

2. The Unique Word OFDM Signaling Scheme

This chapter starts with introducing the principal idea of blockwise transmission techniques based on the insertion of guard intervals (GIs). The cyclic prefix (CP) concept as the most important representative is elaborated further to motivate the main topic of this work: Unique Word OFDM. First considerations focus on the symbol generation and already indicate various challenges in symbol design, justifying the necessity of this work. As a preparatory step for the subsequent chapters, a linear system model describing the UW-OFDM transmission scheme and corresponding linear estimators are presented. The chapter concludes with the presentation of the simulation setups used throughout this work.

Fig. 2.1 provides a discretized basic time-invariant channel model used for describing a typical communication scenario. A signal $x[k]$ shall be transmitted over a

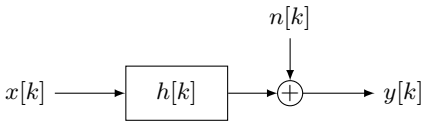


Figure 2.1.: Model of channel propagation.

dispersive propagation channel. This channel is modeled as a multipath environment, which in turn is implemented as a tapped delay line with the channel impulse response (CIR) as its coefficients. Further, transmission is disturbed by additive white Gaussian noise (AWGN). The resulting signal at the receiver $y[k]$ can be expressed as

$$\begin{aligned} y[k] &= x[k] * h[k] + n[k] \\ &= \sum_{n=0}^{\infty} h[n]x[k-n] + n[k], \end{aligned} \tag{2.1}$$

with the CIR $h[k]$, the noise sequence $n[k]$ and the transmit signal $x[k]$. The dispersive nature of the channel results in a linear convolution – denoted by the operator '*' – between the CIR and the transmit signal. For reasons of easier processing, the transmit signal is usually divided into independent transmit blocks

or transmit symbols[1] of length N

$$x^{(l)}[k] = x[k + lN], \quad k = 0, 1, \ldots, N-1. \tag{2.2}$$

Each transmit symbol carries a data sequence $d[k]$ of length N_d, whereas $N_d = N$ is assumed for reasons of simplicity. Unfortunately, the dispersive properties of the channel cause the transient of one symbol ranging into the succeeding symbol illustrated as decaying curve in Fig. 2.2. This phenomenon known as ISI[2] precludes

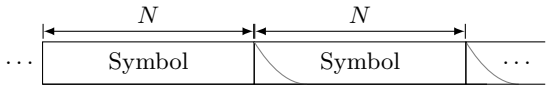


Figure 2.2.: Intersymbol interference caused by a multipath channel.

the original idea of independent symbolwise processing. A common solution is the insertion of guard intervals of length N_g between consecutive transmit symbols, that will be able to fully collect the transient, as long as the CIR does not outlast the GI, cf. Fig. 2.3.

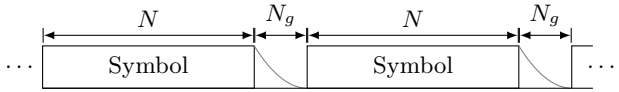


Figure 2.3.: Insertion of guard intervals between consecutive transmit symbols.

So far, the guard interval only fulfills the purpose of preventing ISI, one symbol considered individually is still affected by the channel

$$y^{(l)}[k] = x^{(l)}[k] * h[k] + n^{(l)}[k]. \tag{2.3}$$

The sequence $n^{(l)}[k]$ models additive noise as part of the lth receive sequence. Note that in contrast to $x^{(l)}[k]$ and $x[k]$, a concatenation of $n^{(l)}[k]$ is not equivalent to $n[k]$ (only parts would coincide), as the new model in Fig. 2.3 accounts for additional noise samples within the guard interval. However, this does not affect the following derivations.

In order to recover the transmitted sequence again, the dispersions caused by the channel have to be revoked, an operation commonly known as *equalization*. There are several ways which differ in the specific realization of the GI. All approaches

[1] Both terms – symbol as well as block – are common in the literature to denote the same partitioning of the transmit signal into smaller units. The decision towards one or the other term often arises from the context. In this work, the term *symbol* is the preferred choice.

[2] In the context of transmit blocks instead of transmit symbols, the corresponding term used in the literature is inter-block interference (IBI).

share the motivation though, that at a certain data rate a (linear) equalization in frequency domain provides in general a better performance–complexity tradeoff than in time domain [25]. The most common approach is the use of a *cyclic prefix*, which will be detailed in the following.

Convolution theorem The Fourier transform of a *linear* convolution of two (continuous) signals in time domain translates to a pointwise multiplication of the single Fourier transforms. However, for a time discrete signal with the discrete Fourier transform as corresponding operation, the pointwise multiplication in frequency domain coincides with a *cyclic* convolution in time domain [34]

$$y[k] = x_1[k] \overset{N}{\circledast} x_2[k] \quad \circ\!\!-\!\!\bullet \quad \tilde{y}[n] = \tilde{x}_1[n] \cdot \tilde{x}_2[n], \tag{2.4}$$

$$\tag{2.5}$$

related to sequences of length N. The time discrete cyclic convolution of length N is defined as

$$y[k] = x_1[k] \overset{N}{\circledast} x_2[k] \tag{2.6}$$

$$= \sum_{n=0}^{N-1} x_1[n] x_2[(k-n) \bmod N], \tag{2.7}$$

$$k = 0, 1, \ldots, N-1. \tag{2.8}$$

Unfortunately, the convolution of the transmit signal $x^{(l)}[k]$ with the CIR $h[k]$ in (2.3) is linear. In order to exploit a simple multiplicative relation in frequency domain, the *linear* convolution should appear *cyclic*. The most popular way to achieve this is the cyclic extension of a symbol by copying the last N_g samples and appending them in front of the symbol, as shown in Fig. 2.4. This copy is known

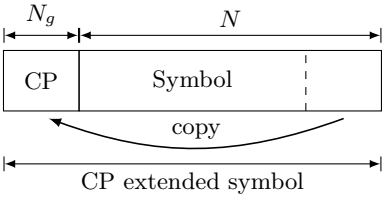


Figure 2.4.: Extension of a transmit symbol using a cyclic prefix.

as CP which enforces

$$y^{(l)}[k] = x^{(l)}[k] \overset{N}{\circledast} h[k] + n^{(l)}[k], \quad k = 0, 1, \ldots, N-1. \tag{2.9}$$

Following the cyclic convolution theorem, the receive sequence in frequency domain is equivalently given as

$$\tilde{y}^{(l)}[n] = \tilde{x}^{(l)}[n] \cdot \tilde{h}[n] + \tilde{n}^{(l)}[n], \quad n = 0, 1, \ldots, N-1. \tag{2.10}$$

An equalization of the channel dispersions can therefore be achieved by an element-wise multiplication of the frequency domain signal $\tilde{y}^{(l)}[n]$ with $1/\tilde{h}[n]$.

In order to demonstrate the transform of the linear into a cyclic convolution based on the cyclic prefix, a vector and matrix notation is introduced in the following.

Each transmit symbol $\mathbf{x}^{(l)} \in \mathbb{C}^{N\times 1}$ shall carry $N_d = N$ data symbols grouped together in a vector $\mathbf{d}^{(l)} \in \mathbb{C}^{N\times 1}$. The way how the data symbols are mapped on a transmit symbol shall mathematically be described by a matrix $\mathbf{G}$ of dimensions $N \times N$ such that

$$\mathbf{x}^{(l)} = \mathbf{G}\mathbf{d}^{(l)}. \tag{2.11}$$

Depending on the design of $\mathbf{G}$, different classes of transmission techniques arise, a very popular categorization follows from the arrangement of the data symbols[3], yielding

- *single-carrier* and
- *multi-carrier* transmission.

In single-carrier transmission, data symbols are placed next to each other in time domain, which corresponds to setting $\mathbf{G} = \mathbf{I}$ in (2.11). In multi-carrier transmission, however, data symbols are arranged adjacently in frequency domain. One of the most prominent candidates out of the multi-carrier family is OFDM, which can be modeled as $\mathbf{G} = \mathbf{F}_N^{-1}$, where $\mathbf{F}_N \in \mathbb{C}^{N\times N}$ and $\mathbf{F}_N^{-1} \in \mathbb{C}^{N\times N}$ denote the DFT and IDFT matrices with elements

$$\begin{aligned} \left[\mathbf{F}_N\right]_{kl} &= \mathrm{e}^{-j\frac{2\pi}{N}kl}, \quad k, l = 0, 1, ..., N-1 \\ \text{and } \left[\mathbf{F}_N^{-1}\right]_{kl} &= \frac{1}{N}\mathrm{e}^{j\frac{2\pi}{N}kl}. \end{aligned} \tag{2.12}$$

Note that the actual design of $\mathbf{G}$ is without relevance for the transform to a cyclic convolution, a detailing of $\mathbf{G}$ at this point should only demonstrate the general validness of this concept for various transmission techniques.

Besides (2.12), very common as well is the definition of *unitary* and therefore energy-invariant DFT and IDFT matrices

$$\begin{aligned} \left[\mathbf{F}'_N\right]_{kl} &= \frac{1}{\sqrt{N}}\mathrm{e}^{-j\frac{2\pi}{N}kl} \\ \text{and } \left[{\mathbf{F}'_N}^{-1}\right]_{kl} &= \frac{1}{\sqrt{N}}\mathrm{e}^{j\frac{2\pi}{N}kl}, \end{aligned} \tag{2.13}$$

[3] Note that aspects such as pulse shaping are neglected here.

featuring the advantageous property

$$\mathbf{F}'_N {\mathbf{F}'_N}^H = \mathbf{I} \quad \Rightarrow \quad {\mathbf{F}'_N}^{-1} = {\mathbf{F}'_N}^H. \tag{2.14}$$

The definition does not have further impact on the operation of OFDM or other mathematical coherence in this work, but in order to coincide with recent publications covering UW-OFDM, this work uses the definition as in (2.12), such that

$$\mathbf{F}_N^{-1} = \frac{1}{N}\mathbf{F}_N^H. \tag{2.15}$$

Two succeeding and extended transmit symbols with index $l-1$ and l are given by

$$\begin{bmatrix} \mathbf{x}_{\text{ext}}^{(l-1)} \\ \mathbf{x}_{\text{ext}}^{(l)} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_{\text{GI}}^{(l-1)} \\ \mathbf{x}^{(l-1)} \\ \mathbf{x}_{\text{GI}}^{(l)} \\ \mathbf{x}^{(l)} \end{bmatrix}, \tag{2.16}$$

each consisting of a guard interval $\mathbf{x}_{\text{GI}} \in \mathbb{C}^{N_g \times 1}$ and a payload part $\mathbf{x} \in \mathbb{C}^{N \times 1}$. Note that the derivations are kept as general as possible, the specific case of realizing the GI as CP is introduced later. Assuming a CIR not exceeding a length[4] of $N_g + 1$ and neglecting AWGN, the transient of the convolution between symbol $l-1$ and the CIR is collected entirely by the guard interval of symbol l. Therefore, the payload of symbol l is not affected by symbol $l-1$ and $\mathbf{x}_{\text{ext}}^{(l-1)}$ can thus be disregarded in the following considerations. Hence, the considered receive vector of length $(N + 2N_g)$ for symbol l after propagating over a multipath channel is given by

$$\mathbf{y}_{\text{conv}}^{(l)} = \mathbf{H}_{\text{conv}} \begin{bmatrix} \mathbf{x}_{\text{GI}}^{(l)} \\ \mathbf{x}^{(l)} \end{bmatrix}, \tag{2.17}$$

with the linear convolution matrix of size $(N+2N_g)\times(N+N_g)$ defined as

$$\mathbf{H}_{\text{conv}} = \begin{bmatrix} h_0 & 0 & \cdots & \cdots & 0 \\ h_1 & h_0 & \ddots & & \vdots \\ \vdots & h_1 & \ddots & \ddots & \vdots \\ h_{N_g} & \vdots & \ddots & \ddots & 0 \\ 0 & h_{N_g} & & \ddots & h_0 \\ \vdots & \ddots & \ddots & & h_1 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & h_{N_g} \end{bmatrix}. \tag{2.18}$$

[4] The length of a guard interval is chosen at design time and adapted according to prevalent channel characteristics. Hence, the condition of a CIR not exceeding the guard interval is generally assumed to be fulfilled.

The transient of the convolution stored in the first N_g and last N_g samples of $\mathbf{y}_{\text{conv}}$ will not be required for recovering the payload and is thus omitted by a selection matrix

$$\mathbf{\Theta}_s = \begin{bmatrix} \mathbf{0}^{(N\times N_g)} & \mathbf{I}^{(N)} & \mathbf{0}^{(N\times N_g)} \end{bmatrix}, \tag{2.19}$$

yielding a receive vector of length N

$$\mathbf{y} = \mathbf{\Theta}_s \mathbf{y}_{\text{conv}}^{(l)} = \underbrace{\mathbf{\Theta}_s \mathbf{H}_{\text{conv}}}_{\mathbf{H}'} \begin{bmatrix} \mathbf{x}_{\text{GI}}^{(l)} \\ \mathbf{x}^{(l)} \end{bmatrix}. \tag{2.20}$$

Matrix $\mathbf{H}'$ is of size $N \times (N + N_g)$ and follows from $\mathbf{H}_{\text{conv}}$ by deleting the first N_g and last N_g rows. Since subsequent considerations are restricted to a single symbol, the superscript (l) is dropped in the following for the sake of simplified notation. For further analysis, an elementwise notation of (2.20) reads as

$$\begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ \vdots \\ \vdots \\ y_{N-1} \end{bmatrix} = \begin{bmatrix} h_{N_g} & \cdots & h_0 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & & \ddots & \ddots & & & \vdots \\ \vdots & & \ddots & & \ddots & \ddots & & \vdots \\ \vdots & & & \ddots & & \ddots & \ddots & \vdots \\ \vdots & & & & \ddots & & \ddots & 0 \\ 0 & & & & 0 & h_{N_g} & \cdots & h_0 \end{bmatrix} \cdot \begin{bmatrix} x_{\text{GI},0} \\ \vdots \\ x_{\text{GI},(N_g-1)} \\ x_0 \\ x_1 \\ \vdots \\ \vdots \\ \vdots \\ x_{N-1} \end{bmatrix}. \tag{2.21}$$

So far, this general transmit-receive relation holds true for any system using guard intervals[5]. However, implementing the guard interval as a cyclic prefix such that

$$x_{\text{GI},k} = x_{N-N_g+k} \qquad\qquad k = 0 \ldots N_g - 1, \tag{2.22}$$

equation (2.21) can be written in the special form

$$\mathbf{y} = \mathbf{H}_c \mathbf{x} \tag{2.23}$$

[5] Note that for ZP-OFDM or KSP-OFDM, the proposed model in (2.21) is incomplete. In order to obtain cyclicity, only the transient at the beginning of $\mathbf{y}_{\text{conv}}$ can be chopped off, the transient at the end is still required. However, this is easily achieved by adapting the selection matrix $\mathbf{\Theta}_s$ accordingly.

$$\begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ \vdots \\ \vdots \\ y_{N-1} \end{bmatrix} = \begin{bmatrix} h_0 & 0 & \cdots & 0 & h_{N_g} & \cdots & h_1 \\ h_1 & h_0 & \ddots & & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & & \ddots & h_{N_g} \\ h_{N_g} & & & \ddots & \ddots & & 0 \\ 0 & \ddots & & & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & & \ddots & 0 \\ 0 & \cdots & 0 & h_{N_g} & \cdots & \cdots & h_0 \end{bmatrix} \cdot \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ \vdots \\ \vdots \\ x_{N-1} \end{bmatrix}. \tag{2.24}$$

Notationally speaking, a cyclic data structure as induced by the cyclic prefix allows incorporating the first N_g columns of $\mathbf{H}'$ into the N_g last ones, resulting in a circulant matrix $\mathbf{H}_c$ of dimension $N \times N$ describing a cyclic convolution. Since any circulant matrix is diagonalized by the Fourier matrix,

$$\tilde{\mathbf{H}} = \mathbf{F}_N \mathbf{H}_c \mathbf{F}_N^{-1} \tag{2.25}$$

is diagonal. Applying the DFT on $\mathbf{y}$ in (2.23), the receive symbol in frequency domain becomes

$$\tilde{\mathbf{y}} = \mathbf{F}_N \mathbf{y} = \mathbf{F}_N \mathbf{H}_c \mathbf{x} = \mathbf{F}_N \mathbf{H}_c \underbrace{\mathbf{F}_N^{-1} \mathbf{F}_N}_{\mathbf{I}} \mathbf{x} = \tilde{\mathbf{H}} \mathbf{F}_N \mathbf{x} = \tilde{\mathbf{H}} \mathbf{F}_N \mathbf{G} \mathbf{d}, \tag{2.26}$$

the last step follows from (2.11). Taking into account the overall channel model in Fig. 2.1, the introduction of a CP provides the simple relation

$$\tilde{\mathbf{y}} = \tilde{\mathbf{H}} \mathbf{F}_N \mathbf{G} \mathbf{d} + \mathbf{v}, \tag{2.27}$$

with an additive white Gaussian noise vector $\mathbf{v} = \mathbf{F}_N \mathbf{n}$ of length N. Multiplying with $\tilde{\mathbf{H}}^{-1}$ equalizes the dispersions caused by the multipath channel, a trivial operation considering the diagonality of $\tilde{\mathbf{H}}^{-1}$. This operation is independent of the design of $\mathbf{G}$, which only plays a role with respect to data estimation[6]. Using CPs in an OFDM system known as CP-OFDM, a data estimate from (2.27) is immediately available after equalization. For single-carrier systems employing CPs, commonly referred to in the literature as single-carrier systems with frequency domain equalization (SC/FDE), a transform back to the time domain after equalization is necessary

$$\hat{\mathbf{d}}_{\text{CP-OFDM}} = \tilde{\mathbf{H}}^{-1} \tilde{\mathbf{y}} = \tilde{\mathbf{H}}^{-1} \tilde{\mathbf{H}} \mathbf{F}_N \mathbf{G} \mathbf{d} + \tilde{\mathbf{H}}^{-1} \mathbf{F}_N \mathbf{n} \tag{2.28}$$

$$= \tilde{\mathbf{H}}^{-1} \tilde{\mathbf{H}} \mathbf{F}_N \mathbf{F}_N^{-1} \mathbf{d} + \tilde{\mathbf{H}}^{-1} \mathbf{F}_N \mathbf{n} \tag{2.29}$$

$$= \mathbf{d} + \tilde{\mathbf{H}}^{-1} \mathbf{F}_N \mathbf{n} \tag{2.30}$$

[6] Note that these derivations shall highlight the purpose of using CPs. Hence, the estimators are kept as simple as possible, thus also neglecting additional information such as noise characteristics, which can further enhance the performance in certain cases.

$$\hat{\mathbf{d}}_{\mathrm{SC/FDE}} = \mathbf{F}_N^{-1}\tilde{\mathbf{H}}^{-1}\tilde{\mathbf{y}} \tag{2.31}$$
$$= \mathbf{F}_N^{-1}\tilde{\mathbf{H}}^{-1}\tilde{\mathbf{H}}\mathbf{F}_N\mathbf{G}\mathbf{d} + \mathbf{F}_N^{-1}\tilde{\mathbf{H}}^{-1}\mathbf{F}_N\mathbf{n} \tag{2.32}$$
$$= \mathbf{F}_N^{-1}\tilde{\mathbf{H}}^{-1}\tilde{\mathbf{H}}\mathbf{F}_N\mathbf{d} + \mathbf{F}_N^{-1}\tilde{\mathbf{H}}^{-1}\mathbf{F}_N\mathbf{n} \tag{2.33}$$
$$= \mathbf{d} + \mathbf{F}_N^{-1}\tilde{\mathbf{H}}^{-1}\mathbf{F}_N\mathbf{n}. \tag{2.34}$$

Using a CP in the guard interval mainly pursues the goal of creating nice mathematical properties, which in turn allow for a simple and low complex equalization in the frequency domain. Due to this limited purpose, the CP is a rather wasteful consumer of transmit energy and transmit time reducing bandwidth efficiency. Unique Word OFDM aims at lifting these limitations by implementing a multipurpose guard interval instead, while still providing the same mathematical properties as a CP. The following section will introduce this novel concept.

2.1. Unique Word OFDM Symbol Generation

Let $\mathbf{x}_u \in \mathbb{C}^{N_u \times 1}$ be a predefined sequence which we call unique word (UW). This unique word shall form the tail of each OFDM time domain symbol vector of length N and occupy the guard interval of equal length $N_g = N_u$, as illustrated in Fig. 2.5. Hence, an OFDM time domain symbol $\mathbf{x}' \in \mathbb{C}^{N \times 1}$ consists of two parts and is of

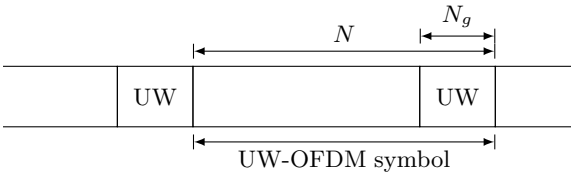


Figure 2.5.: Structure of an OFDM symbol using a unique word.

the form

$$\mathbf{x}' = \begin{bmatrix} \mathbf{x}_p \\ \mathbf{x}_u \end{bmatrix}, \tag{2.35}$$

whereas $\mathbf{x}_p \in \mathbb{C}^{(N-N_u)\times 1}$ carries the payload affected by the data symbols [35]. It is not immediately apparent at this point, that the proposed structure for UW-OFDM symbols will transform a linear into a cyclic convolution the same way a CP does. However, this becomes clear when considering succeeding UW-OFDM symbols and the repeating occurrence of UWs. Following the same arguments as in the previous section, and inserting $\mathbf{x}_{\mathrm{GI}}^{(l)} = \mathbf{x}_u$ and $\mathbf{x}^{(l)} = \begin{bmatrix} \mathbf{x}_p^T & \mathbf{x}_u^T \end{bmatrix}^T$ in (2.20),

the very same cyclic data structure appears and consequently the same properties hold.

The proposed structure for UW-OFDM symbols implies another important property which is not immediately apparent at first sight. Its potential though justifies a slightly off-topic discourse at this stage. Since in UW-OFDM the GI is part of the DFT interval, changing its length N_u will not influence the DFT length N and thus the framing structure. In other words, UW-OFDM enables an easy adaptation to different channel conditions by varying N_u while keeping relevant system structures unchanged. This is in clear contrast to other signaling schemes. As such, UW-OFDM can build the bridge between providing a communication system suitable to a wide range of scenarios while still ensuring high efficiency.

In order to obtain the desired UW-OFDM symbol, it is advantageous [35] to first generate an OFDM time domain symbol with a zero UW

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_p \\ \mathbf{0} \end{bmatrix}. \tag{2.36}$$

The final transmit symbol $\mathbf{x}'$ is then obtained by adding the desired UW in time domain in a second step, yielding

$$\mathbf{x}' = \mathbf{x} + \begin{bmatrix} \mathbf{0} \\ \mathbf{x}_u \end{bmatrix}. \tag{2.37}$$

The decision towards a two-step instead of a direct UW generation approach is due to energy reasons, a detailed reasoning delivers Sec. 3.2. Since the second step of adding the UW is trivial, the following considerations focus only on the first one.

As in conventional OFDM, the data symbols are drawn from a symbol alphabet $\mathcal{A}$, usually quadrature amplitude modulation (QAM), phase shift keying (PSK) or amplitude shift keying (ASK) constellations, and assembled in a data vector $\mathbf{d} \in \mathcal{A}^{N_d \times 1}$. This data vector shall be mapped on the OFDM symbol in frequency domain $\tilde{\mathbf{x}} = \mathbf{F}_N \mathbf{x}$ with $\tilde{\mathbf{x}} \in \mathbb{C}^{N \times 1}$, often together with unused zero subcarriers. Zeros are normally inserted at DC and the band edges to satisfy implementational restrictions of analogue hardware and spectral shaping requirements, respectively. So far, these symbol properties coincide with other OFDM schemes. UW-OFDM, however, additionally demands a zero-word in time domain as part of the vector

$$\mathbf{x} = \mathbf{F}_N^{-1} \tilde{\mathbf{x}} = \begin{bmatrix} \mathbf{x}_p \\ \mathbf{0} \end{bmatrix}. \tag{2.38}$$

Fixing N_u zeros in time domain, the system of equations $\mathbf{F}_N^{-1} \tilde{\mathbf{x}} = \left[\begin{smallmatrix} \mathbf{x}_p \\ \mathbf{0} \end{smallmatrix} \right]$ can only be fulfilled by reducing the number of data symbols in frequency domain by at least N_u, and instead introducing a certain kind of redundancy [35]. There might

arise some confusion at this point, as the idea of *redundancy* is presented without any preliminary motivation. The reader is thus asked for a little patience, an explanation is developed in the following.

Let a generator matrix $\mathbf{G} \in \mathbb{C}^{(N_d+N_r)\times N_d}$ with $N_r \geq N_u$ ($N_r = N_u$ is assumed at the moment) be the fundamental unit that generates UW-OFDM symbols, then the linear mapping

$$\mathbf{c} = \mathbf{G}\mathbf{d} \tag{2.39}$$

delivers a codeword $\mathbf{c} \in \mathbb{C}^{(N_d+N_r)\times 1}$. The term *codeword* has not been selected thoughtless in this context, but with special care to emphasize the parallels to the coding world. Similar as for a linear block code, the generator matrix maps N_d data symbols onto a codeword of length $N_d + N_r$, where N_r accounts for the introduced redundancy. In combination with a subcarrier selection matrix $\mathbf{B} \in \{0,1\}^{N\times(N_d+N_r)}$ that models the insertion of optional N_z zero subcarriers, a UW-OFDM symbol in frequency domain can then be written as

$$\tilde{\mathbf{x}} = \mathbf{B}\mathbf{c} = \mathbf{B}\mathbf{G}\mathbf{d}, \tag{2.40}$$

with the number of total subcarriers of $\tilde{\mathbf{x}}$ given as $N = N_d + N_r + N_z$. Matrix $\mathbf{B}$ is constructed by taking an identity matrix of size $N_d + N_r$ and adding zero rows at the corresponding zero subcarrier positions. Applying an IDFT on $\tilde{\mathbf{x}}$ yields a UW-OFDM symbol in time domain

$$\mathbf{x} = \mathbf{F}_N^{-1}\tilde{\mathbf{x}} = \mathbf{F}_N^{-1}\mathbf{B}\mathbf{G}\mathbf{d} = \begin{bmatrix}\mathbf{x}_p \\ \mathbf{0}\end{bmatrix}. \tag{2.41}$$

The generator matrix $\mathbf{G}$ has not been detailed yet, it is only specified so far that is should map the data symbols on $\mathbf{c}$ in some way, and additionally enable a zero-word at the tail of the time domain symbol $\mathbf{x}$. To achieve this, let us take the approach

$$\mathbf{G} = \mathbf{A}\begin{bmatrix}\mathbf{I} \\ \mathbf{T}\end{bmatrix}, \tag{2.42}$$

whereas $\mathbf{A} \in \mathbb{R}^{(N_d+N_r)\times(N_d+N_r)}$, $\mathbf{T} \in \mathbb{C}^{N_r\times N_d}$ and $\mathbf{I}$ is an identity matrix of size N_d. The reason for this specific decomposition will be apparent at the end of this section. For now it suffices to assume $\mathbf{A}$ having full rank such that it can map all information on $\mathbf{c}$ generated by $\mathbf{I}$ and $\mathbf{T}$. Inserting (2.42) into (2.41) delivers

$$\mathbf{x} = \mathbf{F}_N^{-1}\mathbf{B}\mathbf{A}\begin{bmatrix}\mathbf{I} \\ \mathbf{T}\end{bmatrix}\mathbf{d} = \mathbf{F}_N^{-1}\mathbf{B}\mathbf{A}\begin{bmatrix}\mathbf{d} \\ \mathbf{T}\mathbf{d}\end{bmatrix} \stackrel{!}{=} \begin{bmatrix}\mathbf{x}_p \\ \mathbf{0}\end{bmatrix}. \tag{2.43}$$

Neglecting $\mathbf{A}$ for the moment, (2.43) immediately shows that an UW-OFDM signal incorporates two components, namely the actual data symbols $\mathbf{d}$ as well as additional and therefore *redundant* information about the data symbols in form of the linear mapping $\mathbf{T}\mathbf{d}$. Of course, $\mathbf{A}$ may compound $\mathbf{d}$ and $\mathbf{T}\mathbf{d}$ in a way that lifts the distinction among them, but the basic concept of adding redundancy

remains unaffected. Further, the redundant part $\mathbf{Td}$ will ensure the zero-word constraint in the time domain, if $\mathbf{T}$ is chosen accordingly. With the introduction of

$$\mathbf{M} = \mathbf{F}_N^{-1}\mathbf{BA} = \begin{bmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} \\ \mathbf{M}_{21} & \mathbf{M}_{22} \end{bmatrix}, \tag{2.44}$$

whereas $\mathbf{M}_{11} \in \mathbb{C}^{(N-N_u)\times N_d}$, $\mathbf{M}_{12} \in \mathbb{C}^{(N-N_u)\times N_r}$, $\mathbf{M}_{21} \in \mathbb{C}^{N_u \times N_d}$ and $\mathbf{M}_{22} \in \mathbb{C}^{N_u \times N_r}$, the OFDM time domain symbol can be rewritten as

$$\begin{bmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} \\ \mathbf{M}_{21} & \mathbf{M}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{d} \\ \mathbf{Td} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_p \\ \mathbf{0} \end{bmatrix}. \tag{2.45}$$

It follows that

$$\mathbf{0} = \mathbf{M}_{21}\mathbf{d} + \mathbf{M}_{22}\mathbf{Td} \tag{2.46}$$

and finally

$$\mathbf{T} = -(\mathbf{M}_{22})^{-1}\mathbf{M}_{21}. \tag{2.47}$$

Note that the zero UW can only be generated, if it holds that $N_r \geq N_u$. Sec. 3.6 investigates the special case of $N_r > N_u$, the rest of this work assumes $N_r = N_u$. Hence, with a properly constructed generator matrix $\mathbf{G}$, a UW-OFDM time domain symbol with a zero word at its tail indeed follows as

$$\mathbf{x} = \mathbf{F}_N^{-1}\tilde{\mathbf{x}} = \mathbf{F}_N^{-1}\mathbf{BGd}. \tag{2.48}$$

This equation now denotes an explicit rule for the generation of UW-OFDM symbols. In fact, any time domain symbol $\mathbf{x}$ generated according to (2.48) will always feature a zero UW at its tail, given that the matrices $\mathbf{T}$ and $\mathbf{G}$ are constructed according to (2.47) and (2.42). This holds true independent of the specific realization of $\mathbf{A}$. From another perspective, this means that the choice of $\mathbf{A}$ does not influence $\mathbf{G}$ in qualifying as UW-OFDM generator matrix. However, the right choice of $\mathbf{A}$ is a major contribution of this work, as it essentially determines the applicability of UW-OFDM for communication systems. In order to confirm this, let us carry out a short experiment, at which end the necessity of $\mathbf{A} \neq \mathbf{I}$ is justified by evaluating the mean symbol energy $E_{x'} = \mathrm{E}\left\{\mathbf{x'}^H\mathbf{x'}\right\}$ of a UW-OFDM time domain symbol $\mathbf{x'}$ defined in (2.37).

Matrix $\mathbf{A}$ can be removed from the generator matrix $\mathbf{G}$ by simply choosing $\mathbf{A} = \mathbf{I}$. With the corresponding $\mathbf{G}$, the OFDM symbol in frequency domain translates to

$$\tilde{\mathbf{x}} = \mathbf{B}\begin{bmatrix} \mathbf{d} \\ \mathbf{Td} \end{bmatrix} = \mathbf{B}\begin{bmatrix} \mathbf{d} \\ \mathbf{r} \end{bmatrix}. \tag{2.49}$$

Consequently, the first N_d non-zero subcarriers are loaded with the data symbols in $\mathbf{d}$ and the next N_r subcarriers with dedicated redundant subcarrier symbols grouped together in a vector $\mathbf{r} \in \mathbb{C}^{N_r \times 1}$ as a result of the linear mapping

$$\mathbf{r} = \mathbf{Td}. \tag{2.50}$$

In order to assess its practicability, the proposed subcarrier loading scheme is applied to different setups specified in Tab. 2.2 (see page 35), and the mean symbol energy $E_{x'}$ is evaluated considering $\mathbf{x} = \mathbf{F}_N^{-1}\tilde{\mathbf{x}}$ and (2.37). The setups in Tab. 2.2 are used throughout the whole work to develop different insights. Setup A is inspired by the IEEE 802.11a standard [36], the remaining ones will be detailed later, at the moment it suffices to know that they differ in the specific value for the number of N total, N_d data, N_r redundant and N_z zero subcarriers. Aim of considering several setups is simply to establish an overall valid conclusion. Studying (2.37) and (2.49) suggests a breakdown of the mean symbol energy $E_{x'}$ into three sources: the energy of the data denoted as E_d, the energy of the redundancy E_r and the energy E_{x_u} required for the UW, with the latter vanishing in case of a zero UW. This separation seems reasonable but stands here without proof, a detailed analysis of the symbol energy and its sources in Sec. 3.2 will confirm its correctness though. Tab. 2.1 opposes E_d to E_r, whereas the calculation of E_d assumes a data covariance matrix $\mathbf{C}_{dd} = \sigma_d^2\mathbf{I}$ with $\sigma_d^2 = 1$. A comparison of these two sources reveals that the mean energy required for the redundancy part explodes in case of $\mathbf{A} = \mathbf{I}$, making a practical system unfeasible. Although not explicitly shown, the same problem would arise, if the redundancy would be mapped onto any other N_r concatenated subcarriers within the available bandwidth. Note that in Tab. 2.1 a scaling factor N has been introduced to compensate for $\frac{1}{N}$ which follows from the frequency-time domain relation. Therefore, the scaled energy values can be linked easier to N_r and N_d[7]. Since the redundancy originates from the mapping $\mathbf{r} = \mathbf{T}\mathbf{d}$, the mean energy contribution[8] of $\mathbf{r}$ on the mean energy of the time domain vector $\mathbf{x}'$ simply follows from

$$E_r = \frac{1}{N}\mathrm{E}\left\{\mathbf{r}^H\mathbf{r}\right\} = \frac{1}{N}\mathrm{E}\left\{\mathrm{tr}\left(\mathbf{r}\mathbf{r}^H\right)\right\} \tag{2.51}$$

$$= \frac{1}{N}\mathrm{E}\left\{\mathrm{tr}\left(\mathbf{T}\mathbf{d}\mathbf{d}^H\mathbf{T}^H\right)\right\} = \frac{1}{N}\mathrm{tr}\left(\mathbf{T}\mathrm{E}\left\{\mathbf{d}\mathbf{d}^H\right\}\mathbf{T}^H\right) \tag{2.52}$$

$$= \frac{1}{N}\mathrm{tr}\left(\mathbf{T}\mathbf{C}_{dd}\mathbf{T}^H\right) \tag{2.53}$$

$$= \frac{1}{N}\sigma_d^2\mathrm{tr}\left(\mathbf{T}\mathbf{T}^H\right), \tag{2.54}$$

where $\frac{1}{N}$ again follows from the frequency-time domain relation. Matrix $\mathbf{T}$ obviously influences the properties of the redundant subcarrier symbols and therefore E_r. $\mathbf{T}$ in turn depends on the matrix $\mathbf{A}$, which thus provides degrees of freedom to actively regulate the redundant energy.

[7] At this stage, a scaling by N makes only sense for the mean data energy E_d rather than for E_r, as in the latter case, the resulting values are so high that multiplying or not multiplying with N does not make a noticeable difference. However, in order to ensure consistency within the whole document, E_r is also scaled accordingly.

[8] Following the Parseval theorem [37], the energy can either be calculated in time or in frequency domain. The factor $\frac{1}{N}$ accounts for the transform to the time domain as a consequence of the utilized definition of the N-point DFT in (2.12).

Table 2.1.: Mean data energy E_d and mean redundant energy E_r for setups according to Tab. 2.2 in case of $\mathbf{A} = \mathbf{I}$. The optimum sets $\mathcal{I}_r$ for the redundant subcarrier positions already defined in this table are not considered for this experiment.

	Setup A	Setup B	Setup C	Setup D	Setup E
$E_d N$	36	48	64	100	112
$E_r N$	$7.18 \cdot 10^{23}$	$8.98 \cdot 10^{23}$	$1.13 \cdot 10^{27}$	$7.22 \cdot 10^{28}$	$1.27 \cdot 10^{29}$

The conclusion based on this experiment is that every generator matrix $\mathbf{G}$ constructed according to (2.47) and (2.42) will deliver UW-OFDM time domain symbols with the desired zero UW at the tail. However, the real challenge in UW-OFDM signal design lies in the proper choice of $\mathbf{A}$ to transfer the basic idea of UW-OFDM into a feasible concept. Denoting J as the cost function to determine the feasibility, UW-OFDM signal design can thus be summarized as solving an optimization problem

$$\check{\mathbf{A}} = \underset{\mathbf{A}}{\operatorname{argmin}} \{J\} \quad \text{s.t. } \mathbf{F}_N^{-1}\mathbf{B}\mathbf{G}\mathbf{d} = \begin{bmatrix} \mathbf{x}_p \\ \mathbf{0} \end{bmatrix} \wedge \mathbf{G} = \mathbf{A} \begin{bmatrix} \mathbf{I} \\ \mathbf{T} \end{bmatrix} \tag{2.55}$$

for every possible data vector $\mathbf{d}$, or equivalently

$$\check{\mathbf{A}} = \underset{\mathbf{A}}{\operatorname{argmin}} \{J\} \quad \text{s.t. } \mathbf{F}_N^{-1}\mathbf{B}\mathbf{G} = \begin{bmatrix} \mathbf{\Xi} \\ \mathbf{0} \end{bmatrix} \wedge \mathbf{G} = \mathbf{P} \begin{bmatrix} \mathbf{I} \\ \mathbf{T} \end{bmatrix}, \tag{2.56}$$

whereas $\mathbf{\Xi}$ denotes any arbitrary set of $N - N_u$ rows. Hence, if $\mathbf{m}_k^H$ with $k \in \{N - N_u, \ldots, N-1\}$ denote the N_u lowermost rows of $\mathbf{F}_N^{-1}\mathbf{B}$, then every column vector of $\mathbf{G}$ has to be orthogonal to every vector $\mathbf{m}_k$. $\check{\mathbf{A}}$ denotes an instance out of the possible realizations of $\mathbf{A}$ that minimizes the cost function J. The cost function J to be minimized still needs to be defined, but the results in Tab. 2.1 establish the redundant energy already as cost function as a good initial approach[9].

The design of $\mathbf{A}$ in the context of (2.56) is one of the major contributions of this work. Based on it, two principal classes of UW-OFDM signaling concepts are developed in this work. Implementing $\mathbf{A}$ as a permutation matrix yields systematically encoded UW-OFDM presented in chapter 3, and allowing $\mathbf{A}$ to be any real-valued non-singular matrix results in non-systematically encoded UW-OFDM investigated in chapter 4.

[9] Note that an energy based cost function is not applicable for all classes of UW-OFDM signaling concepts. However, it is a good starting point for UW-OFDM systems with dedicated data and redundant subcarriers, e.g., a system based on the exemplary generator matrix in (2.49), and deviated from there the concept of systematically encoded UW-OFDM presented in the next chapter.

Finally, despite some hidden hints have been given, a dedicated explanation for choosing $\mathbf{G}$ of the form

$$\mathbf{G} = \mathbf{A} \begin{bmatrix} \mathbf{I} \\ \mathbf{T} \end{bmatrix} \tag{2.57}$$

is still missing. As already stated before, calculating $\mathbf{T}$ according to (2.47) and assembling $\mathbf{G}$ as in (2.57) will automatically fulfill the zero-word constraint, independently of the specific realization of $\mathbf{A}$. In the context of (2.56), this means that the *constrained* optimization problem is automatically translated to an *unconstrained* optimization problem. This translation eases the searching process for a solution significantly, thus justifying the selected approach.

2.2. System Model

Based on (2.36)-(2.48), the time domain transmit symbol is given by

$$\mathbf{x}' = \mathbf{F}_N^{-1} \left(\mathbf{BGd} + \tilde{\mathbf{x}}_u \right), \tag{2.58}$$

with the frequency domain version of the unique word $\tilde{\mathbf{x}}_u = \mathbf{F}_N \left[\begin{smallmatrix} \mathbf{0} \\ \mathbf{x}_u \end{smallmatrix} \right]$ of length N. A received UW-OFDM time domain symbol after the transmission over a dispersive (e.g., multipath) channel can be modeled as

$$\mathbf{y}_r = \mathbf{H}_c \mathbf{x}' + \mathbf{n} \tag{2.59}$$

$$= \mathbf{H}_c \mathbf{F}_N^{-1} (\mathbf{BGd} + \tilde{\mathbf{x}}_u) + \mathbf{n}, \tag{2.60}$$

where $\mathbf{H}_c \in \mathbb{C}^{N \times N}$ denotes a cyclic convolution matrix with the zero-padded vector $\mathbf{h}_c \in \mathbb{C}^{N \times 1}$ of channel impulse response coefficients in its first column, and an additive white Gaussian noise vector $\mathbf{n} \in \mathbb{C}^{N \times 1}$. This noise model is very popular in communications, as it accounts for two important aspects. The first aspect is that a typical communication channel experiences various noise sources featuring a white and therefore constant power spectral density, such as thermal noise, shot noise or black body radiation from the earth and other warm objects. The second aspect is based on the central limit theorem (CLT) [37] and accounts for the fact that an accumulation of many random processes approaches a Gaussian distribution. In this work, $\mathbf{n}$ is modelled as a circularly symmetric complex white Gaussian noise vector with statistics $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma_n^2 \mathbf{I})$, with the real and imaginary part of all elements n_k following $\mathrm{Re}\{n_k\} \sim \mathcal{N}\left(\mathbf{0}, \frac{\sigma_n^2}{2}\right)$ and $\mathrm{Im}\{n_k\} \sim \mathcal{N}\left(\mathbf{0}, \frac{\sigma_n^2}{2}\right)$, respectively. Actual values for σ_n^2 required when evaluating the performance of UW-OFDM systems will be determined as a function of E_b/N_0 ratios, a detailed calculation can be found in Sec. 2.4.

After applying a DFT to obtain $\tilde{\mathbf{y}}_r = \mathbf{F}_N \mathbf{y}_r$, the zero subcarriers are excluded from further operation, leading to the downsized vector $\tilde{\mathbf{y}}_d \in \mathbb{C}^{(N_d+N_r)\times 1}$

$$\tilde{\mathbf{y}}_d = \mathbf{B}^T \tilde{\mathbf{y}}_r \tag{2.61}$$

$$= \mathbf{B}^T \mathbf{F}_N \mathbf{H}_c \mathbf{F}_N^{-1} (\mathbf{B}\mathbf{G}\mathbf{d} + \tilde{\mathbf{x}}_u) + \mathbf{B}^T \mathbf{F}_N \mathbf{n}. \tag{2.62}$$

Since any circulant matrix – as such also a cyclic convolution matrix – is diagonalized by the DFT, the transmission channel can be written as a diagonal matrix $\tilde{\mathbf{H}}_c = \mathbf{F}_N \mathbf{H}_c \mathbf{F}_N^{-1}$ ($\tilde{\mathbf{H}}_c \in \mathbb{C}^{N\times N}$) containing the sampled channel frequency response on its main diagonal. Excluding the entries corresponding to the zero subcarriers to obtain

$$\tilde{\mathbf{H}} = \mathbf{B}^T \mathbf{F}_N \mathbf{H}_c \mathbf{F}_N^{-1} \mathbf{B} \tag{2.63}$$

$$\tilde{\mathbf{H}} \in \mathbb{C}^{(N_d+N_r)\times(N_d+N_r)}, \tag{2.64}$$

the received symbol in the frequency domain is given in the form of the *affine* model

$$\tilde{\mathbf{y}}_d = \tilde{\mathbf{H}}\mathbf{G}\mathbf{d} + \tilde{\mathbf{H}}\mathbf{B}^T \tilde{\mathbf{x}}_u + \mathbf{B}^T \mathbf{F}_N \mathbf{n}. \tag{2.65}$$

Subtracting the known portion $\tilde{\mathbf{H}}\mathbf{B}^T\tilde{\mathbf{x}}_u$ originating from the UW (assuming that the channel matrix $\tilde{\mathbf{H}}$ or at least an estimate of the same is available) finally yields the linear model

$$\tilde{\mathbf{y}} = \tilde{\mathbf{y}}_d - \tilde{\mathbf{H}}\mathbf{B}^T \tilde{\mathbf{x}}_u \tag{2.66}$$

$$= \tilde{\mathbf{H}}\mathbf{G}\mathbf{d} + \mathbf{B}^T \mathbf{F}_N \mathbf{n} \tag{2.67}$$

$$= \tilde{\mathbf{H}}\mathbf{G}\mathbf{d} + \mathbf{v}, \tag{2.68}$$

with the noise vector $\mathbf{v} \sim \mathcal{CN}(\mathbf{0}, N\sigma_n^2\mathbf{I})$, and the variance $N\sigma_n^2$ as a result of the definition of the DFT in (2.12).

2.3. Linear Estimators

Based on the model in (2.68), several linear estimators of the form

$$\hat{\mathbf{d}} = \mathbf{E}\tilde{\mathbf{y}}, \tag{2.69}$$

are introduced in the following, whereas $\mathbf{E} \in \mathbb{C}^{N_d\times(N_d+N_r)}$ describes an estimator matrix. Note that this work is restricted to linear estimators by intention, a comprehensive analysis of non-linear receiver concepts for UW-OFDM is provided in [38, 39, 40]. Furthermore, this work focuses on the principle performance capabilities of the estimators rather than on specific details like implementational issues. For these issues, e.g., how to reduce the computational complexity, the reader is referred to [41] and [40].

2.3.1. Classical Data Estimators - Zero Forcing Solutions

A primary property of classical estimation theory is that a data vector is assumed to be deterministic but unknown. For an unbiased data estimator it then holds that

$$\mathrm{E}\left\{\hat{\mathbf{d}}\right\} = \mathrm{E}\left\{\mathbf{E}\tilde{\mathbf{y}}\right\} = \mathbf{E}\tilde{\mathbf{H}}\mathbf{G}\mathbf{d} = \mathbf{d}. \tag{2.70}$$

Consequently, the unbiased constraint takes on the form

$$\mathbf{E}\tilde{\mathbf{H}}\mathbf{G} = \mathbf{I}, \tag{2.71}$$

which is equivalent to the zero forcing (ZF) criterion for linear estimators. The solution to (2.71) is ambiguous. To show this, consider a singular value decomposition of $\tilde{\mathbf{H}}\mathbf{G} \in \mathbb{C}^{(N_d+N_r)\times N_d}$ as

$$\tilde{\mathbf{H}}\mathbf{G} = \mathbf{U}\begin{bmatrix}\mathbf{\Sigma}\\ \mathbf{0}\end{bmatrix}\mathbf{V}^H, \tag{2.72}$$

with unitary matrices $\mathbf{U} \in \mathbb{C}^{(N_d+N_r)\times(N_d+N_r)}$ and $\mathbf{V} \in \mathbb{C}^{N_d\times N_d}$, and with the diagonal matrix $\mathbf{\Sigma} \in \mathbb{R}^{N_d\times N_d}$ having on its main diagonal the singular values of $\tilde{\mathbf{H}}\mathbf{G}$. With (2.72) the unbiased constraint (or ZF criterion) (2.71) becomes

$$\mathbf{E}\mathbf{U}\begin{bmatrix}\mathbf{\Sigma}\\ \mathbf{0}\end{bmatrix}\mathbf{V}^H = \mathbf{I}. \tag{2.73}$$

It is easy to see that (2.73) and therefore also (2.71) is fulfilled by any estimator of the form

$$\mathbf{E} = \mathbf{V}\begin{bmatrix}\mathbf{\Sigma}^{-1} & \mathbf{\Psi}\end{bmatrix}\mathbf{U}^H \tag{2.74}$$

with arbitrary $\mathbf{\Psi} \in \mathbb{C}^{N_d\times N_r}$. Note that the fact of an ambiguous ZF solution distinguishes UW-OFDM from competing block oriented single input single output (SISO) approaches like e.g., CP-OFDM and CP-SC/FDE. For CP-OFDM the channel inversion (CI) receiver $\mathbf{E} = \tilde{\mathbf{H}}^{-1}$ represents the unambiguous ZF solution which also corresponds to the optimum data estimator, cf. [1]. For CP-SC/FDE the ZF solution is also unambiguous as soon as the receiver filter (e.g., a matched filter) preceding the estimator is specified. It is given by the inverse of the diagonal symbol spaced channel matrix which contains the influence of the transmit pulse shaping filter, the dispersive (e.g., multipath) channel and the receiver filter, cf. [28].

Out of the infinitely many solutions to the unbiased constraint, the best linear unbiased estimator (BLUE) and the CI estimator are investigated more in detail.

By applying the Gauss-Markov theorem [42] to (2.68) and with the noise covariance matrix $\mathbf{C}_{vv} = \mathrm{E}\left\{\mathbf{v}\mathbf{v}^H\right\} = N\sigma_n^2\mathbf{I}$, the BLUE and consequently the optimum ZF estimator follows to

$$\mathbf{E}_{\mathrm{BLUE}} = (\mathbf{G}^H\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}\mathbf{G})^{-1}\mathbf{G}^H\tilde{\mathbf{H}}^H. \tag{2.75}$$

$\mathbf{E}_{\mathrm{BLUE}}$ as given in (2.75) represents the pseudoinverse of $\tilde{\mathbf{H}}\mathbf{G}$. Since the noise in (2.68) is assumed to be Gaussian, (2.75) is also the minimum variance unbiased (MVU) estimator. The covariance matrix of $\hat{\mathbf{d}} = \mathbf{E}_{\mathrm{BLUE}}\tilde{\mathbf{y}}$, or equivalently the covariance matrix of the error $\mathbf{e} = \mathbf{d} - \hat{\mathbf{d}}$ is given by

$$\mathbf{C}_{ee} = \mathrm{E}\left\{\mathbf{e}\mathbf{e}^H\right\} = N\sigma_n^2(\mathbf{G}^H\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}\mathbf{G})^{-1}. \tag{2.76}$$

With the singular value decomposition as in (2.72) and after some rearrangements using standard matrix algebra, (2.75) can immediately be rewritten as

$$\mathbf{E}_{\mathrm{BLUE}} = \mathbf{V}\begin{bmatrix}\boldsymbol{\Sigma}^{-1} & \mathbf{0}\end{bmatrix}\mathbf{U}^H. \tag{2.77}$$

Based on (2.74) $\mathbf{E}_{\mathrm{BLUE}}$ corresponds to the solution for the particular case $\boldsymbol{\Psi} = \mathbf{0}$. $\mathbf{E}_{\mathrm{BLUE}}$ is in general a full matrix, which is in contrast to CP-OFDM and CP-SC/FDE, where the BLUE is given by a diagonal matrix.

The CI receiver (which only makes sense for systematically encoded UW-OFDM) is given by

$$\mathbf{E}_{\mathrm{CI}} = \begin{bmatrix}\mathbf{I} & \mathbf{0}\end{bmatrix}\mathbf{P}^T\tilde{\mathbf{H}}^{-1}. \tag{2.78}$$

This estimator inverts the channel $\tilde{\mathbf{H}}$ first, and then extracts the data symbols. A permutation matrix $\mathbf{P} \in \{0,1\}^{(N_d+N_r)\times(N_d+N_r)}$ is part of the generator matrix $\mathbf{G}$ by choosing $\mathbf{A} = \mathbf{P}$ in (2.57). Note that $\mathbf{P}$ is specific to systematically encoded UW-OFDM and will be explained in detail in chapter 3. Clearly, $\mathbf{E}_{\mathrm{CI}}$ fulfills (2.71). The CI receiver represents a low complex solution since $\tilde{\mathbf{H}}$ has a diagonal structure, but it does not take advantage of the correlations introduced by $\mathbf{G}$ at the transmitter side. The covariance matrix of $\hat{\mathbf{d}} = \mathbf{E}_{\mathrm{CI}}\tilde{\mathbf{y}}$, or equivalently the covariance matrix of the error $\mathbf{e} = \mathbf{d} - \hat{\mathbf{d}}$ can easily shown to be

$$\mathbf{C}_{ee} = N\sigma_n^2(\tilde{\mathbf{H}}_1^H\tilde{\mathbf{H}}_1)^{-1}, \tag{2.79}$$

where the diagonal matrix $\tilde{\mathbf{H}}_1 \in \mathbb{C}^{N_d\times N_d}$ contains on its main diagonal the sampled channel frequency response coefficients corresponding to the data subcarriers, and it originates from

$$\mathbf{P}\tilde{\mathbf{H}}\mathbf{P}^T = \begin{bmatrix}\tilde{\mathbf{H}}_1 & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{H}}_2\end{bmatrix}. \tag{2.80}$$

2.3.2. Linear Bayesian Data Estimators – LMMSE Solutions

Contrary to classical estimation theory, where the data vector is assumed to be deterministic and unknown, Bayesian theory thinks of the data vector as a realization of a random vector. By applying the Bayesian Gauss-Markov theorem [42] to (2.68), where assuming $\mathbf{d}$ to be the realization of a random vector, and by using $\mathbf{C}_{dd} = \sigma_d^2\mathbf{I}$ and $\mathbf{C}_{vv} = N\sigma_n^2\mathbf{I}$, the linear minimum mean square error (LMMSE) estimator follows to

$$\mathbf{E}_{\mathrm{LMMSE}} = \mathbf{W}\tilde{\mathbf{H}}^{-1}. \tag{2.81}$$

Here, $\mathbf{W}$ represents a Wiener smoothing matrix given by

$$\mathbf{W} = \mathbf{G}^H \left(\mathbf{G}\mathbf{G}^H + \frac{N\sigma_n^2}{\sigma_d^2} (\tilde{\mathbf{H}}^H \tilde{\mathbf{H}})^{-1} \right)^{-1}. \tag{2.82}$$

Equation (2.81) allows the following interpretation of the mode of operation: The LMMSE estimator acts as a composition of a simple channel inversion stage (multiplication with $\tilde{\mathbf{H}}^{-1}$ as in (2.78)) and a Wiener smoothing operation (multiplication with $\mathbf{W}$). The Wiener smoothing operation exploits the correlations between subcarrier symbols which have been introduced by $\mathbf{G}$ at the transmitter and acts as a noise reduction operation on the subcarriers. By applying the matrix inversion lemma, it can be shown that the estimator can equivalently be determined by

$$\mathbf{E}_{\text{LMMSE}} = (\mathbf{G}^H \tilde{\mathbf{H}}^H \tilde{\mathbf{H}} \mathbf{G} + \frac{N\sigma_n^2}{\sigma_d^2} \mathbf{I})^{-1} \mathbf{G}^H \tilde{\mathbf{H}}^H. \tag{2.83}$$

Eq. (2.83) shows strong similarities to the BLUE in (2.75). For $\sigma_n^2 = 0$ the expressions for LMMSE estimator and BLUE coincide. The error $\mathbf{e} = \mathbf{d} - \hat{\mathbf{d}}$ has zero mean, and its covariance matrix is given by

$$\mathbf{C}_{ee} = N\sigma_n^2 (\mathbf{G}^H \tilde{\mathbf{H}}^H \tilde{\mathbf{H}} \mathbf{G} + \frac{N\sigma_n^2}{\sigma_d^2} \mathbf{I})^{-1}. \tag{2.84}$$

It is important to note at this point that in contrast to the BLUE, an LMMSE estimator is not unbiased anymore in the classical sense (2.70). Unbiasedness is only given in the Bayesian sense when averaging over the data symbols as well. If not compensated or adequately taken into account, this bias introduction causes a significant bit error degradation when using higher order transmit constellations. In this work, the LMMSE bias is taken care of as part of the reliability information fed to the channel decoder, which is conducted in form of log-likelihood ratio (LLR) values, see appendix A.

2.4. Simulation Setup

The performance of UW-OFDM will be evaluated by analytical considerations as well as numerical simulations, e.g., in terms of the BER behavior. This part provides all information necessary to reproduce the evaluation framework used in this work.

2.4.1. Schematic of Simulation Chain

The UW-OFDM concept has thus been embedded into a full communication chain visualized as block diagram in Fig. 2.6. The different components are discussed in the following.

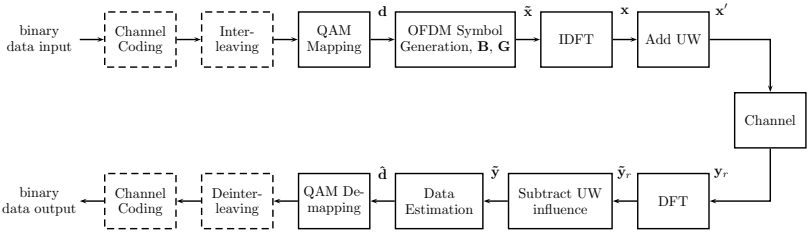


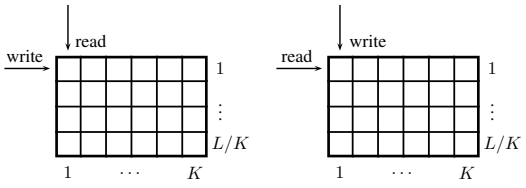
Figure 2.6.: Block diagram of the simulated Unique Word OFDM systems.

Channel Coding and Interleaving The binary information symbols are assumed to be uniformly as well as independent and identically distributed (iid), i.e., with entropy 1. Since this work considers both, coded and uncoded transmission, the information symbols are either transmitted directly, or channel coded and interleaved before further processing. The optionality of these operations is indicated by dashed boxes.

For coded transmission, a convolutional encoder with standard rate $r = 1/2$, constraint length 7 and the generator polynomials $(133, 171)_8$ in octal representation [43] are used. Additional results are provided for a code rate of $r = 3/4$, which is obtained by puncturing the $r = 1/2$ code based on the pattern $\left(\begin{smallmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{smallmatrix}\right)$. Both codes have been used during the standardization process of the IEEE 802.11a standard [36].

In order to avoid bundles of errors, the encoded information bits are interleaved by a block interleaver of block size L. The interleaver consists of a memory block, modelled as a matrix in Fig. 2.7a for reasons of visualization, featuring K columns and L/K rows, where K denotes the so-called *interleaving factor*. The binary input symbols are written horizontally in subsequent rows, and they are read vertically from subsequent columns. Hence, an interleaver spreads adjacent coded bits by as many positions as determined by the interleaving factor. This factor shall ensure that firstly neighboring bits are spread further apart than the constraint length of the code, and that secondly these bits are placed on nonadjacent subcarriers to avoid effects resulting from correlated channel frequency response coefficients. The second requirement can be challenging for burst-wise transmission in time-invariant channels (which is the case in this work), as all OFDM symbols experience the same channel conditions. A large interleaving factor may place neighboring bits on subcarriers with correlated channel conditions in one of the *next* OFDM symbols. To exclude this pitfall in the first place, the block size L has been set to the number of bits carried within a single OFDM symbol, forcing the interleaving distance to be limited to one OFDM symbol due to $K < L$ (known as in-place interleaving, with the advantage of not increasing the latency of a data stream). The interleaving factor K of the IEEE 802.11a standard (see setup A

in Tab. 2.3) has served as initialization value and has been slightly adapted for the different setups of UW-OFDM and CP-OFDM detailed in Tab. 2.2 and 2.3. Moreover, experiments showed that for the receivers used in this work, a limitation of the interleaving process to one symbol did not significantly degrade the BER performance. For iterative estimators as presented in [44], however, a block length and interleaving factor beyond one OFDM symbol would pay off and translate to an improved BER behavior.



(a) Interleaving (b) Deinterleaving

Figure 2.7.: Reordering of binary symbols after channel encoder (interleaving) and before channel decoder (deinterleaving).

Mapping The sequence of binary symbols is then mapped to transmit symbols. In this work, QPSK is the preferred modulation alphabet, having constellation points at $\{1+j, 1-j, -1+j, 1-j\}$. Selected results are also presented for the case of 16-QAM as modulation alphabet, with constellation points formed from values out of $\{-3, -1, 1, 3\}$ in both, real and imaginary part. The resulting QPSK and 16-QAM symbols are scaled with $\rho = 1/\sqrt{2}$ and $\rho = 1/\sqrt{10}$ to obtain unit variance, respectively. Gray mapping [45] is applied in all cases.

OFDM symbol generation A UW-OFDM symbol is generated in frequency domain by carrying out $\tilde{\mathbf{x}} = \mathbf{BGd}$ according to (2.40), where $\mathbf{B}$ models the insertion of optional zero subcarriers. The investigation of the generation process is the main contribution of this work, and the design of $\mathbf{G}$ and its implications will be detailed in the following chapters.

IDFT and unique word addition In a next step, the UW-OFDM symbol is transformed to the time domain by applying an inverse DFT (IDFT) as described in (2.38), followed by an optional addition of a desired non-zero UW (2.37) that completes the generation of an UW-OFDM time domain symbol. Since the non-zero UW is orthogonal to the UW-OFDM time domain symbol with zero tail, the signals do not influence each other and can thus be considered independently. This work focuses on the principles of the UW-OFDM generation process rather than

on the specific properties of UWs, hence no additional UW is added by default, yielding UW-OFDM time domain symbols with a zero word at the tail. Only in cases where a non-zero UW is needed or orthogonality is lost between the additive UW term and the UW-OFDM symbol with zero word (and therefore more in-depth investigations are required), a UW different from the zero word will be used.

The orthogonality between both terms enables a design of the non-zero UW just for the purpose of providing optimal synchronization and estimation properties. As stressed later in this work, however, the UW will in general overlay the whole signal bandwidth. Hence, caution is required to maintain spectral shaping properties.

Several UW-OFDM time domain symbols are assembled to form together a transmission burst as illustrated in Fig. 2.8. One burst comprises 8 000 information bits, the number of OFDM symbols per burst varies based on coding rate and modulation alphabet. Note that in front of the first UW-OFDM symbol, an additional UW is inserted to ensure cyclicity. Furthermore, a preamble as defined in the IEEE 802.11a standard [36] is appended at the beginning of the burst for channel estimation purposes investigated in Sec. 4.3.4.

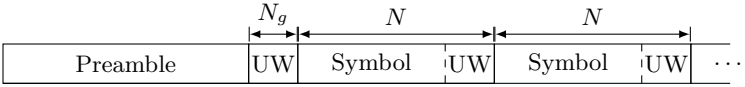


Figure 2.8.: Transmission burst of investigated UW-OFDM systems.

Channel The channel consists of a multipath environment and additive white Gaussian noise. The multipath propagation environment is modeled as a linear convolution of the transmit signal with a discrete channel impulse response (CIR) represented by a vector $\mathbf{h}$. A realization of the CIR is generated according to the channel model in [46], which has also been used for the IEEE 802.11a standardization process. The channel model is presented in detail in Sec. 2.4.2.

Noise is modeled at the receiver input as circularly symmetric complex additive white Gaussian noise (which implies zero mean) with variance σ_n^2, which is set according to a desired E_b/N_0 value measured at the receiver input. In case of an AWGN channel only, the CIR reduces to $\mathbf{h} = [1\,0\ldots0]^T$.

The relation E_b/N_0 serves as a signal-to-noise ratio (SNR) measure, with E_b denoting the average energy per bit of information and $N_0/2$ the double-sided noise power spectral density of a bandpass noise signal. Following the derivations in [33],

the variance σ_n^2 of the (complex) noise term is given by

$$\sigma_n^2 = 2\frac{mP_s}{2(E_b/N_0)rb}, \tag{2.85}$$

whereas m denotes the number of samples per (OFDM) symbol, with $m = N$ and $m = N + N_g$ in case of UW-OFDM and CP-OFDM, respectively. P_s represents the average power per sample, which generally varies with the concept and the specific setup. One way to determine P_s is based on the average transmit symbol energy. In case of UW-OFDM and without pilot tones, the average power per sample is given as $P_s := P_{s,\mathrm{uw}} = \frac{E_{x'}}{m}$ with the mean transmit symbol energy $E_{x'} = \mathrm{E}\left\{\mathbf{x}'^H\mathbf{x}'\right\}$ of an UW-OFDM time domain symbol $\mathbf{x}'$ defined in (2.37). For pilot based UW-OFDM systems as presented in chapter 5, an additional energy term $E_p^{(r)}$ specified in (5.13) has to be taken into account. In case of CP-OFDM, the average power per sample is given as $P_s := P_{s,\mathrm{cp}} = \frac{E_{x,\mathrm{cp}}}{m}$ with $E_{x,\mathrm{cp}} = \mathrm{E}\left\{\mathbf{x}_{\mathrm{cp}}^H\mathbf{x}_{\mathrm{cp}}\right\}$ denoting the mean transmit symbol energy of a corresponding CP-OFDM time domain symbol $\mathbf{x}_{\mathrm{cp}}$. For details, the reader is referred to Sec. 4.3.3, which will provide expressions for $E_{x'}$ and $E_{x,\mathrm{cp}}$ specifically evaluated for the system configurations used in this work, cf. (4.35) and (4.42).

As usual for digital signal considerations, a normalization of the sampling time T_s to 1 applies in this context to enable an analysis independent of the specific sampling frequency. This normalization does not influence the results w.r.t. the SNR measure, as long as the sampled signal and the noise sequence are treated in the same way. Finally, r represents the coding rate of an optional channel code, and b corresponds to the number of bits per (OFDM) symbol.

DFT and unique word subtraction At the receiver, a DFT is applied in a first step to transform the OFDM symbol to frequency domain, followed by subtracting the constant offset of a non-zero UW in a second step as described in (2.66).

Data estimation The linear data estimators presented in Sec. 2.3 are used to recover the data symbols at the receiver side, the LMMSE estimator serves as default option. Furthermore, the corresponding error covariance matrix provides reliability information for the channel decoder.

Perfect channel knowledge in terms of the CIR and E_b/N_0 is assumed at the receiver in all cases except for Sec. 4.3.4. This part explicitly investigates the effects of channel estimation errors on the BER performance.

QAM demapping The process of demapping differs for uncoded and coded transmission. In the uncoded case, a data estimate is demapped to a sequence of binary

symbols associated with the transmit data symbol, that is closest to the data estimate in terms of the Euclidean distance. A demapping to binary symbols is also known as *hard decision*.

In the coded case, however, a data estimate is demapped to a sequence of reliability information instead, whereas each reliability information indicates for a received and therefore disturbed binary symbol the most likely corresponding binary transmit symbol. This kind of demapping is denoted as *soft decision*. The reliability for the receive binary symbols (after QAM demapping) can meaningfully be described by the ratio of the probabilities of each symbol to be 0 and 1. Starting from the data symbol vector estimate in (2.69), the kth data symbol estimate can be modelled as

$$\hat{d}_k = \mathbf{e}_k^H \tilde{\mathbf{y}}, \tag{2.86}$$

with $\mathbf{e}_k^H$ denoting the kth column of the linear estimator matrix $\mathbf{E}$. For this general estimator, the so-called LLRs of a bit mapped onto any symbol constellation can be written as

$$\mathrm{L}\left(b_{ik}\middle|\hat{d}_k\right) = \ln \frac{\Pr\left(b_{ik}=1\middle|\hat{d}_k\right)}{\Pr\left(b_{ik}=0\middle|\hat{d}_k\right)}, \tag{2.87}$$

where b_{ik} is the ith bit of the kth received symbol. Hence, $\mathrm{L}\left(b_{ik}\middle|\hat{d}_k\right)$ denotes the LLR of the ith bit within the kth estimated data symbol.

For $\mathrm{L}\left(b_{ik}\middle|\hat{d}_k\right) > 0$, it is more likely that $b_{ik} = 1$ has been transmitted, $\mathrm{L}\left(b_{ik}\middle|\hat{d}_k\right) < 0$ suggests $b_{ik} = 0$, and $\mathrm{L}\left(b_{ik}\middle|\hat{d}_k\right) = 0$ identifies both possibilities as equiprobable. The magnitude $\left|\mathrm{L}\left(b_{ik}\middle|\hat{d}_k\right)\right|$ represents the certainty, whereas absolute certainty would translate to an LLR of $\pm\infty$.

Eq. (2.87) denotes a general definition of the LLR. Appendix A provides a detailed derivation of the LLRs, when evaluated for specific estimators and symbol constellations used throughout this work.

Deinterleaving and decoding These operations are optional and only conducted, if the dual operations channel coding and interleaving have been applied at the transmitter. The deinterleaver recovers the original order of the binary symbol sequence as it was given in front of the interleaver. Its functionality is best described by a matrix of the same size as for the interleaver, but now the incoming data is written in the vertical and read in the horizontal direction, as depicted in Fig. 2.7a.

A Viterbi algorithm is applied for decoding of convolutional codes [47]. Since reliability information for the output of all considered data estimators is available,

and due to performance advantages, only soft-decision rather than hard-decision input is considered. The required soft information is provided by the demapper.

2.4.2. Multipath Channel Model

The CIRs used throughout this work to form a multipath propagation environment have been generated according to a discretized exponentially decaying power delay profile $P(\tau)$ [48] model from [46], which will be presented next. The power delay profile is specified at N_h sampling points for the time instants $\tau_k = kT_s; \quad k = 0, 1, \ldots, N_h - 1$ denoted as

$$P(\tau_k) = P_k, \tag{2.88}$$

where T_s denotes the sampling time of the corresponding UW-OFDM system. Assuming now N_h CIR coefficients stacked together in a CIR vector

$$\mathbf{h} = \left[h_0, h_1, \ldots, h_{N_h-1}\right]^T, \tag{2.89}$$

each coefficient h_k is modelled as a circularly symmetric complex white Gaussian random variable with $h_k \sim \mathcal{CN}(0, \sigma_k^2)$ and therefore $\mathrm{Re}\{h_k\}, \mathrm{Im}\{h_k\} \sim \mathcal{N}\left(0, \frac{\sigma_k^2}{2}\right)$. The variances of the coefficients follow an exponential decay with

$$\sigma_k^2 = P_k = \sigma_0^2 \mathrm{e}^{-k\frac{T_s}{\tau_{\mathrm{RMS}}}}, \tag{2.90}$$

whereas the variance σ_0^2 of the first coefficient – usually representing the line-of-sight path – is chosen as

$$\sigma_0^2 = 1 - \mathrm{e}^{-\frac{T_s}{\tau_{\mathrm{RMS}}}} \tag{2.91}$$

to ensure $\sum_{k=0}^{N_h-1} \sigma_k^2 = 1$. The parameter τ_{RMS} represents the root mean square delay spread of the channel. τ_{RMS} is the second central moment of $P(\tau)$ defined as [49]

$$\tau_{\mathrm{RMS}} = \sqrt{\frac{\sum_{k=0}^{N_h-1} P_k \tau_k^2}{\sum_{k=0}^{N_h-1} P_k} - \frac{\left(\sum_{k=0}^{N_h-1} P_k \tau_k\right)^2}{\left(\sum_{k=0}^{N_h-1} P_k\right)^2}}, \tag{2.92}$$

which can be interpreted as a general means to describe the time spreading of a signal due to the multipath channel. Furthermore, in an exponentially decaying power delay profile of the form $P(\tau) = P_0 \mathrm{e}^{-\beta\tau}$ with P_0 denoting the power at $\tau = 0$, τ_{RMS} is inversely proportional to the decaying constant β [50], see (2.90).

In order to simplify the model by excluding channel coefficients without any significant contribution, i.e., neglectable power level σ_k^2, the number of channel coefficients N_h has been very generously limited [46] to

$$N_h = 10\frac{\tau_{\mathrm{RMS}}}{T_s}. \tag{2.93}$$

In this work, CIR realizations are drawn from the channel model featuring $\tau_{\mathrm{RMS}} = 100\,\mathrm{ns}$ and $\tau_{\mathrm{RMS}} = 200\,\mathrm{ns}$, assuming a given sampling time of $T_s = 50\,\mathrm{ns}$ (as a consequence of $T_s = T_{\mathrm{OFDM}}/N$, see Tab. 2.2). The first case with $\tau_{\mathrm{RMS}} = 100\,\mathrm{ns}$ denotes the default option. Note that in this case (2.93) delivers $N_h = 20$, which exceeds the guard interval length $N_g = 16$ used throughout this work (Tab. 2.2 and 2.3). However, evaluating (2.90) immediately reveals that the tail of these CIRs will be practically zero, which can therefore also be chopped off. It can thus safely be assumed that a CIR vector $\mathbf{h}$ drawn from the channel model with $\tau_{\mathrm{RMS}} = 100\,\mathrm{ns}$ does not exceed the guard interval, that is a length of $N_g + 1$ taps.

2.4.3. Simulation

Tab. 2.2 summarizes different UW-OFM setups investigated in this work, whereas setup A determines the default option. Conventional CP based OFDM serves as main reference system. Tab. 2.3 summarizes the physical parameters of two exemplary CP-OFDM setups, with the first one again determining the default option. In any case, UW-OFDM and CP-OFDM are always compared within the same setup class, i.e., A versus A or C versus C.

Note that setup A in Tab. 2.3 corresponds to the physical layer model of the IEEE 802.11a standard, cf. [36]. This standard implements four pilot subcarriers dedicated to estimation and synchronization tasks. In case of UW-OFDM, a UW might take over these tasks from the pilot subcarriers. In order to enable a fair comparison, both systems should thus invest the same amount of energy for dedicated pilot symbols. Consequently, a UW should incorporate the same energy as the pilot subcarriers, the actual energy value is irrelevant though. Since in most cases a zero UW is applied, w.l.o.g. the pilot subcarrier symbols in the CP-OFDM systems are chosen to be zero unless stated otherwise.

The performance of UW-OFDM and CP-OFDM is evaluated by means of BER and MSE simulations, the latter is mainly used to quantify CFO effects in chapter 6. Simulation results are obtained by averaging over a fixed set of 10 000 CIR realizations, which are generated according to Sec. 2.4.2. The channel is assumed to be quasi-static, meaning that it stays constant during the transmission of one burst and then changes to an independent channel realization for the next burst. Each realization is used equally often in the simulations. A burst is assembled according to Fig. 2.8. Each burst contains 8 000 information bits, the number of OFDM

symbols and therefore the length of the burst depends on the applied coding rate and modulation alphabet.

Unless specified otherwise, performance differences between two systems in terms of the BER are measured at a BER $= 10^{-6}$, and the quotient between the required E_b/N_0 values is stated by the difference in dB.

For all operations arbitrarily accurate arithmetic is assumed, which means that quantization effects and numerical effects due to arithmetic implementation are neglected.

Default setup In order to provide a quick overview, the following part summarizes the most important default options of the presented simulation setup, which hold for this work unless stated explicitly otherwise. UW-OFDM and CP-OFDM are configured according to setup A in Tab. 2.2 and Tab. 2.3, respectively, with zero UW, and zero energy for the pilot symbols in case of CP-OFDM. QSPK normalized to unit variance serves as modulation alphabet. The multipath channel is modeled with CIR realizations featuring a channel delay spread of $\tau_{\mathrm{RMS}} = 100\,\mathrm{ns}$. Data estimation is performed by an LMMSE estimator.

Table 2.2.: Summary of the main PHY parameters of the investigated UW-OFDM setups.

		Setup A	Setup B	Setup C	Setup D	Setup E	Setup F
DFT size	N	64	64	80	128	128	64
data subcarriers	N_d	36	48	64	100	112	32
red. subcarriers	N_r	16	16	16	16	16	16
zero subcarriers	N_z	12	-	-	12	-	12
pilot subcarriers	N_p	-	-	-	-	-	4
unique word samples	N_u	16	16	16	16	16	16
zero subcarrier indices	$\mathcal{I}_z$	{0,27,...,37}	{}	{}	{0,59,...,69}	{}	{0,27,...,37}
red. subcarrier indices	$\mathcal{I}_r$	{2,6,10,14, 17,21,24,26, 38,40,43,47, 50,55,58,62}	{0,4,8,12, 16,20,24,28, 32,36,40,44, 48,52,56,60}	{0,5,10,15, 20,25,30,35, 40,45,50,55, 60,65,70,75}	{4,12,20,28,36, 44,51,58,70, 77,84,92,100, 108,116,124}	{0,8,16,24,32, 40,48,56,64, 72,80,88,96, 104,112,120}	{2,5,9,13, 17,20,24,26, 38,40,44,47, 51,54,58,62}
pilot subcarrier indices	$\mathcal{I}_p$	{}	{}	{}	{}	{}	{7,21,43,57}
interleaving factor		12	16	16	-	-	12
DFT length	T_{DFT}	3.2 μs	3.2 μs	4 μs	3.2 μs	3.2 μs	3.2 μs
guard interval length	T_{GI}	0.8 μs	0.8 μs	0.8 μs	0.8 μs	0.8μs	0.8 μs
OFDM symbol length	T_{OFDM}	3.2 μs	3.2 μs	4 μs	3.2 μs	3.2 μs	3.2 μs
subcarrier spacing	Δ_f	312.5 kHz	312.5 kHz	250 kHz	156.25 kHz	156.25 kHz	312.5 kHz

Table 2.3.: Summary of the main PHY parameters of the investigated CP-OFDM setups.

		Setup A	Setup C
DFT size	N	64	64
data subcarriers	N_d	48	64
zero subcarriers	N_z	12	0
pilot subcarriers	N_p	4	0
guard interval samples	N_g	16	16
zero subcarrier indices	$\mathcal{I}_z$	{0,27,28,...,37}	{}
pilot subcarrier indices	$\mathcal{I}_p$	{7,21,43,57}	{}
DFT length	T_{DFT}	3.2 μs	3.2 μs
guard interval length	T_{GI}	0.8 μs	0.8 μs
OFDM symbol length	T_{OFDM}	4 μs	4 μs
subcarrier spacing	Δ_f	312.5 kHz	312.5 kHz

3. Systematically Encoded Unique Word OFDM

The previous chapter has introduced a principle way to generate Unique Word based OFDM signals. The idea is based on a generator matrix of the form $\mathbf{G} = \mathbf{A}\begin{bmatrix}\mathbf{I} & \mathbf{T}^T\end{bmatrix}^T$, where $\mathbf{I}$ takes care of the data symbols, $\mathbf{A}$ is a freely selectable matrix detailed later in this chapter, and $\mathbf{T}$ is a given matrix responsible for generating OFDM signals with the desired time domain properties. More in detail, this matrix introduces a certain redundancy in the frequency domain to generate a UW with the desired properties in the time domain. Very important to note in this context is the dependence of $\mathbf{T}$ and thus the generated redundancy on $\mathbf{A}$. Experiments with $\mathbf{A} = \mathbf{I}$ performed in the previous chapter delivered OFDM symbols with N_d adjacent subcarriers loaded with data, followed by N_r adjacent subcarriers loaded with redundancy

$$\tilde{\mathbf{x}} = \mathbf{B}\begin{bmatrix}\mathbf{d}\\ \mathbf{r}\end{bmatrix}. \tag{3.1}$$

The results summarized in Tab. 2.1 strongly suggest a proper handling of the energy required for generating the redundancy, otherwise a feasible system cannot be achieved. This problem has already been discussed in [20] from a coding theory point of view, but remained unsolved.

The underlying idea of the concept presented in this chapter is to dissolve blocks of data and redundancy in frequency domain, and instead distribute the redundancy among the available bandwidth, while still enabling distinction among both in frequency domain. Such a distribution shall reduce the energy required for generating the redundancy, yielding a cost function

$$J_{\mathrm{E}} = \sigma_d^2 \mathrm{tr}\left(\mathbf{T}\mathbf{T}^H\right). \tag{3.2}$$

This cost function is derived from the redundant energy E_r in(2.54) by omitting $\frac{1}{N}$ and taking into account the usual assumption of independent and identically distributed (iid) data symbols with zero mean and covariance matrix $\mathbf{C}_{dd} = \sigma_d^2\mathbf{I}$. In order to model the distribution of redundancy over the bandwidth, a permutation matrix $\mathbf{P} \in \{0,1\}^{(N_d+N_r)\times(N_d+N_r)}$ is introduced and $\mathbf{A} = \mathbf{P}$ is chosen to yield a UW-OFDM generator matrix of the form

$$\mathbf{G} = \mathbf{P}\begin{bmatrix}\mathbf{I}\\ \mathbf{T}\end{bmatrix}. \tag{3.3}$$

The optimization problem follows as

$$\check{\mathbf{P}} = \underset{\mathbf{P}}{\operatorname{argmin}} \{J_{\mathrm{E}}\} \quad \text{s.t. } \mathbf{F}_N^{-1}\mathbf{B}\mathbf{G} = \begin{bmatrix} \boldsymbol{\Xi} \\ \mathbf{0} \end{bmatrix} \wedge \mathbf{G} = \mathbf{P} \begin{bmatrix} \mathbf{I} \\ \mathbf{T} \end{bmatrix}. \tag{3.4}$$

The resulting generator matrix $\mathbf{G}$ should then optimally map the data and the generated redundant symbols onto the subcarriers, whereas the actual assignment is the available degree of freedom. To say it in other words, the optimization problem transforms into finding a subset of N_r subcarriers out of $N - N_z$ possible ones that minimizes J_{E} while fulfilling the constraint in (3.4). Unfortunately, an optimum set by means of analytical analysis can only be found for a very few and simple setups, cf. [51], thus suggesting to apply numerical optimization methods. It turns out that for reasonable choices of N and N_r, an exhaustive search is unfeasible from the perspective of computational complexity. For instance, setup A already offers $\binom{(N-N_z)}{N_r} = \binom{52}{16} \approx 1.04 \cdot 10^{13}$ different combinations to be evaluated. Sec. 3.3 will investigate the problem of finding an optimal set in detail, however, the searching process is neglected for the moment and the optimal sets are simply assumed to be given. Tab. 2.2 summarizes the best redundant subcarrier sets for the exemplary setups used throughout this work. Note that in this table the optimal redundant subcarriers are specified by index sets $\mathcal{I}_r$ out of $\mathcal{I} = \{0 \ldots (N-1)\}$, the corresponding permutation matrix $\mathbf{P}$ can easily be derived from them. In order to ensure an unambiguous definition of $\mathbf{P}$, from now on it is always assumed that a permutation matrix $\mathbf{P}$ is derived from an ordered index set

$$\mathcal{I}_{r,o} = (\mathcal{I}_r, <). \tag{3.5}$$

Once a permutation matrix is given, the generator matrix $\mathbf{G}$ is easily determined using (3.3) and (2.47), and an UW-OFDM frequency domain symbol $\tilde{\mathbf{x}}$ follows to

$$\tilde{\mathbf{x}} = \mathbf{B}\mathbf{P} \begin{bmatrix} \mathbf{I} \\ \mathbf{T} \end{bmatrix} \mathbf{d} = \mathbf{B}\mathbf{G}\mathbf{d}. \tag{3.6}$$

In principle, the generator matrix $\mathbf{G}$ takes the data symbols vector, additionally produces redundant subcarrier symbols based on

$$\mathbf{r} = \mathbf{T}\mathbf{d}, \tag{3.7}$$

scrambles them and maps the resulting word on the subcarriers. There are now two popular ways of interpreting this generation process. Although both are based on the insertion of the redundancy, these different interpretations may ease the understanding of the beneficial properties of UW-OFDM later on.

- From a statistical point of view, $\mathbf{r}$ introduces correlations into the frequency domain symbol $\tilde{\mathbf{x}}$.

- From a coding theory point of view, the outcome

$$\mathbf{c} = \mathbf{G}\mathbf{d} \tag{3.8}$$

may be interpreted as a *systematic* code word $\mathbf{c} \in \mathbb{C}^{(N_d+N_r)\times 1}$, thus justifying the naming *systematically encoded* UW-OFDM.

In a systematic code word $\mathbf{c} = \mathbf{G}\mathbf{d}$, the data symbols are embedded into the encoded output, i.e., the data symbols are recognizable in the code word $\mathbf{c}$ as illustrated in Fig. 3.1. Since UW-OFDM time domain symbols contain with the UW a block of fixed samples, the set of all corresponding vectors in discrete frequency domain forms in fact a coset to a Reed Solomon (RS) code. Usually RS codes of length n are defined for a finite field $\mathbb{F}_Q$ using an element $w \in \mathbb{F}_Q$ of order n, $n \cdot l = Q - 1$, with $n, l, Q \in \mathbb{N}$ to define a discrete Fourier transform $\mathbb{F}_Q^n \rightarrow \mathbb{F}_Q^n$ in $\mathbb{F}_Q$. The set of codewords is specified by the fact, that the (inverse) DFT of all codewords contains a block of $d_{min} - 1$ successive zeros, where d_{min} is the minimum Hamming distance of the RS code. If this block of $d_{min} - 1$ successive symbols differs from zero, but is also fixed for all codewords, a coset code to an RS code is generated in the other domain w.r.t. this Fourier transform with the same minimum distance d_{min}, cf. [19]. This RS property of course allows for algebraic RS decoding methods which will be investigated more in detail in Sec. 3.5.1.

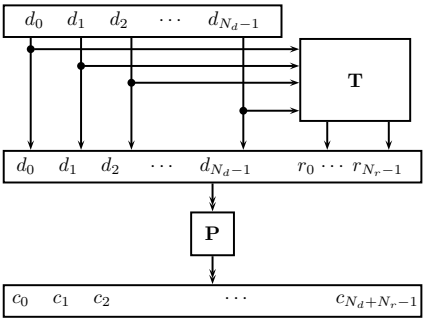


Figure 3.1.: Interpretation of an UW-OFDM symbol as a systematic code word.

Finally, the generation of a systematically encoded UW-OFDM time domain symbol can be summarized as

$$\mathbf{x} = \mathbf{F}_N^{-1}\mathbf{B}\mathbf{c} =, \mathbf{F}_N^{-1}\mathbf{B}\mathbf{P}\begin{bmatrix}\mathbf{d}\\ \mathbf{r}\end{bmatrix}. \tag{3.9}$$

To highlight the differences to conventional OFDM, Fig. 3.2 visualizes the symbol generation by detailing in- and output of the IDFT operation.

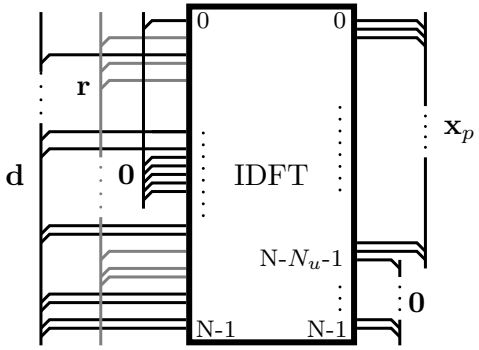


Figure 3.2.: Time- and frequency-domain view of a systematically encoded UW-OFDM symbol with zero UW.

3.1. Signal Properties

The following section compasses a study of the signal properties specific to this concept. This study is based on a top-down approach consisting of three levels, with the degree of detailing increasing incrementally. The first level denotes thus the highest degree of abstraction and focuses on the overall generation process. Insights are based on the whole OFDM symbol, i.e., data and redundancy together. Making one step closer to the symbol, the second level investigates only the redundancy part. Conclusions are drawn treating the redundancy as one entity. Finally, the third level presents results based on the properties of single redundant subcarrier symbols.

Properties drawn from an OFDM symbol. Fig. 3.3 reveals a number of interesting properties of a generator matrix $\mathbf{G}$ of the form (3.3) that solves the optimization problem in (3.4). It is already known that a data symbol d_i is mapped one-to-one onto a dedicated code word symbol c_i and is additionally spread over the redundant subcarrier symbols. Studying the rows of $\mathbf{G}$ shows that a single redundant symbol only depends on a few neighboring data symbols. This property leads to a relatively sparse matrix $\mathbf{G}$, see Fig. 3.3. Additionally, the generator matrix $\mathbf{G}$ also features the symmetry property

$$\mathbf{G} = [\mathbf{g}_0 \cdots \mathbf{g}_{N_d/2-1}\, \mathrm{flip}\{(\mathbf{g}_{N_d/2-1})^*\} \cdots \mathrm{flip}\{(\mathbf{g}_0)^*\}]. \tag{3.10}$$

Here, $\mathbf{g}_i$ with $i = 0, 1, ..., N_d/2 - 1$ represent the first $N_d/2$ columns of $\mathbf{G}$, and flip denotes an operator that flips the elements in vertical direction. Both properties – sparsity as well as symmetry – are of interest when implementing UW-OFDM resource efficiently on real-world hardware.

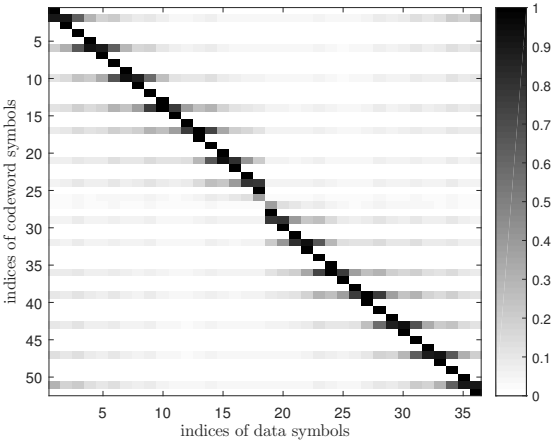


Figure 3.3.: $|\mathbf{G}|$, magnitude of entries of generator matrix $\mathbf{G}$.

Another interpretation of the generator matrix $\mathbf{G}$ is the combination of a linear dispersive preprocessor (or a channel-independent precoder, cf. [52]), and a channel coder. Analogously to precoding, a data symbol is spread over several subcarriers (due to the redundant subcarriers). In case of fading holes, a data symbol might still be detectable. Analogously to channel coding, additional redundancy is added. These two properties can explain the performance boost over conventional OFDM presented in subsequent sections.

An UW-OFDM frequency domain symbol immediately follows by loading the subcarriers according to $\tilde{\mathbf{x}} = \mathbf{Bc} = \mathbf{BGd}$. Fig. 3.4 illustrates the mean power values on these subcarriers evaluated for the two exemplary UW-OFDM setups A and B from Tab. 2.2 and their respective optimal sets. The mean power values of the non-zero subcarriers correspond to

$$\operatorname{diag}\left(\mathrm{E}\left\{\mathbf{c}\mathbf{c}^H\right\}\right) = \operatorname{diag}\left(\mathrm{E}\left\{\mathbf{G}\mathbf{d}\mathbf{d}^H\mathbf{G}^H\right\}\right) = \sigma_d^2 \operatorname{diag}\left(\mathbf{G}\mathbf{G}^H\right). \tag{3.11}$$

Properties drawn from the redundancy part in an OFDM symbol. Setup A with zero and setup B without zero subcarriers serve as representatives. It turns out that in the presence of zero subcarriers, the redundant subcarrier symbols experience different mean power values. Furthermore, the mean power of redundant subcarriers close to zero subcarriers decreases, i.e., next to DC and guard bands. In case no zero subcarriers are employed yielding an equidistant distribution, all

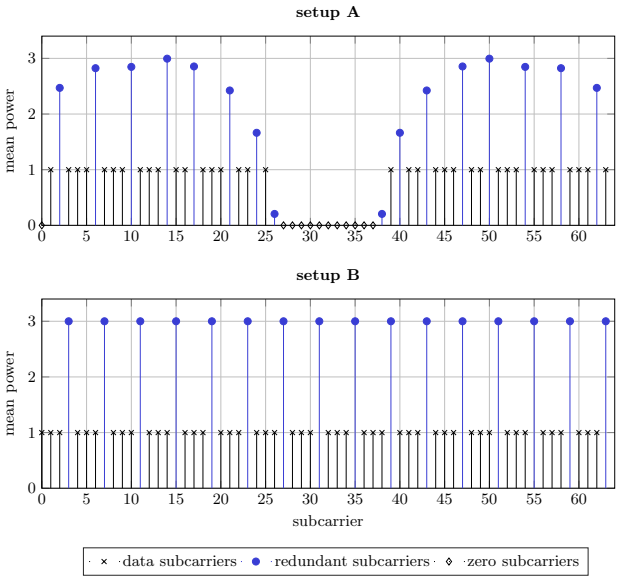


Figure 3.4.: Mean power of individual subcarrier symbols for setup A (upper plot) and setup B (lower plot).

redundant subcarriers have the same mean power value or variance of $\sigma_r^2 = \sigma_d^2 \frac{N_d}{N_r}$, leading to

$$E_r = \frac{\sigma_d^2}{N} \mathrm{tr}\left(\mathbf{T}\mathbf{T}^H\right) = \sigma_r^2 \frac{N_r}{N} = \sigma_d^2 \frac{N_d}{N_r} \frac{N_r}{N} = \sigma_d^2 \frac{N_d}{N} = E_d. \tag{3.12}$$

This equality only holds approximately for setup A with $E_r = \frac{36.56}{N} \approx E_d = \frac{36}{N}$. These results suggest that the optimum distribution of the redundant subcarriers leads to UW-OFDM symbols, where on average (at least approximately) half the transmit energy is spent for data and half the energy is spent for redundant subcarriers.

These findings motivate a more in-depth analysis of the optimal distribution of the redundant subcarriers and the resulting energy. The author in [53] shows that for a UW-OFDM setup without zero subcarriers, the minimum redundant energy as a function of the number of redundant symbols is symmetrical in N_r around $N/2$, i.e.,

$$E_r(N_r) = E_r(N - N_r). \tag{3.13}$$

This can exemplarily be seen in the upper plot of Fig. 3.5 for varying N_r and the case $N_r = N_u$. Note that a scaling factor N has been introduced in the plot to

easier link the energy values to N_r and N_d. It turns out that (3.13) does not hold in the presence of zero subcarriers as depicted in the lower plot of Fig. 3.5. In this case, a zero subcarrier has been introduced at DC. Although not explicitly illustrated, experiments confirm the loss of the symmetry property for any setup with zero subcarrier(s), regardless of the actual number and/or the placement of them.

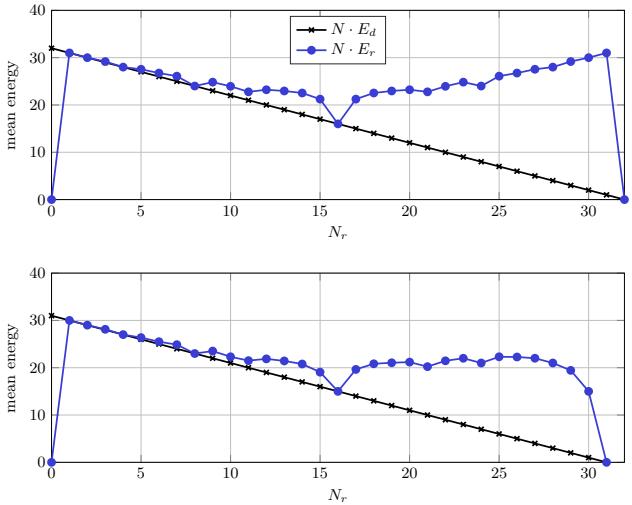


Figure 3.5.: Mean redundant energy E_r for varying N_r with $N_r = N_u$, evaluated for the optimal set of an UW-OFDM setup with $N = 32$ and either $N_z = 0$ (upper plot) or $N_z = 1$ with a zero at DC (lower plot).

Furthermore, [53] also presents an analytical lower bound for the redundant energy

$$E_r \leq \frac{N_d N_u}{N_r} E_s, \tag{3.14}$$

with $E_s = E_{x'}/N_d$ denoting the mean energy per data symbol in case of a zero UW. The bound is met with equality in case N_r is a power of 2. This result can even be generalized to the case

$$\mathrm{mod}(N, N_r) = 0 \qquad \wedge \qquad \mathcal{I}_r = \left\{ l\frac{N}{N_r} + n_0 \right\} \qquad l = 0 \ldots N_r - 1, \tag{3.15}$$

with any constant offset $n_0 \in \{0 \ldots N/N_r - 1\}$. It follows that as long as N is a multiple of N_r, and additionally the N_r redundant subcarriers can equidistantly be distributed with distance N/N_r, the lower bound is exactly met. In order to

support this statement, Fig. 3.6 depicts six exemplary UW-OFDM setups. All six share the same DFT size of $N = 27$ and the same number of $N_r = 9$ redundant subcarriers, but may differ in the cardinality N_z as well as in the distribution of the zero subcarriers. In contrast to all others, setup 1 does not incorporate any zero subcarriers in the spectrum at all. It is obvious that here a redundant subcarrier features an average power of 2, thus leading to $E_r = N_r \cdot 2/N = 18/N$, which corresponds to the energy[1] of the data symbols $E_d = N_d \cdot 1/N = 18/N$. Setup 2 and setup 3 incorporate $N_z = 3$ and $N_z = 8$ zero and therefore $N_d = 15$ and $N_d = 10$ data subcarriers, respectively. Different to the first setup, the average power varies among the redundant subcarriers in both cases. Although not explicitly extractable from Fig. 3.6 due to a granular resolution of the ordinate, the bound in (3.14) is exactly met with $E_r = 15/N$ and $E_r = 10/N$, respectively. For the latter this can clearly be seen in Fig. 3.7 by evaluating E_r at $N_r = 9$. Despite completely different configurations, $E_r = E_d$ holds in all three cases.

Setup 3 and 4 share the same system parameter values, both have $N = 27$, $N_r = 9$ and $N_z = 8$. Even the distribution of the zero subcarriers is identical except for a single one which is shifted from index 10 to 11. This minor difference has a remarkable impact though, as it prevents an equidistant arrangement with spacing N/N_r of the redundant subcarriers over the available spectrum. The consequence is that $E_r > E_d$ as observable in Fig. 3.7 when considering E_r of setup 4 at $N_r = 9$.

Setup 4 and setup 5 have the same number of zero subcarriers, in the latter case, however, these zeros appear in a bundle instead of in a distributed manner. Completely unexpected, Fig. 3.8 illustrates that this even leads to $E_r < E_d$ in some cases. Various experiments indicate that the basic effect is independent of the position of the zero subcarrier bundle. Although not proven theoretically and only supported by means of simulations, the difference between E_r and E_d seems to scale with the length of a zero subcarrier bundle, i.e., the longer the bundle, the larger the difference. A comparison of setup 5 and 6 in Fig. 3.6 and Fig. 3.8 should exemplarily confirm this observation. Furthermore, simulations suggest that the effect of $E_r < E_d$ only appears, if the number of concatenated zero subcarriers is at least $\lceil N/N_r \rceil$.

[1] The factor $\frac{1}{N}$ accounts for the transform to the time domain (motivated by applying the Parseval theorem) as a consequence of the utilized definition of the N-point DFT in (2.12).

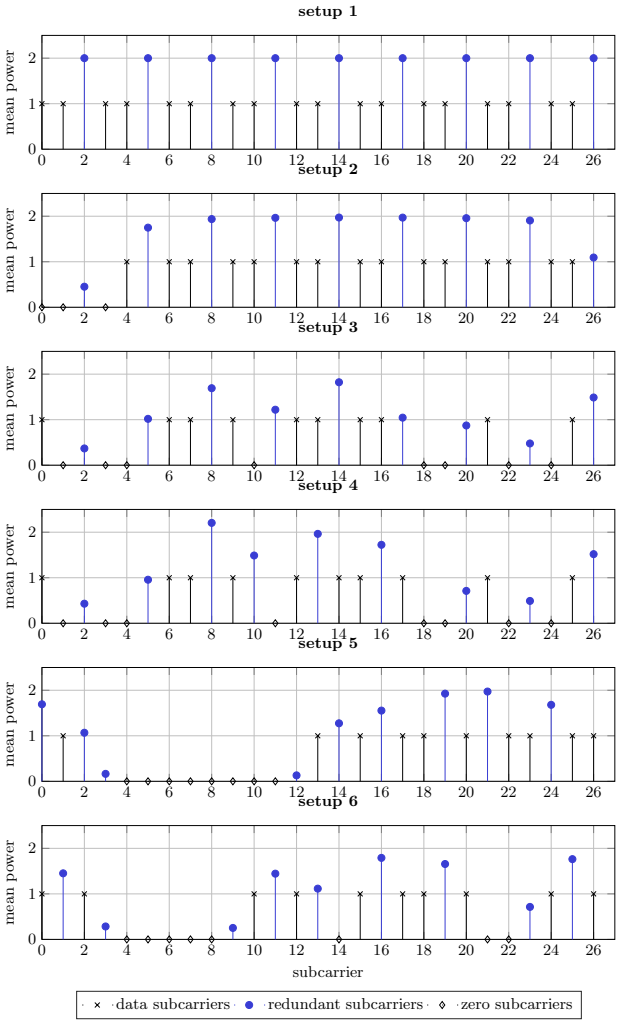


Figure 3.6.: Mean power on individual subcarriers for UW-OFDM setups with $N = 27$, $N_r = 9$, and varying number N_z and placement of zero subcarriers.

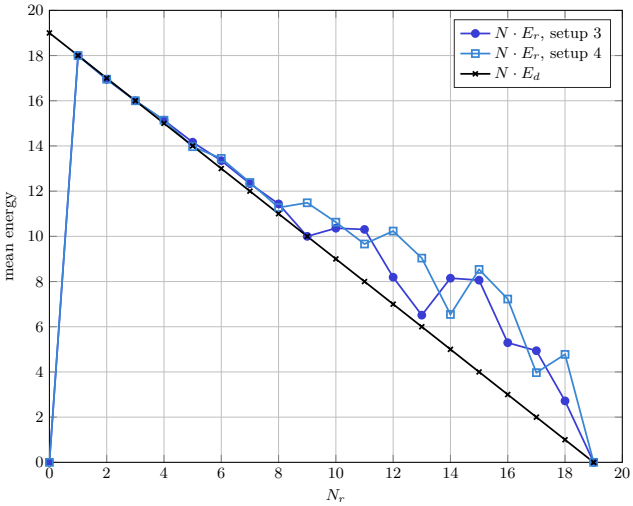


Figure 3.7.: Mean data energy E_d and mean redundant energy E_r for setup 3 and 4 as defined in Fig. 3.6.

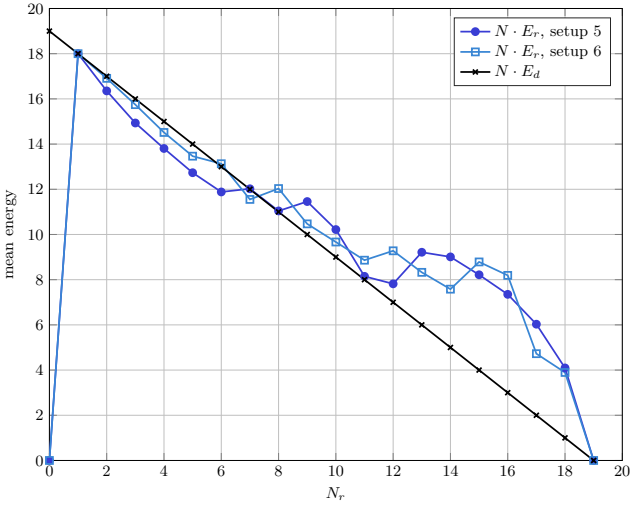


Figure 3.8.: Mean data energy E_d and mean redundant energy E_r for setup 5 and 6 as defined in Fig. 3.6.

In summary, the following statements regarding the placement of the redundant subcarriers and the resulting energy can be drawn:

- The energy E_r of the redundant subcarriers is symmetric in N_r around $N/2$ as described in (3.13). This symmetry is lost as soon as any zero subcarrier is present in the system.
- The energy E_r of the redundant subcarriers equals the energy E_d of the data symbols as long as the redundant subcarriers can be arranged equidistantly with an integer spacing of N/N_r.
- The energy E_r is always larger than E_d, if the redundant subcarriers cannot be arranged equidistantly with an integer spacing of N/N_r and no zero subcarriers are present in the system.
- The energy E_r is always larger than E_d, if the redundant subcarriers cannot be arranged equidistantly with an integer spacing of N/N_r and possible zero subcarriers do not appear in bundles.
- The energy E_r may be smaller than E_d, if the redundant subcarriers cannot be arranged equidistantly with an integer spacing of N/N_r and possible zero subcarriers appear in a bundle.
- In case that $E_r < E_d$, the difference between the energy values seems to increase with increasing bundle size.

Properties drawn from single redundant subcarrier symbols. In systematically encoded UW-OFDM the code words $\mathbf{c}$ contain data symbols d_i drawn from probability mass functions (PMFs) given by the underlying finite complex alphabet. This is identical to conventional OFDM, however, there are also redundant subcarrier symbols r_i. These symbols result from a weighted sum of data symbols, thus motivating a more in-depth analysis. The following investigations assume iid data symbols with zero mean and variance σ_d^2, and modulation alphabets that produce data vectors with rotation invariant elements as defined in [54]. This includes constellations like QPSK, 16-QAM and 64-QAM, but excludes e.g., BPSK modulation. With these assumptions $\mathbf{r} = \mathbf{T}\mathbf{d}$ has zero mean and a covariance matrix $\mathbf{C}_{rr} = \mathrm{E}\left\{\mathbf{r}\mathbf{r}^H\right\} = \sigma_d^2\mathbf{T}\mathbf{T}^H$, and $\mathbf{r}$ fulfills the rotation invariance condition as well, as it is preserved by linear complex transformations. A redundant symbol r_i is generated by a weighted sum of N_d iid data symbols, consequently, r_i is a discrete complex valued random variable (RV). Nevertheless, in the following r_i is approximated by a continuous RV represented by a probability density function (PDF). Following CLT arguments, the PDF of r_i may be well approximated by a complex Gaussian PDF $\mathcal{CN}(0, \sigma_{r_i}^2)$ with zero mean and variance $\sigma_{r_i}^2 = [\mathbf{C}_{rr}]_{ii}$:

$$p_{r_i}(r_i) = \frac{1}{\pi\sigma_{r_i}^2} \cdot \exp\left(-\frac{1}{\sigma_{r_i}^2}|r_i|^2\right). \tag{3.16}$$

Due to the rotational invariance condition, the real and imaginary part of r_i (represented by u_i and v_i, respectively) are uncorrelated, and the PDFs of both can be approximated by $\mathcal{N}(0, \sigma_{r_i}^2/2)$. The upper plot in Fig. 3.9 displays an estimated PDF for the real part of the code word symbol c_1, which corresponds to r_0 in case of setup A, by plotting a distribution of relative frequencies. These relative frequencies have been obtained from simulating 10^5 UW-OFDM symbols with QPSK data symbols. Note that $\sigma_{r_0}^2 = [\mathbf{C}_{rr}]_{00} = 2.47$. It turns out that an approximation with a Gaussian distribution $\mathcal{N}(0, \sigma_{r_0}^2/2)$ per quadrature component perfectly matches the simulation. The lower plot of Fig. 3.9 shows the PMF of the real part of a data symbol, in this case of c_0. Consequently, there is an important difference between conventional OFDM and systematically encoded UW-OFDM when considering the energy of one OFDM symbol. In case all elements of a modulation alphabet have the same power, e.g., as it is for QPSK but not for 16-QAM, the energy of a conventional OFDM symbol is constant and the same for every symbol. This is clearly not the case for systematically encoded UW-OFDM, since the energy of the redundant subcarriers generally varies from symbol to symbol.

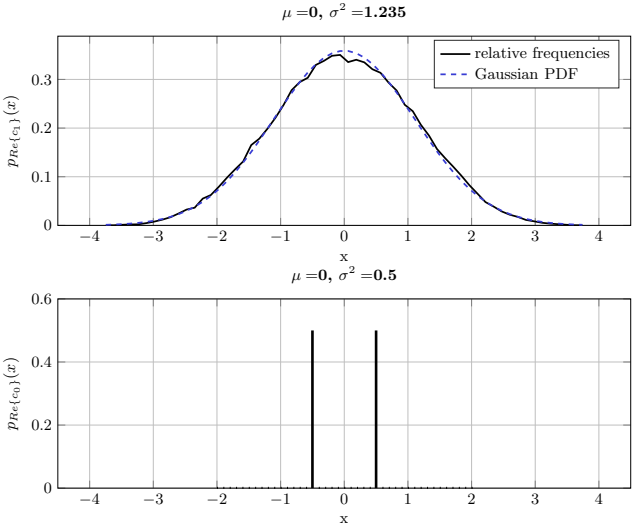


Figure 3.9.: Upper plot: Distribution of relative frequencies to estimate the PDF of the real part of c_1. A fitted Gaussian PDF serves as reference. Lower plot: PMF of the real part of c_0.

The next step investigates the resulting energy distribution of systematically encoded UW-OFDM symbols. The transmit symbol energy for one dedicated realization

of an OFDM symbol is given by

$$E_{\mathbf{x}''}^{(r)} = \frac{1}{N}\mathbf{d}^H\mathbf{d} + \frac{1}{N}\mathbf{r}^H\mathbf{r} + \mathbf{x}_u^H\mathbf{x}_u. \tag{3.17}$$

The interesting part worth investigating here is the energy contribution[2]

$$y = \mathbf{r}^H\mathbf{r} = \sum_{i=0}^{N_r-1} |r_i|^2 \tag{3.18}$$

of the redundant subcarriers. The discrete RV y is again approximated by a continuous one, and approximate PDFs are derived. The PDF $p_{x_i}(x)$ of $x_i = |r_i|^2 = u_i^2 + v_i^2$ corresponding to the squared magnitude of the ith redundant subcarrier symbol serves as a starting point. Since u_i and v_i denote uncorrelated random variables which are both approximated by a Gaussian PDF $\mathcal{N}(0, \sigma_{r_i}^2/2)$, it immediately follows that $p_{x_i}(x)$ can be well approximated by an exponential distribution of the form

$$p_{x_i}(x) = \frac{1}{\sigma_{r_i}^2}\mathrm{e}^{-x/\sigma_{r_i}^2} \quad \text{for} \quad x \geq 0, \tag{3.19}$$

cf. [55]. Note that the redundant subcarrier symbols are correlated among each other, leading to non-zero off-diagonal elements in $\mathbf{C}_{rr}$. For the approximation of the PDF of y the off-diagonal elements in $\mathbf{C}_{rr}$ are ignored and uncorrelated redundant subcarrier symbols r_i are assumed, cf. [56]. Because of the Gaussian model the r_i's can further be assumed to be iid, consequently the x_i's can be assumed to be iid as well. Making these assumptions, the PDF of the sum $y = \mathbf{r}^H\mathbf{r}$ results in an $(N_r - 1)$-fold convolution of the PDFs $p_{x_i}(x)$:

$$p_y(y) = p_{x_1}(y) * p_{x_2}(y) * \cdots * p_{x_{N_r}}(y). \tag{3.20}$$

Due to its practical relevance, the general case with zero subcarriers is considered in the following. The variance of the redundant symbols is not constant in this case (Fig. 3.4), but due to symmetry properties the redundant symbols r_i and r_{N_r-i-1} for $i = 0, 1, \cdots N_r/2 - 1$ feature the same variance. Therefore, y can also be written as $y = \sum_{i=0}^{N_r-1} x_i = \sum_{i=0}^{N_r/2-1} z_i$ with $z_i = x_i + x_{N_r-i-1}$ and $z_i \sim \mathrm{Erl}(2, \beta_i)$, where $\mathrm{Erl}(\cdot)$ denotes an Erlang distribution, $\beta_i = 1/\sigma_{r_i}^2$ and $\beta_i \neq \beta_j$ for $i \neq j$. An analytical expression of the approximate PDF of y is given by

$$\begin{aligned} p_y(y) = &\sum_{i=0}^{N_r/2-1} \beta_i^2 e^{-\beta_i y} \sum_{j=1}^{2} \frac{(-1)^{2-j}}{(j-1)!} y^{j-1} \\ &\times \sum_{\substack{m_0+m_1+\cdots+m_{N_r/2-1}=2-j \\ m_i=0}} \prod_{\substack{l=0 \\ l\neq i}}^{N_r/2-1} \binom{2+m_l-1}{m_l} \frac{\beta_l^2}{(\beta_l-\beta_i)^{2+m_l}} \end{aligned} \tag{3.21}$$

[2] Note that for the actual energy, a factor $\frac{1}{N}$ is missing. This factor arises when going from the frequency to the time domain (motivated by applying the Parseval theorem) as a consequence of the utilized definition of the DFT in (2.12). However, this factor has been intentionally left out for reasons of a compact notation.

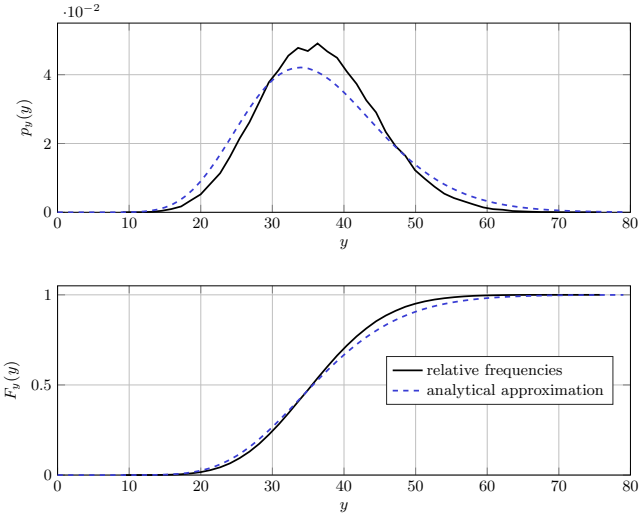


Figure 3.10.: Upper plot: Distribution of the relative frequencies and analytical approximation to estimate the PDF of the redundant energy. Lower plot: Cumulative frequency distribution to estimate the CDF of the redundant energy.

for $y > 0$, cf. [57]. This distribution is also known as a particular form of the generalized chi-squared distribution. Fig. 3.10 depicts the estimated and the analytically approximated PDFs and cumulative distribution functions (CDFs) of the redundant energy. Again the estimated PDF is obtained from simulating 10^5 UW-OFDM symbols with QPSK data symbols. The analytical approximation has been derived by evaluating (3.21). The reason for the deviation of the analytical approximation from the estimated PDF mainly originates from the simplified assumption of uncorrelated redundant symbols. The true mean of the redundant subcarriers' energy contribution is $\mathrm{E}\{y\} = \sigma_d^2 \mathrm{tr}\left(\mathbf{T}\mathbf{T}^H\right) = 36.5658$, calculating the mean with the help of the analytical approximation yields 36.5573, and using the PDF derived by the particular simulation run yields 36.5557.

3.2. Addition of UW: Two-Step versus Direct Approach

A UW-OFDM time domain signal is by definition of the form $\begin{bmatrix}\mathbf{x}_p^T & \mathbf{x}_u^T\end{bmatrix}^T \in \mathbb{C}^{N\times 1}$, with $\mathbf{x}_p$ representing the payload and $\mathbf{x}_u$ determining the desired unique word. There are in principle two ways, known as *direct* and *two-step approach* [58], to generate a UW at the tail of the time domain symbol. Both approaches yield

different UW-OFDM symbol waveforms, since the $\mathbf{x}_p$ sections differ significantly. It has already been mentioned in the first chapter that the two-step approach is the preferable one, which generates a zero UW in the first step, and then adds the desired UW in a second step. This section will now verify this statement by proving the two-step approach to generate OFDM symbols with much less redundant energy than a single step or direct UW generation approach. The difference might be surprising at this point, as the DFT is a linear operation and consequently the result should be invariant to an addition (superposition principle). In fact, it will turn out that both approaches differ in the circumstance whether the UW influences the data subcarriers.

3.2.1. Symbol Energy for Direct Approach

The straightforward or direct approach to produce UW-OFDM symbols with a unique word tail is to force $\mathbf{F}_N^{-1}\tilde{\mathbf{x}}'' = \mathbf{x}''$, or equivalently

$$\mathbf{x}'' = \mathbf{F}_N^{-1}\mathbf{B}\mathbf{P}\begin{bmatrix}\mathbf{d}\\ \mathbf{r}''\end{bmatrix} = \begin{bmatrix}\mathbf{x}_p''\\ \mathbf{x}_u\end{bmatrix}. \tag{3.22}$$

This approach is in fact closely related to the proposal in [32]. With (3.22) and the definition $\mathbf{M} = \mathbf{F}_N^{-1}\mathbf{B}\mathbf{P} = \left[\begin{smallmatrix}\mathbf{M}_{11} & \mathbf{M}_{12}\\ \mathbf{M}_{21} & \mathbf{M}_{22}\end{smallmatrix}\right]$ from (2.44) for $\mathbf{A} = \mathbf{P}$, it immediately follows that

$$\mathbf{M}_{21}\mathbf{d} + \mathbf{M}_{22}\mathbf{r}'' = \mathbf{x}_u. \tag{3.23}$$

Note that $\mathbf{M}_{22}$ is quadratic with permuted Vandermonde structure, hence it is always invertible. With $\mathbf{T} \in \mathbb{C}^{N_r \times N_d}$ as defined in (2.47) the vector of redundant subcarrier symbols can thus be determined from the data vector $\mathbf{d}$ and the unique word $\mathbf{x}_u$ by

$$\mathbf{r}'' = \mathbf{T}\mathbf{d} + \mathbf{M}_{22}^{-1}\mathbf{x}_u. \tag{3.24}$$

By inserting (3.24) into (3.22), the transmit time domain symbol $\mathbf{x}''$ follows as

$$\mathbf{x}'' = \mathbf{F}_N^{-1}\mathbf{B}\mathbf{P}\begin{bmatrix}\mathbf{I}\\ \mathbf{T}\end{bmatrix}\mathbf{d} + \mathbf{F}_N^{-1}\mathbf{B}\mathbf{P}\begin{bmatrix}\mathbf{0}\\ \mathbf{M}_{22}^{-1}\end{bmatrix}\mathbf{x}_u. \tag{3.25}$$

The first term in (3.25) produces a zero word guard space, while the second term generates the desired UW. Note that the second term only affects the redundant subcarriers, while the data subcarriers remain untouched. As a preparatory step for comparing both approaches, the mean UW-OFDM symbol energy $E_{\mathbf{x}''} = \mathrm{E}\left\{\mathbf{x}''^H\mathbf{x}''\right\} = \frac{1}{N}\mathrm{E}\left\{\tilde{\mathbf{x}}''^H\tilde{\mathbf{x}}''\right\}$ is investigated. Assuming uncorrelated data symbols from an alphabet $\mathcal{A}$ with zero mean and covariance matrix $\mathbf{C}_{dd} = \mathrm{E}\left\{\mathbf{d}\mathbf{d}^H\right\} = \sigma_d^2\mathbf{I}$, and exploiting the fact that the linear mapping $\tilde{\mathbf{x}}'' = \mathbf{B}\mathbf{P}\left[\begin{smallmatrix}\mathbf{d}\\ \mathbf{r}''\end{smallmatrix}\right]$ does not change

the mean energy of the vector $\left[\begin{smallmatrix}\mathbf{d}\\ \mathbf{r}''\end{smallmatrix}\right]$, yields

$$\begin{aligned} E_{\mathbf{x}''} &= \frac{1}{N}\mathrm{E}\left\{\begin{bmatrix}\mathbf{d}^H & \mathbf{r}''^H\end{bmatrix}\begin{bmatrix}\mathbf{d}\\ \mathbf{r}''\end{bmatrix}\right\} \\ &= \frac{1}{N}\mathrm{E}\left\{\mathbf{d}^H\mathbf{d}\right\} + \frac{1}{N}\mathrm{E}\left\{\mathbf{r}''^H\mathbf{r}''\right\} \\ &= \frac{N_d\sigma_d^2}{N} + \frac{1}{N}\mathrm{E}\left\{\mathbf{r}''^H\mathbf{r}''\right\}. \end{aligned} \tag{3.26}$$

With (3.24) the term $\mathrm{E}\left\{\mathbf{r}''^H\mathbf{r}''\right\}$ can be rewritten as

$$\begin{aligned} \mathrm{E}\left\{\mathbf{r}''^H\mathbf{r}''\right\} &= \mathrm{E}\left\{\mathrm{tr}\left(\mathbf{r}''\mathbf{r}''^H\right)\right\} \\ &= \mathrm{tr}\left(\mathrm{E}\left\{\mathbf{r}''\mathbf{r}''^H\right\}\right) \\ &= \mathrm{tr}\left(\mathrm{E}\left\{(\mathbf{T}\mathbf{d}+\mathbf{M}_{22}^{-1}\mathbf{x}_u)(\mathbf{T}\mathbf{d}+\mathbf{M}_{22}^{-1}\mathbf{x}_u)^H\right\}\right) \\ &= \mathrm{tr}\Big(\mathrm{E}\Big\{\mathbf{T}\mathbf{d}\mathbf{d}^H\mathbf{T}^H + \mathbf{M}_{22}^{-1}\mathbf{x}_u\mathbf{d}^H\mathbf{T}^H \\ &\qquad + \mathbf{T}\mathbf{d}\mathbf{x}_u^H(\mathbf{M}_{22}^{-1})^H + \mathbf{M}_{22}^{-1}\mathbf{x}_u\mathbf{x}_u^H(\mathbf{M}_{22}^{-1})^H\Big\}\Big) \\ &= \sigma_d^2\mathrm{tr}\left(\mathbf{T}\mathbf{T}^H\right) + \mathrm{tr}\left(\mathbf{M}_{22}^{-1}\mathbf{x}_u\mathbf{x}_u^H(\mathbf{M}_{22}^{-1})^H\right), \end{aligned}$$

with the last step originating from $\mathrm{E}\{\mathbf{d}\} = \mathbf{0}$. Finally, the symbol energy is

$$E_{\mathbf{x}''} = \underbrace{\frac{N_d\sigma_d^2}{N}}_{E_d} + \underbrace{\frac{\sigma_d^2}{N}\mathrm{tr}\left(\mathbf{T}\mathbf{T}^H\right)}_{E_r} + \underbrace{\frac{1}{N}\mathbf{x}_u^H(\mathbf{M}_{22}^{-1})^H\mathbf{M}_{22}^{-1}\mathbf{x}_u}_{E_u}. \tag{3.27}$$

Here, E_d describes the mean contribution of the data subcarrier symbols, E_r depicts the mean contribution of the redundant subcarrier symbols for the case $\mathbf{x}_u = \mathbf{0}$, and E_u denotes the contribution of a non-zero UW. E_d can be treated as fixed and given. E_r and E_u strongly depend on the permutation matrix $\mathbf{P}$ (since $\mathbf{T}$ and $\mathbf{M}_{22}$ depend on $\mathbf{P}$), E_u additionally depends on the particular shape of $\mathbf{x}_u$. E_r can reasonably be minimized by a proper choice of $\mathbf{P}$. A comparison of the initial results in Tab. 2.1 with the results gained in Sec. 3.1 confirm this statement. Hence, E_r should therefore be of a minor problem, leaving E_u as the last subject for investigation in the following two sections.

3.2.2. Symbol Energy for Two-Step Approach

In order to get rid of the term $\mathbf{x}_u^H(\mathbf{M}_{22}^{-1})^H\mathbf{M}_{22}^{-1}\mathbf{x}_u$ in (3.27), we proposed a simple, yet highly efficient approach in [59],[60], which we call two-step approach. As already motivated in chapter 2, the idea is to first generate a zero word such that

$\mathbf{x} = \begin{bmatrix}\mathbf{x}_p^T & \mathbf{0}^T\end{bmatrix}^T$, and in a second step to generate the desired output symbol $\mathbf{x}' = \mathbf{x} + \begin{bmatrix}\mathbf{0}^T & \mathbf{x}_u^T\end{bmatrix}^T$.

The first step is a special case of the direct approach described in the previous section for $\mathbf{x}_u = \mathbf{0}$. According to (3.24) the redundant subcarrier symbols result from the linear mapping

$$\mathbf{r} = \mathbf{T}\mathbf{d}. \tag{3.28}$$

With (2.37), (3.6) and $\tilde{\mathbf{x}}_u$, the transmit symbol $\mathbf{x}'$ is given as

$$\mathbf{x}' = \mathbf{F}_N^{-1}\mathbf{B}\mathbf{P}\begin{bmatrix}\mathbf{I}\\ \mathbf{T}\end{bmatrix}\mathbf{d} + \mathbf{F}_N^{-1}\tilde{\mathbf{x}}_u. \tag{3.29}$$

A comparison of (3.25) with (3.29) reveals that the first term generating the zero UW is identical for both approaches, while the second term generating the actually desired UW differs. Moreover, the second term in (3.25) only affects the redundant subcarriers, whereas $\tilde{\mathbf{x}}_u$ in (3.29) may in general overlay the redundant subcarrier symbols $\mathbf{r}$ as well as the data symbols $\mathbf{d}$. Hence, in the two-step approach the UW distorts the data symbols and thus requires a correct subtraction of the UW before data detection, which is not necessary in case of the direct approach. However, $\tilde{\mathbf{x}}_u$ is deterministic and known, and these distortions can therefore easily be reversed at the receiver, cf. (2.66).

As for the direct approach, the mean transmit symbol energy $E_{\mathbf{x}'}$ is evaluated in the following. With (3.27) (and still having in mind that the first step of the two-step approach is a special case of the direct approach for $\mathbf{x}_u = \mathbf{0}$) the mean energy $E_{\mathbf{x}} = \mathrm{E}\left\{\mathbf{x}^H\mathbf{x}\right\}$ becomes

$$E_{\mathbf{x}} = \frac{N_d\sigma_d^2}{N} + \frac{\sigma_d^2}{N}\mathrm{tr}\left(\mathbf{T}\mathbf{T}^H\right). \tag{3.30}$$

Further, as the terms $\mathbf{x}$ and $\begin{bmatrix}\mathbf{0}^T & \mathbf{x}_u^T\end{bmatrix}^T$ in (2.37) are orthogonal, the mean transmit symbol energy $E_{\mathbf{x}'}$ immediately follows to

$$E_{\mathbf{x}'} = \underbrace{\frac{N_d\sigma_d^2}{N}}_{E_d} + \underbrace{\frac{\sigma_d^2}{N}\mathrm{tr}\left(\mathbf{T}\mathbf{T}^H\right)}_{E_r} + \underbrace{\mathbf{x}_u^H\mathbf{x}_u}_{E_{x_u}}. \tag{3.31}$$

E_d and E_r describe the contributions of the data and the redundant subcarrier symbols to the total mean symbol energy before the addition of the UW, respectively, and E_{x_u} describes the contribution of the UW. Note that different to the direct approach in (3.27), the energy contribution of the UW to the total mean transmit symbol energy is now only determined by the energy of the UW $E_{x_u} = \mathbf{x}_u^H\mathbf{x}_u$ itself, but not by its particular shape.

3.2.3. Comparison of Both Approaches

This section provides the analytical proof of $E_{\mathbf{x}''} \geq E_{\mathbf{x}'}$ for all possible UWs and for all possible permutation matrices $\mathbf{P}$. This is achieved by proving that $E_u \geq E_{x_u}$, cf. (3.27) and (3.31), or equivalently

$$\frac{1}{N}\mathbf{x}_u^H(\mathbf{M}_{22}^{-1})^H\mathbf{M}_{22}^{-1}\mathbf{x}_u \geq \mathbf{x}_u^H\mathbf{x}_u. \tag{3.32}$$

Note that the derivations are restricted to the case of $N_r = N_u$, however, the following proof also holds in case of $N_r > N_u$, cf. [61] and Sec. 3.6, where $\mathbf{M}_{22}$ becomes rectangular and $\mathbf{M}_{22}^{-1}$ has to be replaced by the Moore-Penrose Pseudo-Inverse $\mathbf{M}_{22}^{\dagger}$.

In fact, for any vector $\mathbf{x}_u$ it holds that

$$\begin{aligned}
\mathbf{x}_u^H\mathbf{x}_u &= \|\mathbf{x}_u\|_2^2 \\
&= \left\|\mathbf{M}_{22}\mathbf{M}_{22}^{-1}\mathbf{x}_u\right\|_2^2 \\
&\leq \|\mathbf{M}_{22}\|_S^2 \left\|\mathbf{M}_{22}^{-1}\mathbf{x}_u\right\|_2^2 && (3.33)\\
&\leq \left\|\mathbf{F}_N^{-1}\right\|_S^2 \left\|\mathbf{M}_{22}^{-1}\mathbf{x}_u\right\|_2^2 && (3.34)\\
&= \frac{1}{N}\left\|\mathbf{M}_{22}^{-1}\mathbf{x}_u\right\|_2^2 && (3.35)\\
&= \frac{1}{N}\mathbf{x}_u^H(\mathbf{M}_{22}^{-1})^H\mathbf{M}_{22}^{-1}\mathbf{x}_u.
\end{aligned}$$

Clearly, equality is given for $\mathbf{x}_u = \mathbf{0}$. Equation (3.33) is true since the Euclidean vector norm $\|\cdot\|_2$ is compatible with the spectral matrix norm $\|\cdot\|_S$, and for a vector norm being compatible to a matrix norm, the inequality $\|\mathbf{Ax}\| \leq \|\mathbf{A}\|\,\|\mathbf{x}\|$ holds for every square matrix $\mathbf{A}$ and every vector $\mathbf{x}$ (which match in their dimensions), cf. [62]. To (3.34) and (3.35) we note that the spectral matrix norm $\|\mathbf{A}\|_S$ is defined as $\|\mathbf{A}\|_S = \sqrt{\lambda_{max}\left(\mathbf{A}^H\mathbf{A}\right)}$, which denotes the largest singular value of $\mathbf{A}$, and consequently $\lambda_{max}\left(\mathbf{A}^H\mathbf{A}\right)$ the largest eigenvalue of $\mathbf{A}^H\mathbf{A}$. The spectral norm of any submatrix cannot exceed the spectral norm of the matrix it has been extracted from, cf. [63]. Since $\mathbf{M}_{22}$ is a submatrix of $\mathbf{F}_N^{-1}$, we have $\|\mathbf{M}_{22}\|_S \leq \left\|\mathbf{F}_N^{-1}\right\|_S$. For the IDFT matrix $\mathbf{F}_N^{-1}$, the spectral norm becomes $\left\|\mathbf{F}_N^{-1}\right\|_S = \frac{1}{\sqrt{N}}$.

In [60] we compared E_u and E_{x_u} for various potential UW sequences and observed that the direct approach requires substantially more energy to generate a desired UW in time domain. Furthermore, E_u heavily varies with the shape of the UW. To conclude, the two-step approach resolves this problem, as it allows to get rid of the problematic term $\mathbf{x}_u^H(\mathbf{M}_{22}^{-1})^H\mathbf{M}_{22}^{-1}\mathbf{x}_u$ in (3.27), and the mean transmit symbol energy becomes independent of the particular shape of the UW. As a consequence of the two-step approach, the UW also uses now the data spectrum.

The derivations in Sec. 2.2 show that the influence of the UW has to be eliminated before data estimation is applied on the linear model in (2.68). The UW does not provide any gain w.r.t. the data estimation task. From this point of view, it is thus desirable to spend as less energy as possible for a UW. Taking into account that the very same linear model is achieved for the direct approach when replacing $\tilde{\mathbf{x}}_u$ in (2.60) by

$$\tilde{\mathbf{x}}'_u = \mathbf{BP} \begin{bmatrix} \mathbf{0} \\ \mathbf{M}_{22}^{-1} \end{bmatrix} \mathbf{x}_u, \tag{3.36}$$

both approaches show the same data estimation performance (e.g., in terms of MSE). However, this comes in case of the direct approach at the price of a much higher transmit energy. Hence, the two-step approach will always outperform the direct approach in terms of the BER-over-E_b/N_0 performance.

To provide a feeling for the performance difference between both approaches, $E_u = \frac{1}{N}\mathbf{x}_u^H(\mathbf{M}_{22}^{-1})^H\mathbf{M}_{22}^{-1}\mathbf{x}_u$ from the direct approach and $E_{x_u} = \mathbf{x}_u^H\mathbf{x}_u$ from the two-step approach are evaluated for various potential UW sequences, cf. [60]. The UW will be normalized to $E_{x_u} = \mathbf{x}_u^H\mathbf{x}_u = 1$ for reasons of graphical representation. Fig. 3.11 summarizes the results for the following UW sequences:

1. The generalized Barker sequence [64] of length 12 padded with zeros to the final length of 16;
2. A CAZAC sequence (constant amplitude, zero autocorrelation) from [65];
3. The length 16 Frank-Zadoff sequence from [66], which also has CAZAC properties.

The direct approach obviously requires substantially more energy to generate a desired UW in time domain than the two-step approach. Furthermore, E_u varies with the shape of the UW. In summary, the two-step approach resolves the problem of the energy contribution of the UW in two ways:

- the problematic term $\mathbf{x}_u^H(\mathbf{M}_{22}^{-1})^H\mathbf{M}_{22}^{-1}\mathbf{x}_u$ in (3.27) is avoided, and
- the mean transmit symbol energy becomes independent of the particular shape of the UW.

3.3. Optimization Algorithm for Redundant Subcarrier Distribution

In our previous works, e.g., in [59],[41], we have stressed the necessity of finding the optimum positions of the redundant subcarrier symbols and hence the optimum permutation matrix $\mathbf{P}$. The results in Tab. 2.1 strongly suggest an optimization w.r.t. minimizing the mean redundant energy.

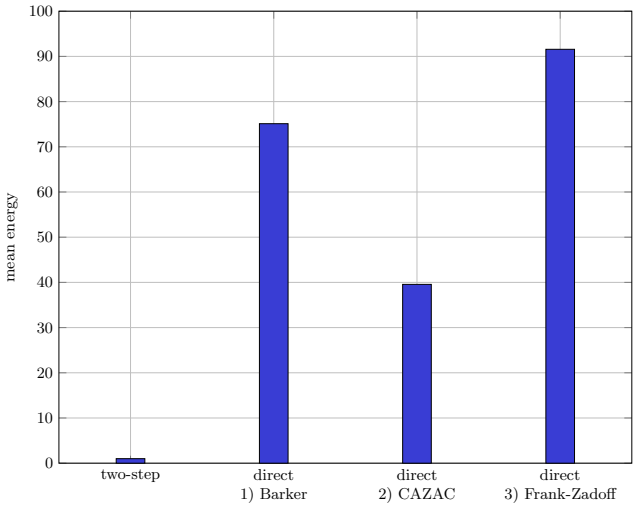


Figure 3.11.: Mean energy contribution of non-zero UWs in the direct (E_u) and the two-step approach (E_{x_u}).

Unfortunately, only for the special case that the redundant subcarriers can be equidistantly arranged with integer spacing N/N_r, cf. (3.15), an optimal solution can be derived analytically. For all other cases, an optimum by means of analytical derivations is not available. Experiments summarized in Fig. 3.6–3.8 suggest though that an optimum or near-to-optimum redundant subcarrier distribution will lead to $E_r = E_d + E_\Delta$ with a relatively small E_Δ (at least for $N_u < N/2$), however, the exact value of E_Δ and thus the global optimum remains unknown. To make matters worse, an exhaustive search is computationally unmanageable for reasonable choices of N and N_r (setup A already provides $\binom{(N-N_z)}{N_r} = \binom{52}{16} \approx 1.04 \cdot 10^{13}$ different combinations).

3.3.1. Split Distribution

In order to circumvent this computational burden, [51] presents a so-called split distribution as a means to select the redundant subcarrier positions according to a simple analytical expression. In this case N_r is modelled as the sum of powers of 2 such that

$$N_r = \sum_{l=0}^{L_r-1} N_{r,l}, \tag{3.37}$$

whereas $N_{r,l} = 2^{x_l}$, $x_l \in \mathbb{N}_0$ with $\mathbb{N}_0$ denoting the set of natural numbers including 0 and $N_{r,1} > N_{r,2} > \cdots > N_{r,L_r}$. The set $\mathcal{I}_r$ of N_r subcarrier indices is split into L_r subsets $\mathcal{I}_{r,l}$ consisting of $N_{r,l}$ redundant subcarrier positions. Each subset follows a uniform distribution with spacing $\Delta_l = \frac{N}{N_{r,l}}$ and an integer constant $n_{0,l}$ such that

$$\mathcal{I}_{r,l} = \{n_{0,l} + m\Delta_l\} \qquad m = 0, \ldots, N_{r,l} - 1. \tag{3.38}$$

A comparison of the results in Fig. 3.5 and [53] demonstrates that in most cases the split distribution is far away from delivering the optimum distribution. In fact the resulting redundant energy E_r exceeds the optimum by 50-150% and is thus not applicable in practice.

3.3.2. Quasi-Uniform Distribution

In [67] another analytical approach called quasi-uniform (QU) distribution is introduced. Here, the redundant carriers are generated according to

$$\mathcal{I}_r = \left\{ \left[\frac{lN}{N_r} \right]_R \right\} \qquad l = 0, \ldots, N_r - 1, \tag{3.39}$$

whereas $[x]_R$ rounds to the nearest integer. This method achieves optimum or at least near-optimum results, but works only as long as no subcarriers are excluded from the optimization process [68]. This might occur, if e.g., pilot or zero subcarriers are present in an OFDM symbol or also in case of dynamic and discontiguous spectrum allocation required in cognitive radios. In order to apply the QU distribution approach for this scenario as well, the formula may be adapted to

$$\mathcal{I}_r = \{i_m\} \qquad m = \left[\frac{l(N - N_z)}{N_r} \right]_R \qquad l = 0, \ldots, N_r - 1, \tag{3.40}$$

whereas i_m addresses the mth element of the ordered set $\mathcal{I}_{nz,o}$ defined as

$$\mathcal{I}_{nz,o} = \{i_0, i_1, i_2 \ldots i_{N-N_z-1}\} = (\mathcal{I}_{nz}, <), \tag{3.41}$$

with $\mathcal{I}_{nz} = \mathcal{I}_N \backslash \mathcal{I}_z$ representing the set of all possible indices $\mathcal{I}_N = \{0, \ldots, N-1\}$ excluding the positions of zero subcarriers $\mathcal{I}_z$. Hence, the redundant subcarriers are distributed as uniformly as possible, but only over the remaining bandwidth. W.l.o.g. it is assumed in (3.40) that zero subcarriers are the only source for exclusion. The second part of Tab. 3.1 reveals, however, that the QU approach does not work well, since $E_r >> E_d$ as soon as zero subcarriers are present.

Table 3.1.: Mean redundant energy E_r of redundant subcarrier distributions obtained from the heuristic and QU approach for the setups in Tab. 2.2.

	Setup A	Setup B	Setup C	Setup D	Setup E
E_dN	36	48	64	100	112
N_z	12	0	0	12	0
E_rN QU distribution	1000	48	64	198.24	112
E_rN Heuristic	36.56	48	64	98.55	112

3.3.3. Heuristic Distribution

The heuristic algorithm presented in the following is able to provide very good results even for these cases, as the fourth row of Tab. 3.1 indicates. This algorithm provides full flexibility and works independently of how many or which subcarriers may be excluded from the optimization process. Furthermore, it allows using any arbitrary cost function J in the optimization routine, e.g., $J = J_{\mathrm{E}}$ as defined in (3.2). For reasons of easier notation, index sets and index vectors instead of permutation matrices are used in the following. The index sets of the redundant and the data subcarriers are represented by $\mathcal{I}_r$ and $\mathcal{I}_d$. Further, $\mathbf{i}_r$ and $\mathbf{i}_d$ denote the corresponding (in ascending order) sorted index vectors. Note that the permutation matrix $\mathbf{P}$ can unambiguously be derived from the sorted index vectors $\mathbf{i}_r$ and $\mathbf{i}_d$. Below the script `heuristic_optimization` and the function `optimize_index_vectors` are reproduced as pseudocode.

Algorithm 1 `heuristic_optimization`

choose valid index vectors $\mathbf{i}_r$ and $\mathbf{i}_d$ randomly
$J_{\mathrm{old}} \leftarrow \infty$
stop $\leftarrow$ false
while not stop **do**
    $(\mathbf{i}_r, \mathbf{i}_d, J_{\mathrm{new}}) \leftarrow$ OPTIMIZE_INDEX_VECTORS$(\mathbf{i}_r, \mathbf{i}_d)$
    if $J_{\mathrm{new}} < J_{\mathrm{old}}$ **then**
        $J_{\mathrm{old}} \leftarrow J_{\mathrm{new}}$
    else
        stop $\leftarrow$ true
    end if
end while
$J_{\mathrm{opt}} \leftarrow J_{\mathrm{old}}$
sort $\mathbf{i}_r$ and $\mathbf{i}_d$ and determine $\mathbf{P}$

`heuristic_optimization` starts with randomly chosen but valid index vectors $\mathbf{i}_r$ and $\mathbf{i}_d$. Next, the function `optimize_index_vectors`, which tries to exchange one

Algorithm 2 `optimize_index_vectors`

```
1: function OPTIMIZE_INDEX_VECTORS(i_r, i_d)
2:     calculate T and G (cf. (2.44), (2.47) and (2.42)) using i_r and i_d
3:     calculate cost function J
4:     J_new ← J
5:     i_r,new ← i_r
6:     i_d,new ← i_d
7:     for k = 0, 1, ..., N_r − 1 do
8:         for l = 0, 1, ..., N_d − 1 do
9:             i_r,tmp ← i_r
10:            i_d,tmp ← i_d
11:            tmp ← i_r,tmp[k]
12:            i_r,tmp[k] ← i_d,tmp[l]
13:            i_d,tmp[l] ← tmp
14:            update T, G using i_r,tmp and i_d,tmp
15:            update J
16:            if J < J_new then
17:                i_r,new ← i_r,tmp
18:                i_d,new ← i_d,tmp
19:                J_new ← J
20:            end if
21:        end for
22:    end for
23:    return (i_r,new, i_d,new, J_new)
24: end function
```

element (index) of $\mathbf{i}_r$ with one element (index) of $\mathbf{i}_d$ such that the cost function decreases by a maximum amount, is repeatedly called until at least a local minimum is found. There is no guarantee to find the global minimum, however, investigations suggest that either the optimum or a near-optimum set is found after executing `heuristic_optimization` a few times. This conclusion is drawn based on two observations: First of all, for the investigated cases a verification by means of exhaustive search due to manageable complexity is possible, it turns out that the results obtained coincide with the global optimum. This has been proven for various setups, e.g., also for setup 3–6 in Fig. 3.7 and Fig. 3.8. Second, in case an exhaustive search is unfeasible, it holds for the found subcarrier distribution that $E_r \approx E_d$. In combination with the results gained in Sec. 3.1, the assumption of having found at least a near to optimum seems valid.

In contrast to an analytical approach like the QU distribution, the deployment of the proposed heuristic may be limited by the parameter set, as `optimize_index_vectors` experiences a quadratic complexity of $N_r \cdot N_d$. Fortunately, finding the optimum distribution is a task which has to be done only once during system

design time and not during run time, hence making this approach applicable for many cases.

3.4. Alternative Optimization Criteria

The experiments summarized in Tab. 2.1 suggest to optimize the permutation matrix $\mathbf{P}$ based on minimizing the cost function $J_{\mathrm{E}}(\mathbf{P})$ in (3.2), i.e., minimizing the redundant energy E_r in (3.31). The BER simulation results in previous publications like [59, 60, 41] as well as in Sec. 3.5 confirm this decision. However, note that the cost function $J_{\mathrm{E}}(\mathbf{P})$ only takes the (mean) energy of the transmit symbol into account and the particular receiver does influence the optimization. On the one hand, this seems preferable, since $\mathbf{P}$ is optimized regardless of the specific receiver structure, e.g., independent of the utilized estimator, and thus applicable for a broad range of scenarios. On the other hand, the question arises, whether this choice is effectively optimum in terms of the overall transceiver performance. The idea is now to find a permutation matrix $\mathbf{P}$ such that the sum of the error variances after a BLUE or an LMMSE estimator at a fixed SNR is minimized instead of the mean energy [58]. These variances can be derived from the main diagonal of the error covariance matrices $\mathbf{C}_{ee}$ as defined in (2.76) and (2.84), respectively. Starting with the BLUE, the sum of the error variances follows to

$$J'(\mathbf{P}) = N\sigma_n^2 \mathrm{tr}\left((\mathbf{G}^H\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}\mathbf{G})^{-1}\right), \tag{3.42}$$

with the generator matrix $\mathbf{G} = \mathbf{P}\begin{bmatrix}\mathbf{I} & \mathbf{T}^T\end{bmatrix}^T$ defined in (3.3). $J'(\mathbf{P})$ depends on the particular channel instance $\tilde{\mathbf{H}}$, meaning that a resulting permutation matrix $\mathbf{P}$ may be perfectly suitable only for one specific transmission channel, but may fail in other situations. This consequently requires the design of several permutation matrices and a channel specific selection algorithm, which in turn demands a certain amount of channel state information at the transmitter side. Since a permutation matrix considered in this work shall be designed only once during system design, the dependence of a cost function on the particular channel $\tilde{\mathbf{H}}$ is disadvantageous. Let us therefore assume the AWGN case with $\tilde{\mathbf{H}} = \mathbf{I}$ yielding

$$J'(\mathbf{P}) = N\sigma_n^2 \mathrm{tr}\left((\mathbf{G}^H\mathbf{G})^{-1}\right). \tag{3.43}$$

With (3.3) and

$$\mathbf{G}^H\mathbf{G} = \begin{bmatrix}\mathbf{I} & \mathbf{T}^H\end{bmatrix}\mathbf{P}^T\mathbf{P}\begin{bmatrix}\mathbf{I}\\ \mathbf{T}\end{bmatrix} = \mathbf{I} + \mathbf{T}^H\mathbf{T}, \tag{3.44}$$

the expression for J' transforms to

$$J'(\mathbf{P}) = N\sigma_n^2 \mathrm{tr}\left((\mathbf{T}^H\mathbf{T} + \mathbf{I})^{-1}\right). \tag{3.45}$$

Introducing and fixing the ratio $\gamma = \frac{E_s}{\sigma_n^2}$ during the optimization allows a performance comparison of different code generator matrices $\mathbf{G}$ at a fixed SNR value. From (3.31) the mean energy per data symbol $E_s = \frac{E_{\mathbf{x}'}}{N_d}$ as in (3.14) immediately follows to $E_s = \left(\sigma_d^2 N_d + \sigma_d^2 \mathrm{tr}\left(\mathbf{T}^H\mathbf{T}\right)\right)/(NN_d)$ for the case of a zero UW, leading to the following expression for σ_n^2:

$$\sigma_n^2 = \frac{E_s}{\gamma} = \frac{\sigma_d^2 N_d + \sigma_d^2 \mathrm{tr}\left(\mathbf{T}^H\mathbf{T}\right)}{\gamma N N_d}. \tag{3.46}$$

Inserting (3.46) into (3.45) finally delivers the cost function

$$J_{\mathrm{BLUE}}(\mathbf{P}) = \frac{\sigma_d^2}{\gamma N_d}\left(\mathrm{tr}\left(\mathbf{T}^H\mathbf{T}\right) + N_d\right)\mathrm{tr}\left(\left(\mathbf{T}^H\mathbf{T} + \mathbf{I}\right)^{-1}\right). \tag{3.47}$$

Obviously, the particular value of $\gamma = \frac{E_s}{\sigma_n^2}$ does not influence the optimization process and the resulting optimum permutation matrix $\mathbf{P}$.

Applying the same steps, the LMMSE estimator based cost function follows to

$$J_{\mathrm{LMMSE}}(\mathbf{P}) = \frac{\sigma_d^2}{\gamma N_d}\left(\mathrm{tr}\left(\mathbf{T}^H\mathbf{T}\right) + N_d\right)\mathrm{tr}\left(\left(\mathbf{T}^H\mathbf{T} + \frac{(\gamma+1)N_d + \mathrm{tr}\left(\mathbf{T}^H\mathbf{T}\right)}{\gamma N_d}\mathbf{I}\right)^{-1}\right). \tag{3.48}$$

Alternatively, the cost function can also be written as

$$J_{\mathrm{LMMSE}}(\mathbf{P}) = \sigma_d^2 \mathrm{tr}\left(\left(\mathbf{I} + \frac{\gamma N_d}{N_d + \mathrm{tr}\left(\mathbf{T}^H\mathbf{T}\right)}(\mathbf{T}^H\mathbf{T} + \mathbf{I})\right)^{-1}\right). \tag{3.49}$$

For sufficiently large γ it holds that $J_{\mathrm{LMMSE}}(\mathbf{P}) \approx J_{\mathrm{BLUE}}(\mathbf{P})$, and the particular choice of γ is again irrelevant for the searching of an optimum permutation matrix. However, this is not immediately apparent for small values of γ.

In the following paragraphs, optimum permutation matrices $\mathbf{P}$, or equivalently ascendingly ordered optimum index sets $\mathcal{I}_{r,o} = (\mathcal{I}_r, <)$, are presented for the UW-OFDM setups in Tab. 2.2 by minimizing the cost functions $J_{\mathrm{E}}(\mathbf{P})$ from (3.2), $J_{\mathrm{BLUE}}(\mathbf{P})$ and $J_{\mathrm{LMMSE}}(\mathbf{P})$. Tab. 3.2 shows the optimum index sets of the redundant subcarriers obtained with the heuristic optimization algorithm (Sec. 3.3) for setup A-C. As noted before, the choice of $\gamma = \frac{E_s}{\sigma_n^2}$ is irrelevant for the minimization of J_{BLUE}. In contrast, the solution of J_{LMMSE} might depend on γ and is thus solved for $\gamma = 1, 2, \ldots, 40$.

For setup A the solutions to the three cost functions yield exactly the same optimum index set $\mathcal{I}_r$, and in case of J_{LMMSE} additionally independent of γ.

For setup B and C again all three optimization criteria deliver the very same optimum index set(s), for $J_{\mathrm{LMMSE}}(\mathbf{P})$ again independent of the specific value of γ. However, as the redundant subcarriers are now equidistantly distributed among

Table 3.2.: Best index sets for redundant subcarriers when optimizing towards different cost functions for setups according to Tab. 2.2, $N = 64$ and $N = 80$.

	γ	Setup A – Best index set $\mathcal{I}_r$
J_{E},	-	{2,6,10,14,17,21,24,26,38,40,43,47,50,54,58,62}
J_{BLUE},	-	-"-
J_{LMMSE}	1,...,40	-"-

	γ	Setup B – Best index set $\mathcal{I}_r$
J_{E},	-	{0,4,8,12,16,20,24,28,32,36,40,44,48,52,56,60}
J_{BLUE},	-	{1,5,9,13,17,21,25,29,33,37,41,45,49,53,57,61}
J_{LMMSE}	1,...,40	{2,6,10,14,18,22,26,30,34,38,42,46,50,54,58,62}
		{3,7,11,15,19,23,27,31,35,39,43,47,51,55,59,63}

	γ	Setup C – Best index set $\mathcal{I}_r$
J_{E},	-	{0,5,10,15,20,25,30,35,40,45,50,55,60,65,70,75}
J_{BLUE},	-	{1,6,11,16,21,26,31,36,41,46,51,56,61,66,71,76}
J_{LMMSE}	1,...,40	{2,7,12,16,22,27,32,37,42,47,52,57,62,67,72,77}
		{3,8,13,18,23,28,33,38,43,48,53,58,63,68,73,78}

the available frequency band, an optimum set is not unique anymore, but every set originating from a cyclic shift of an optimum set minimizes the cost functions as well, cf. [67].

Table 3.3.: Best index sets for redundant subcarriers when optimizing towards different cost functions for setups according to Tab. 2.2, $N = 128$.

	γ	Setup D – Best index set $\mathcal{I}_r$
J_{E},	-	{4,12,20,28,36,44,51,58,70,77,84,92,100,108,116,124}
J_{BLUE}	-	-"-
J_{LMMSE}	1	{4,12,20,28,36,44,51,57,71,77,84,92,100,108,116,124}
	6	{5,13,21,29,37,45,52,58,71,78,85,93,101,109,117,125}
	34,...,40	{4,12,20,28,36,44,51,58,70,77,84,92,100,108,116,124}

	γ	Setup E – Best index set $\mathcal{I}_r$
J_{E},	-	{0,8,16,24,32,40,48,56,64,72,80,88,96,104,112,120}
J_{BLUE},	-	{1,9,17,25,33,41,49,57,65,73,81,89,97,105,113,121}
J_{LMMSE}	1,...,40	...

Tab. 3.3 summarizes the optimization results for setup D and setup E, respectively. Setup E does not provide any new insights, all statements made for setup B and C also hold for this case. However, setup D shows some new aspects. Whereas minimizing $J_{\mathrm{E}}(\mathbf{P})$ and $J_{\mathrm{BLUE}}(\mathbf{P})$ results in the same optimum index set $\mathcal{I}_r$, this is not always true for $J_{\mathrm{LMMSE}}(\mathbf{P})$. Here, the optimum set slightly changes depending on the specific value of γ. Only for $\gamma \geq 34$, minimizing $J_{\mathrm{LMMSE}}(\mathbf{P})$ provides the same optimum set as for $J_{\mathrm{E}}(\mathbf{P})$ and $J_{\mathrm{BLUE}}(\mathbf{P})$. At first sight, these results seem now to be in some contrast to the previous outcomes. Let us thus examine the

cost function $J_{\mathrm{LMMSE}}(\mathbf{P})$ for this situation in detail. Tab. 3.4 presents $J_{\mathrm{LMMSE}}(\mathbf{P})$ for $\gamma = 1$ and $\gamma = 6$, and in each case evaluated for the optimum index set $\mathcal{I}_{r,opt}$ as well as for the set $\mathcal{I}_{r,J_{\mathrm{E}},J_{\mathrm{BLUE}}}$ minimizing $J_{\mathrm{E}}(\mathbf{P})$ and $J_{\mathrm{BLUE}}(\mathbf{P})$. Note that the difference in the cost function is basically negligible. Hence, $\mathcal{I}_{r,J_{\mathrm{BLUE}},J_{\mathrm{E}}}$ can practically be seen as the optimum set that minimizes all the three cost functions. In conclusion, in all cases the intuitive choice of finding $\mathcal{I}_r$ by minimizing $J_{\mathrm{E}}(\mathbf{P})$ is in fact also optimum in terms of the new performance measures J_{BLUE} in (3.47) and J_{LMMSE} in (3.49), which take the whole transceiver performance into account.

Table 3.4.: Evaluation of J_{LMMSE} for setup D according to Tab. 2.2.

	$\mathcal{I}_{r,opt}$	$\mathcal{I}_{r,J_{\mathrm{E}},J_{\mathrm{BLUE}}}$	Difference
$J_{\mathrm{LMMSE}}, \gamma = 1$	59.5122	59.5769	0.0647
$J_{\mathrm{LMMSE}}, \gamma = 6$	21.6837	21.6897	0.0060

3.5. Performance Evaluation

The performance evaluation presented in this section is split into two main parts. The first part investigates the principle mode of operation and highlights scenarios in which UW-OFDM can show its strengths. An MSE analysis after data estimation serves as metric. Based on this analysis, the class of estimators suitable for UW-OFDM is identified. The second part of this section evaluates the performance of UW-OFDM against CP-OFDM. The latter is currently the most popular multicarrier technique and therefore constitutes a perfect benchmark. The comparison of both signaling schemes comprises spectral properties, bandwidth efficiency and BER simulations.

3.5.1. Principles of UW-OFDM

As already mentioned, UW-OFDM can in fact be interpreted as an RS code, the difference to the usual case is the definition of the code over the field of the complex numbers instead of over a Galois field. Thus, the RS code property immediately motivates for an algebraic decoding approach to recover the data symbols disturbed by the channel [21]. However, subsequent elaborations will identify problems in terms of the practical applicability of this approach, paving the way for estimators such as BLUE or LMMSE, which will in Sec. 3.5.3 demonstrate the potential of the UW-OFDM signaling concept.

As a preparatory step for the algebraic approach, let us equalize the channel effects first by multiplying with $\tilde{\mathbf{H}}^{-1}$, yielding

$$\tilde{\mathbf{y}}' = \tilde{\mathbf{H}}^{-1}\tilde{\mathbf{y}} \tag{3.50}$$

$$= \mathbf{G}\mathbf{d} + \tilde{\mathbf{H}}^{-1}\mathbf{v} \tag{3.51}$$

$$= \mathbf{c} + \tilde{\mathbf{H}}^{-1}\mathbf{v}. \tag{3.52}$$

Now a similar principle as for the generation of the redundant subcarriers $\mathbf{r}$ at the transmitter side is applied, cf. (3.7), but this time the aim is the recovery of highly attenuated subcarrier symbols $\mathbf{c}_b \in \mathbb{C}^{(m\times 1)}$ from well-received subcarrier symbols $\mathbf{c}_g \in \mathbb{C}^{(N_d+N_u-m)\times 1}$ with $m \leq N_u$ denoting the number of highly attenuated subcarriers. Let $\mathbf{P}' \in \mathbb{C}^{(N_d+N_u)\times(N_d+N_u)}$ be a permutation matrix that splits up $\mathbf{c}$ into good $\mathbf{c}_g$ and bad $\mathbf{c}_b$ subcarrier symbols such that

$$\mathbf{c}' = \begin{bmatrix}\mathbf{c}_g \\ \mathbf{c}_b\end{bmatrix} = \mathbf{P}'\mathbf{c}. \tag{3.53}$$

It holds that

$$\mathbf{F}_N^{-1}\mathbf{B}\mathbf{c} = \mathbf{F}_N^{-1}\mathbf{B}\mathbf{P}'^T\begin{bmatrix}\mathbf{c}_g \\ \mathbf{c}_b\end{bmatrix} = \begin{bmatrix}* \\ \mathbf{0}\end{bmatrix}. \tag{3.54}$$

With

$$\mathbf{M}' = \mathbf{F}_N^{-1}\mathbf{B}\mathbf{P}'^T = \begin{bmatrix}\mathbf{M}'_{11} & \mathbf{M}'_{12} \\ \mathbf{M}'_{21} & \mathbf{M}'_{22}\end{bmatrix}, \tag{3.55}$$

where $\mathbf{M}'_{ij}$ are appropriately sized sub-matrices, it follows that

$$\mathbf{M}'_{21}\mathbf{c}_g + \mathbf{M}'_{22}\mathbf{c}_b = \mathbf{0}. \tag{3.56}$$

Assuming that $\mathbf{c}_g$ is known, then $\mathbf{c}_b$ can be derived by $\mathbf{c}_b = -\mathbf{M}_{22}^{'-1}\mathbf{M}'_{21}\mathbf{c}_g$ for the case $m = N_u$, and by $\mathbf{c}_b = -\mathbf{M}_{22}^{'\dagger}\mathbf{M}'_{21}\mathbf{c}_g$ for $m < N_u$. With $\mathbf{T}' = -\mathbf{M}_{22}^{'\dagger}\mathbf{M}'_{21}$, the code word is given as

$$\mathbf{c} = \mathbf{P}'^T\begin{bmatrix}\mathbf{c}_g \\ \mathbf{c}_b\end{bmatrix} = \mathbf{P}'^T\begin{bmatrix}\mathbf{I} \\ \mathbf{T}'\end{bmatrix}\mathbf{c}_g = \mathbf{G}'\mathbf{c}_g. \tag{3.57}$$

In practice, $\mathbf{c}_g$ is not perfectly known, however, with $\mathbf{P}' = \begin{bmatrix}\mathbf{P}_1'^T & \mathbf{P}_2'^T\end{bmatrix}^T$ and $\mathbf{P}'_1 \in \mathbb{C}^{(N_d+N_u-m)\times(N_d+N_u)}$, an estimate $\hat{\mathbf{c}}_g$ is obtained from (3.52) by $\hat{\mathbf{c}}_g = \mathbf{P}'_1\tilde{\mathbf{y}}'$. Using the estimate $\hat{\mathbf{c}}_b = \mathbf{T}'\hat{\mathbf{c}}_g$ leads to

$$\hat{\mathbf{c}} = \mathbf{G}'\mathbf{P}'_1\tilde{\mathbf{y}}' = \mathbf{G}'\mathbf{P}'_1\tilde{\mathbf{H}}^{-1}\tilde{\mathbf{y}}. \tag{3.58}$$

As a first step, the impact of the channel is inverted. Next, the well-received symbols are extracted by applying $\mathbf{P}'_1$ and finally the heavily distorted symbols are recovered by $\mathbf{G}'$. Since there is a zero UW of length N_u in the time domain, the decoder is able to recover up to N_u erased subcarrier symbols. In case of $m = N_u$, this system of equations for $\hat{\mathbf{c}}_b$ has a uniquely defined solution (i.e.,

the Pseudo-Inverse turns into an inverse). In case of $m < N_u$, this leads to an overdetermined system of equations for $\hat{\mathbf{c}}_b$ which is solved in a least square sense. One can show that the error $\mathbf{e} = \mathbf{c} - \hat{\mathbf{c}}$ has zero mean, and its covariance matrix is given by

$$\mathbf{C}_{ee} = \sigma_d^2 \left(\mathbf{C}_{cc} - \mathbf{G}'\mathbf{P}_1'\mathbf{C}_{cc} - \mathbf{C}_{cc}(\mathbf{G}'\mathbf{P}_1')^H + \mathbf{G}'\mathbf{P}_1'\mathbf{C}_{cc}(\mathbf{G}'\mathbf{P}_1')^H \right) + \mathbf{G}'\mathbf{P}_1'\mathbf{C}_{vv}(\mathbf{P}_1')^H. \tag{3.59}$$

$\mathbf{C}_{ee}$ can further be used in the case when additional channel coding is applied. Finally, the data part

$$\hat{\mathbf{d}} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \end{bmatrix} \mathbf{P}^T \hat{\mathbf{c}} \tag{3.60}$$

is processed as usual.

In the successive example, all non-zero subcarrier symbols are taken into account and considerations are thus based on (3.58). Two exemplary multipath snapshots shown in Fig. 3.12 will help to reveal the potential of this decoding method. Channel A represents a pleasant environment with moderate fading holes and is close to a frequency flat channel. Channel B is a quite bad environment for communication due to its two deep spectral notches.

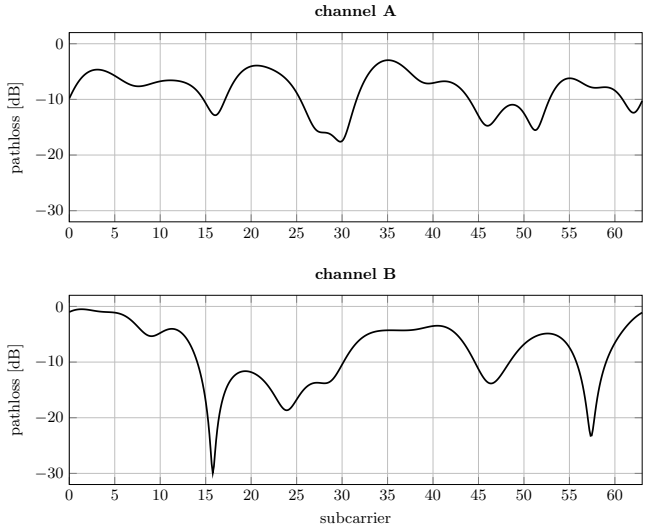


Figure 3.12.: Frequency response of exemplary multipath channels A and B.

Starting considerations based on channel B, it makes sense to assume that all subcarrier symbols have been well-received except for those at the subcarriers with index no. 16 and 58. Hence, these two subcarrier symbols will be recalculated from

the well-received ones applying the estimator in (3.58). Note that the approach of dividing into good and bad received symbols shows strong parallels to successive interference cancelation (SIC)/decision feedback equalization (DFE) [69, 70]. Fig. 3.13 shows the Bayesian MSE (BMSE) before and after algebraic RS decoding, and with and without slicing. In this context, 'Bayesian' denotes an averaging not only over noise, but over data as well. Slicing means that a received and noisy data symbol is mapped to the closest (in terms of Euclidean distance) element of the underlying data symbol alphabet. For this analysis, perfect slicing is assumed, i.e., the noisy data symbols within $\hat{\mathbf{c}}_g = \mathbf{P}'_1\tilde{\mathbf{y}}'$ are replaced by the actually transmitted data symbols. For the redundant subcarriers within $\hat{\mathbf{c}}_g = \mathbf{P}'_1\tilde{\mathbf{y}}'$, slicing is assumed to be practically unfeasible due to the vast amount of elements of the underlying symbol alphabet (Fig. 3.9), hence the noisy symbols after the CI stage are used. In case no slicing is applied, the values gained after the CI stage ($\hat{\mathbf{c}}_g = \mathbf{P}'_1\tilde{\mathbf{y}}'$) are used for all subcarrier symbols. The algebraic RS decoding approach is able to slightly improve the BMSE on subcarrier no. 16. Interestingly, with perfect slicing the BMSE cannot be reduced significantly anymore. As such, the noise on the redundant subcarriers dominates the performance of the algebraic RS decoding receiver. Moreover, the decoder will even increase the BMSE in case of subcarrier no. 58.

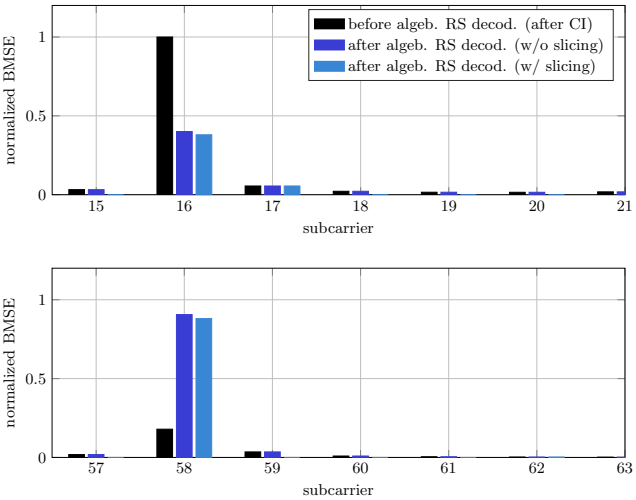


Figure 3.13.: Noise reduction/enhancement effect of the algebraic RS decoder for channel B at $E_b/N_0 = 20\,\mathrm{dB}$. Above: zoomed ordinate around carrier no. 16; below: zoomed ordinate around carrier no. 58.

In summary, algebraic RS decoding leads to solving a very ill-conditioned system of equations which cannot be solved adequately anymore, as soon as only little noise ($E_b/N_0 = 20\,\mathrm{dB}$ for this simulation) is present. Higher noise due to moderate E_b/N_0 values can sometimes stabilize the system to a certain extent such that the BMSE will at least not increase, however, there is still no countable gain in terms of minimization. Unfortunately, also applying common regularization techniques like e.g., the Tikhonov regularization [71] does not improve the performance. Furthermore, channel B has been identified to be the most pleasant scenario for algebraic decoding among many simulated channel instances, in case of channel A for instance, the resulting BMSE after RS decoding will always skyrocket, independently of the present noise. Hence, this receiver concept is not applicable for practical UW-OFDM systems.

However, the next paragraphs will demonstrate that receivers based on classical and Bayesian estimation theory such as BLUE and LMMSE, respectively, work well in contrast to the presented concept based on algebraic decoding. Following the definition in (2.81), the LMMSE estimator can be interpreted as the concatenation of a channel inversion and a smoothing stage, leading to

$$\mathbf{E}_{\mathrm{LMMSE}} = \mathbf{W}\tilde{\mathbf{H}}^{-1}. \tag{3.61}$$

$\tilde{\mathbf{H}}^{-1}$ reverses the impact of the transmission channel, while the smoothing matrix $\mathbf{W}$ exploits the characteristic correlations of an UW-OFDM signal. This supersedes conventional CP-OFDM, where no correlations are present and thus the inversion of the channel influence already represents the optimum linear receiver. Hence, comparing input and output of the smoothing stage allows to quantify the additional potential that is provided by the (channel-independent) precoding (in the sense of a linear dispersive preprocessing [52]) and channel coding properties introduced in UW-OFDM. Furthermore, this partitioning into two stages allows an easy comparison with the algebraic decoding approach. Applying $\mathbf{E}_{\mathrm{LMMSE}}$ on $\tilde{\mathbf{y}} = \tilde{\mathbf{H}}\mathbf{G}\mathbf{d} + \mathbf{v}$ results in

$$\hat{\mathbf{d}} = \mathbf{E}_{\mathrm{LMMSE}}\tilde{\mathbf{y}}. \tag{3.62}$$

In order to visualize the effect of the smoothing operation on all non-zero subcarriers, i.e., on $\mathbf{d}$ as well as on $\mathbf{r}$, let

$$\hat{\mathbf{c}} = \mathbf{G}\hat{\mathbf{d}}. \tag{3.63}$$

Fig. 3.14 displays the BMSE before and after the smoothing stage when transmitting at $E_b/N_0 = 20\,\mathrm{dB}$ over channel A (Fig. 3.12). As expected from LMMSE theory [42], smoothing improves the BMSE on every subcarrier. Furthermore, it seems that the level of improvement, in relative as well as in absolute terms, rises with increasing BMSE at the output of the channel inversion stage. *Relative* in this case refers to the relative improvement w.r.t. the BMSE before smoothing.

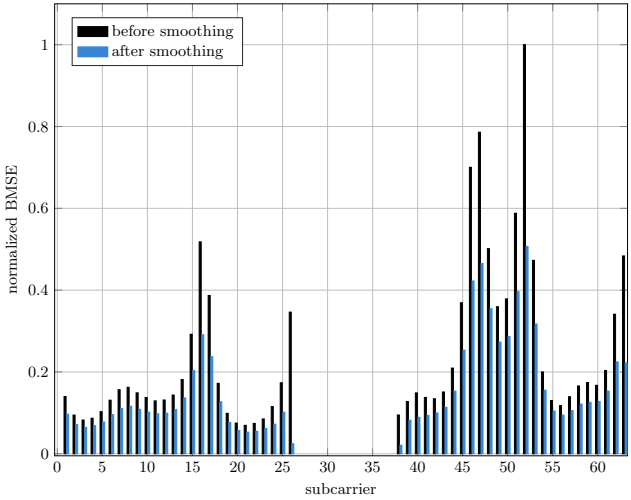


Figure 3.14.: Noise reduction effect of the Wiener smoother for exemplary multipath channel A at $E_b/N_0 = 20\,\mathrm{dB}$.

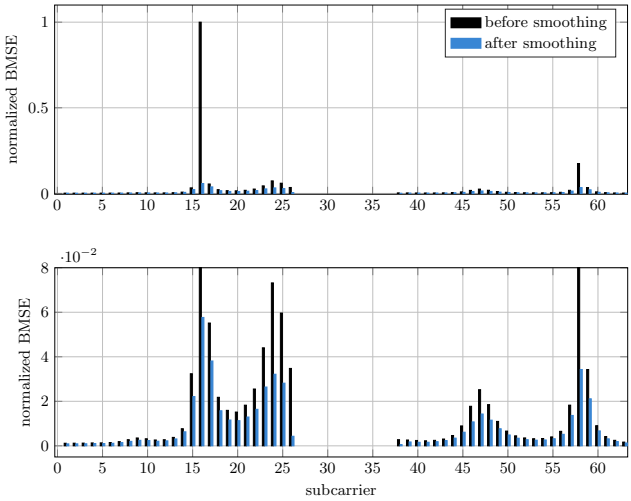


Figure 3.15.: Noise reduction effect of the Wiener smoother for exemplary multipath channel B at $E_b/N_0 = 20\,\mathrm{dB}$. Upper plot: full scale; lower plot: zoomed ordinate.

In order to investigate this assumption more in detail, Fig. 3.15 displays the BMSE in case of transmission channel B. Here, the lower plot represents a zoomed version of the upper plot. Again the performance is enhanced on every subcarrier after smoothing, but the main message can be derived from subcarrier no. 16. Considering channel B, this subcarrier has obviously experienced substantial attenuation by a deep fading hole, consequently leading to a significant noise enhancement when applying channel inversion. Nevertheless, the inherent correlations of an UW-OFDM signal can be exploited in a very powerful way by applying $\mathbf{W}$, such that the effects of spectral notches are compensated to a large extent. Note that the *before smoothing* in Fig. 3.15 exactly corresponds to the *before algebraic RS decoding* in Fig. 3.13, a comparison thus emphasizes the potential of LMMSE estimation. The high E_b/N_0 value has been chosen for comparison reasons with algebraic decoding. Nevertheless, the LMMSE estimator performs significantly better in every E_b/N_0 region.

In conclusion, the presented algebraic decoding approach is not a serious option for UW-OFDM receivers. However, the results for an LMMSE receiver in Fig. 3.14 and 3.15 already indicate the performance potential of UW-OFDM over CP-OFDM[3]. This conclusion follows from the already stated fact that in CP-OFDM a channel inverter is the best linear estimator possible. It turns out that the performance enhancement in UW-OFDM generally increases with the frequency selectivity of the transmission channel. In this context, an increase denotes an increase of the number and depth of fading holes. In order to support this statement, Fig. 3.16 illustrates BER curves of setup A for uncoded transmission and QPSK as modulation alphabet, whereas one time $\mathbf{E}_{\mathrm{CI}}$ and the other time $\mathbf{E}_{\mathrm{LMMSE}}$ is applied as estimator. The simulations are conducted for AWGN, channel A and channel B, and confirm an increasing gap from 1.5 dB over 2.7 dB to 11.6 dB, all measured at a BER of 10^{-6}.

3.5.2. Power Spectral Density

Fig. 3.17 illustrates the estimated power spectral densities (PSDs) of a simulated UW-OFDM and a CP-OFDM burst in case of setup A. For that, a simulated burst is composed of a preamble (in all cases the IEEE 802.11a preamble [36]), and a data part comprising 8000 information bits, which are processed by an outer channel code with coding rate $r = 1/2$. In this case no additional filters for spectral shaping have been applied. For a better comparison, the PSDs have been normalized such that the passband of each spectral mask is centered at 0 dB. The UW-OFDM spectrum features a significantly better sidelobe suppression compared to the CP-OFDM[4] spectrum. The out-of-band emissions of UW-OFDM are more

[3] Note that in-depth BER comparisons with CP-OFDM follow in Sec. 3.5.3.

[4] The displayed CP-OFDM spectrum shows four additional spectral holes besides the one at DC, which result from setting the pilot symbols to zero.

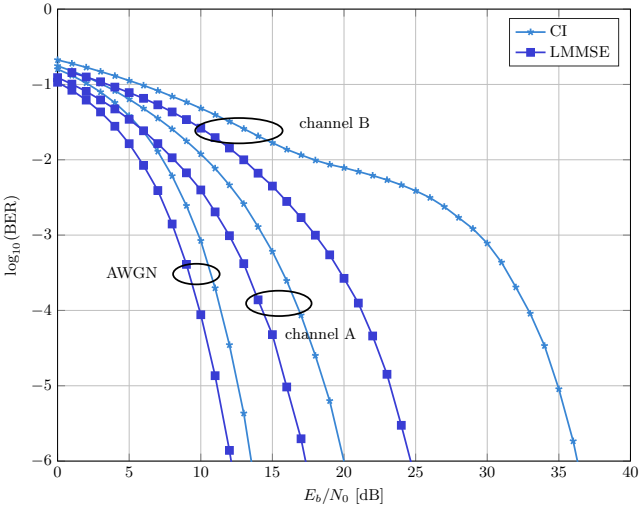


Figure 3.16.: BER comparison for UW-OFDM between LMMSE and CI estimator in case of uncoded transmission in AWGN, multipath channel A and B.

than 15 dB below the emissions of the CP-OFDM system. One reason for this property is of course the decreasing mean power level of subcarriers close to zero bands, cf. Fig. 3.4, however, the main cause originates from the correlations introduced by the redundant subcarriers. This can easily be verified by a simple experiment, namely replacing the redundant subcarrier symbols by data symbols, generating therefore conventional OFDM symbols and weighting the subcarriers according to the mean power profile in Fig. 3.4. Consequently, the effect of decreasing mean power values is covered while the correlations are excluded. The resulting spectrum of these OFDM symbols labeled as 'Weighted OFDM' in Fig. 3.17 misses the superior suppression of out-of-band-emissions, revealing the correlations as responsible source. In [72] the authors additionally investigate the impact of the specific redundant subcarrier distribution on the spectral properties.

The PSDs also show that the UW-OFDM system requires about 0.4 MHz less bandwidth when measured at -10 dB. In combination with the shorter symbol duration T_{OFDM}, the UW-OFDM system features almost the same bandwidth efficiency η as the CP-OFDM system, despite less data subcarriers per symbol are occupied. Tab. 3.5 displays the bandwidth efficiency $\eta = (N_d/T_{\mathrm{OFDM}})/B$ measured in megasymbols per seconds per Hertz for setup A. Additionally, η_{cp} determines the bandwidth efficiency of the UW-OFDM system related to the CP-OFDM reference system.

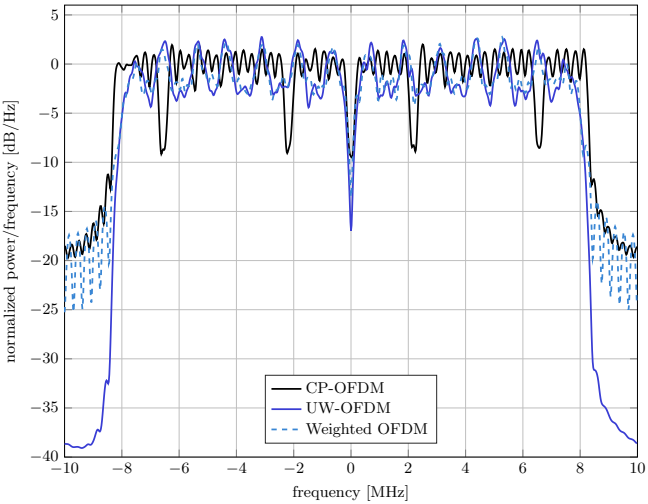


Figure 3.17.: Comparison of normalized Welch power spectral density between CP-OFDM and UW-OFDM for setup A.

	CP-OFDM	UW-OFDM
N_d	48	36
N_r (N_p)	4	16
B [MHz]	16.70	16.30
η [Msymb/s/Hz]	0.719	0.690
η_{cp}	1.000	0.960
E_b	0.00977	0.01575
AWGN loss [dB]	0.0000	2.0754

Table 3.5.: Setup comparison between CP-OFDM and UW-OFDM for setup A.

3.5.3. Bit Error Ratio Simulations

Obviously, OFDM has been designed for data transmission in a frequency selective environment. Nevertheless, simulation results in the AWGN channel serve as a starting point, since these results provide first interesting insights. Fig. 3.18 shows that CP-OFDM outperforms UW-OFDM with simple channel inversion (which in this case corresponds to the trivial estimator $\mathbf{E} = \mathbf{I}$) by 2.1 dB for uncoded transmission, i.e., without additional channel code. This performance gap exactly originates from the energy excess required in UW-OFDM to load the redundant subcarriers. However, the introduced redundancy will only pay off, if the thereby enabled coding gain in the receiver exceeds the excess energy of 2.1 dB. This AWGN loss can also be derived analytically by calculating $10\log_{10}\left(\frac{E_{b,\mathrm{uw}}}{E_{b,\mathrm{cp}}}\right)$, whereas $E_{b,\mathrm{cp}}$ determines the mean energy per data bit for the CP-OFDM system and $E_{b,\mathrm{uw}}$ that of the UW-OFDM system. Tab. 3.5 shows the corresponding energy values when evaluated for setup A. With LMMSE estimation, the performance of UW-OFDM improves substantially, however, it still lacks 0.6 dB to keep up with CP-OFDM. Hence, in case of AWGN and linear estimators, systematically encoded UW-OFDM is not the preferable choice from a BER point of view.

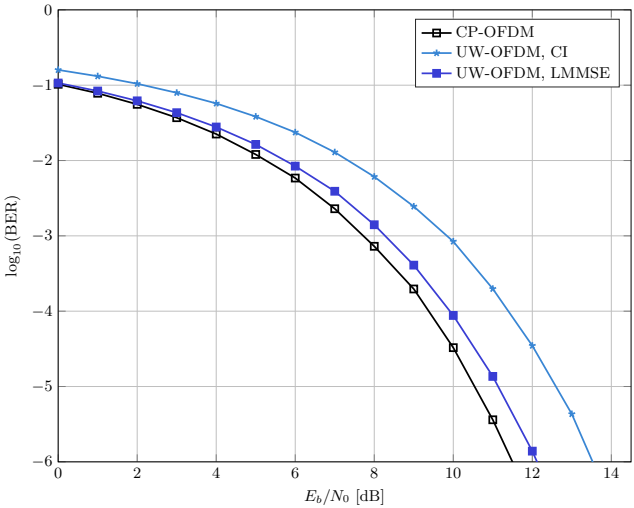


Figure 3.18.: BER comparison of UW-OFDM with CI or LMMSE estimator and CP-OFDM for uncoded transmission in AWGN.

For frequency selective environments, the channel model according to [46] applies which has also been used during the IEEE 802.11a standardization process. The

channel impulse responses are modeled as tapped delay lines, each tap as a circularly symmetric complex Gaussian variable with zero mean and variances decaying exponentially along the taps, and the number of taps depending on the channel delay spread τ_{RMS}. A detailed description of the underlying channel model can be found in 2.4.2. The simulations have been averaged over a fixed set of 10^4 different channel instances (see simulation settings in Sec. 2.4.3) featuring a delay spread of $\tau_{\mathrm{RMS}} = 100\,\mathrm{ns}$ and a total length not exceeding the guard interval. The channel is assumed to be quasi-static, meaning that it is constant for the transmission of one burst and then changes to an independent new instance.

Fig. 3.19 plots the BER performance in case of a frequency selective environment and uncoded transmission. In this scenario, UW-OFDM reveals its strengths over conventional CP-OFDM, the inherent pre- [52] and channel coding translates to a diversity gain, i.e., an increase of the slope of the BER curve, and thus to a performance gain of 20.5 dB. In the coded case as illustrated in Fig. 3.20, the gain

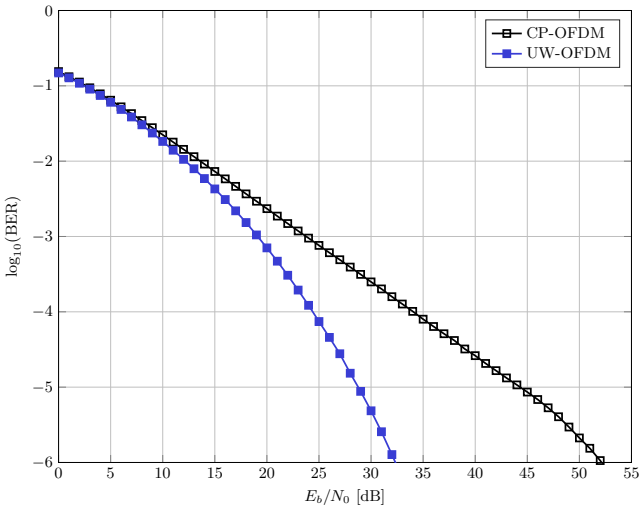


Figure 3.19.: BER comparison of UW-OFDM and CP-OFDM for uncoded transmission in a multipath environment with $\tau_{\mathrm{RMS}} = 100\,\mathrm{ns}$.

reduces but still reaches 1.1 dB for a coding rate of $r = 3/4$ and 0.5 dB for $r = 1/2$.

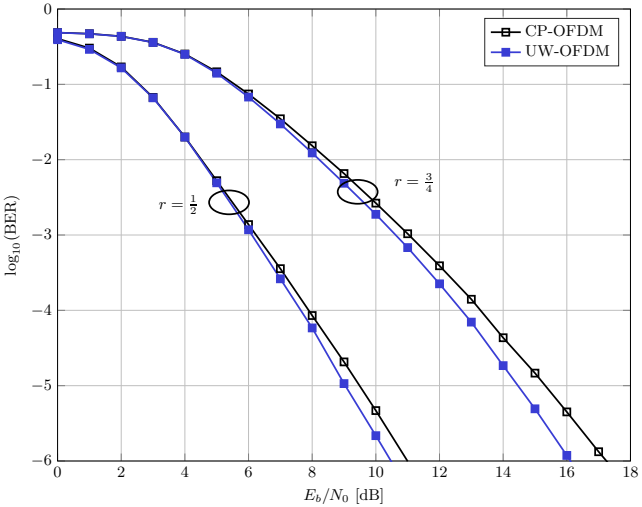


Figure 3.20.: BER comparison of UW-OFDM and CP-OFDM for coded transmission ($r = 1/2$ and $r = 3/4$) in a multipath environment with $\tau_{\mathrm{RMS}} = 100\,\mathrm{ns}$.

3.6. Introduction of Additional Redundant Subcarriers

The optimization of $\mathbf{P}$ in Sec. 3.3 has already reduced the redundant energy to an acceptable level. The BER simulation results in the previous section – especially for the frequency selective scenarios – confirm the potential of UW-OFDM utilizing an optimized $\mathbf{P}$. It cannot be stressed enough that these results would not be possible without minimizing the redundant energy. Thus, this minimization determines the key to a practical UW-OFDM system in the first place. The question arises, if there is still some potential left for minimization.

Up till now, only the case $N_r = N_u$ has been considered. For this particular situation, the problem of calculating and loading the redundant subcarriers translates to solving a linear system of equations with N_u equations for N_u unknowns and is thus uniquely solvable. In case of $N_r > N_u$, the problem to be solved translates to an underdetermined linear system of equations and thus provides infinitely many solutions [61]. In order to elaborate on this in detail, let us recap the general formulation of the optimization problem to find an appropriate generator matrix for systematically encoded UW-OFDM

$$\check{\mathbf{G}} = \underset{\mathbf{G}}{\operatorname{argmin}} \{J_{\mathrm{E}}\} \quad \text{s.t. } \mathbf{F}_N^{-1}\mathbf{B}\mathbf{G} = \begin{bmatrix} \boldsymbol{\Xi} \\ \mathbf{0} \end{bmatrix} \wedge \mathbf{G} = \mathbf{P}\begin{bmatrix} \mathbf{I} \\ \mathbf{T} \end{bmatrix}. \tag{3.64}$$

In detail, the generator matrix $\mathbf{G}$ should generate a zero UW in the time domain, be composed of an identity matrix $\mathbf{I}$, a permutation matrix $\mathbf{P}$ and a matrix $\mathbf{T}$ that introduces a redundancy $\mathbf{r}$, and should further minimize the mean redundant energy represented by the cost function J_{E} defined in (3.2). Taking all this into account and reformulating the zero word constraint for the sake of easier understanding afterwards, the optimization problem translates to

$$\check{\mathbf{P}}, \check{\mathbf{T}} = \underset{\mathbf{P},\mathbf{T}}{\operatorname{argmin}} \left\{ \operatorname{tr}\left(\mathbf{T}\mathbf{T}^H\right) \right\} \quad \text{s.t. } \mathbf{F}_N^{-1}\mathbf{B}\mathbf{P}\begin{bmatrix}\mathbf{d}\\ \mathbf{r}\end{bmatrix} = \begin{bmatrix}\boldsymbol{\Xi}\\ \mathbf{0}\end{bmatrix} \wedge \mathbf{G} = \mathbf{P}\begin{bmatrix}\mathbf{I}\\ \mathbf{T}\end{bmatrix}, \tag{3.65}$$

thus identifying $\mathbf{P}$ and $\mathbf{T}$ as optimization parameters available for minimizing the cost function. With the introduction of $\mathbf{M} = \mathbf{F}_N^{-1}\mathbf{B}\mathbf{P} = \left[\begin{smallmatrix}\mathbf{M}_{11} & \mathbf{M}_{12}\\ \mathbf{M}_{21} & \mathbf{M}_{22}\end{smallmatrix}\right]$ as known from (2.45), the problem becomes

$$\check{\mathbf{P}}, \check{\mathbf{T}} = \underset{\mathbf{P},\mathbf{T}}{\operatorname{argmin}} \left\{ \operatorname{tr}\left(\mathbf{T}\mathbf{T}^H\right) \right\} \quad \text{s.t. } \mathbf{M}_{21}\mathbf{d} + \mathbf{M}_{22}\mathbf{r} = \mathbf{0} \wedge \mathbf{G} = \mathbf{P}\begin{bmatrix}\mathbf{I}\\ \mathbf{T}\end{bmatrix}. \tag{3.66}$$

At this point, a distinction between both cases is useful. In case of $N_r = N_u$, matrix $\mathbf{M}_{22}$ of dimension $N_u \times N_r$ becomes quadratic and the constraint is fulfilled by choosing $\mathbf{r} = \mathbf{T}\mathbf{d}$ with the unambiguously[5] determined matrix $\mathbf{T} = -\mathbf{M}_{22}^{-1}\mathbf{M}_{21}$. Consequently, the permutation matrix $\mathbf{P}$ is the only degree of freedom to minimize the cost function. This is the status already known so far.

In case of $N_r > N_u$, however, the construction of $\mathbf{T}$ is ambiguous, leaving infinitely many solutions that fulfill the zero word constraint, thus raising the question how to optimally choose $\mathbf{T}$. Let us remember in this context that a proper $\mathbf{G}$ aims at minimizing the mean energy of the redundant subcarrier symbols $\mathbf{r}$. Second, let us also remember that for underdetermined linear systems of the form $\mathbf{C}\mathbf{x} = \mathbf{b}$, the Moore-Penrose Pseudo-Inverse delivers the minimum Euclidean norm $\|\mathbf{x}\|_2 = \sqrt{\mathbf{x}^H\mathbf{x}}$ among all solutions. In the context of UW-OFDM with $\mathbf{C} = \mathbf{M}_{22}$, $\mathbf{x} = \mathbf{r} = \mathbf{T}\mathbf{d}$ and $\mathbf{b} = -\mathbf{M}_{21}\mathbf{d}$, the Moore-Penrose Pseudo-Inverse delivers the minimum redundant energy for a given but arbitrary data vector instance $\mathbf{d}$. As a consequence, if it provides the resulting minimum redundant energy for each single data vector, it automatically provides the minimum energy when averaged over all data vectors, thus also minimizing the mean redundant energy. Hence, assuming a given permutation matrix $\mathbf{P}$, (3.66) is met with optimality by the Moore-Penrose Pseudo-Inverse defined as $\mathbf{M}_{22}^{\dagger} = \mathbf{M}_{22}^H\left(\mathbf{M}_{22}\mathbf{M}_{22}^H\right)^{-1}$ with $\mathbf{M}_{22}^{\dagger} \in \mathbb{C}^{N_r \times N_u}$, leading to

$$\mathbf{r} = -\mathbf{M}_{22}^{\dagger}\mathbf{M}_{21}\mathbf{d}. \tag{3.67}$$

With $\mathbf{T} = -\mathbf{M}_{22}^{\dagger}\mathbf{M}_{21}$ ($\mathbf{T} \in \mathbb{C}^{N_r \times N_d}$), the same linear mapping as in (3.7) follows with

$$\mathbf{r} = \mathbf{T}\mathbf{d}. \tag{3.68}$$

[5] Note that unambiguousness in this context assumes a given permutation matrix $\mathbf{P}$.

N_r	Index set
16	2,6,10,14,17,21,24,26,38,40,43,47,50,55,58,62
20	1,4,7,10,13,15,18,21,24,26,38,40,43,46,48,51,54,57,59,62
24	2,4,6,9,11,13,16,18,20,22,24,26,38,40,42,44,46,48, 51,54,56,58,60,63

Table 3.6.: Optimal index set of redundant subcarriers for $N_r > N_u$ and constant N_u according to Tab. 3.7.

	CP-OFDM	UW-OFDM		
N_d	48	36	32	28
N_r (N_p)	4	16	20	24
N_u (N_g)	16	16	16	16
B [MHz]	16.70	16.30	16.33	16.35
η [Msymb/s/Hz]	0.719	0.690	0.613	0.535
η_{rel}	1.000	0.960	0.853	0.744
NE_r		36.57	24.91	18.41
E_b	0.00977	0.01575	0.01390	0.01295
AWGN loss [dB]	0.0000	2.0754	1.5318	1.2254

Table 3.7.: Setup comparison of CP-OFDM and UW-OFDM with increasing N_r and constant N_u.

From that point on, all steps are identical to the case $N_r = N_u$, and the very same optimization routines and receiver concepts can be applied.

Starting from setup A with $N_r = 16$, Tab. 3.6 illustrates the subcarrier sets when gradually increasing N_r while keeping the length of the guard interval $N_u = 16$ constant. Each set has been derived by applying the optimization routine in Sec. 3.3 and minimizing the cost function J_{E} in (3.2).

Tab. 3.7 compares the resulting UW-OFDM setups against CP-OFDM setup A. Increasing N_r leads of course to a lower bandwidth efficiency on one hand, but decreases the mean energy E_r of the redundant subcarrier symbols and hence the mean energy per transmit bit E_b, and the AWGN loss on the other hand. With channel inversion only, the UW-OFDM system loses around 2.08 dB against CP-OFDM in case of $N_r = 16$, 1.53 dB for $N_r = 20$ and 1.23 dB for $N_r = 24$. However, a Wiener smoother as in (3.61) exploits the correlations between the subcarriers of an UW-OFDM symbol and improves the performance in the AWGN channel by around 1.5 dB for all investigated modes, cf. Fig. 3.21, leading to a residual performance loss compared to CP-OFDM of only 0.6 dB for $N_r = 16$ and 0.05 dB for $N_r = 20$. In case of $N_r = 24$, the UW-OFDM system even outperforms the

CP-OFDM system by 0.2 dB, but at the price of a significantly reduced bandwidth efficiency.

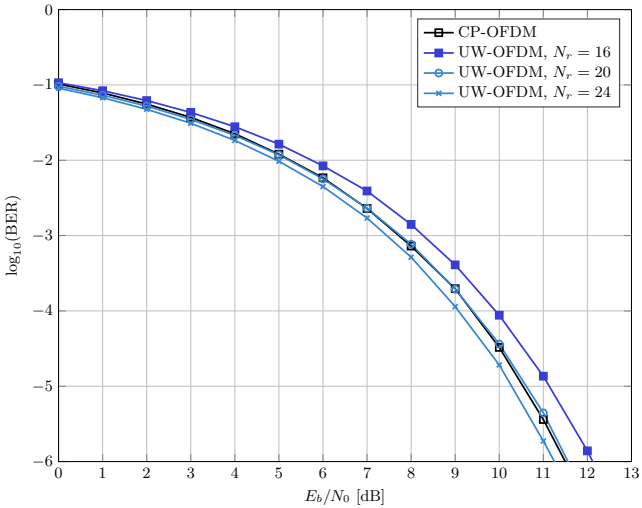


Figure 3.21.: BER comparison of UW-OFDM systems with increasing N_r and constant N_u for uncoded transmission in AWGN.

Fig. 3.22 presents the BER results in a frequency selective environment for the uncoded case. Increasing the number of redundant subcarriers N_r from 16 to 20 and 24 enhances the performance by 3.5 and 4.5 dB, respectively. However, this is only a moderate additional improvement considering the gain of 20.5 dB an UW-OFDM system with $N_r = 16$ already provides over CP-OFDM. Interestingly, a big portion of the gain originating from increasing N_r can even be transferred into the coded scenario, namely 1.3 dB and 1.8 dB for $r = 3/4$, and 0.7 and 1.2 dB for $r = 1/2$, cf. Fig. 3.23. This is especially astonishing, since the original gain of UW-OFDM with $N_r = 16$ over CP-OFDM reduces substantially when going from the uncoded to the coded case, as can be seen in Fig. 3.19 and 3.20.

In conclusion, increasing the number of redundant subcarriers N_r while keeping the length N_u of the UW constant can be interpreted as a means to vary the *inner coding rate* of an UW-OFDM system, while the *outer coding rate* is determined by the channel code. However, as UW-OFDM already has a slightly lower bandwidth efficiency than the competing CP-OFDM concept, lowering the redundant energy E_r by trading data against redundant subcarriers may not be the preferable way to improve the BER behavior.

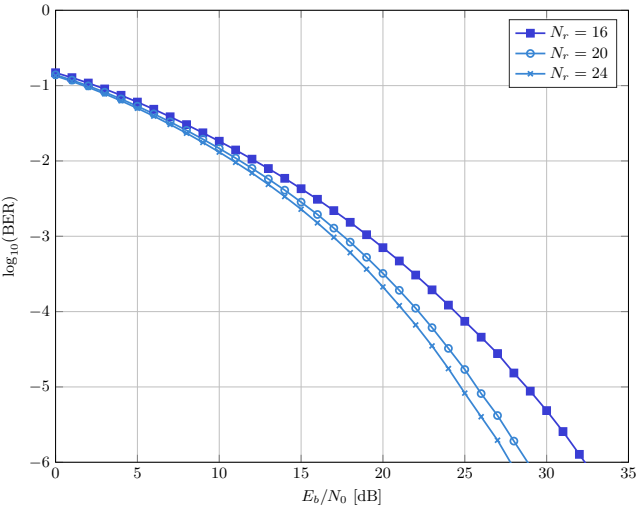


Figure 3.22.: BER comparison of UW-OFDM systems with increasing N_r and constant N_u for uncoded transmission in a multipath environment with $\tau_{\mathrm{RMS}} = 100\,\mathrm{ns}$.

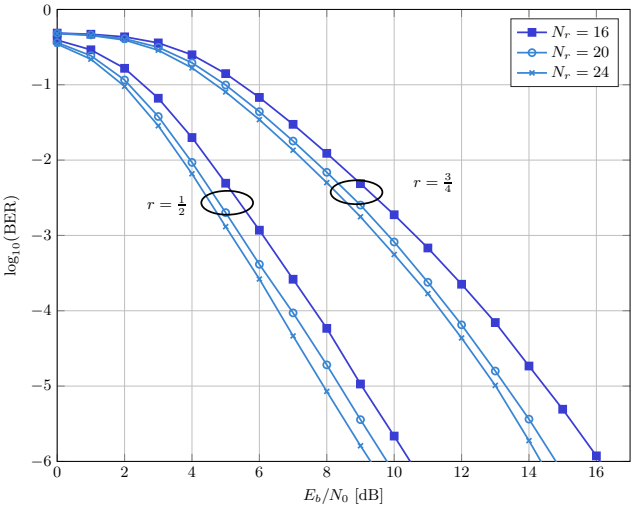


Figure 3.23.: BER comparison of UW-OFDM systems with increasing N_r and constant N_u for coded transmission ($r = 1/2$ and $r = 3/4$) in a multipath environment with $\tau_{\mathrm{RMS}} = 100\,\mathrm{ns}$.

3.7. Systematic Noise in UW-OFDM

The previous section showed that the redundant energy can only be reduced, if additional degrees of freedom are spent to ease the fulfillment of the zero word constraint. Clearly, sacrificing bandwidth efficiency seems too costly and should thus be avoided. The idea is now to establish an underdetermined system of equations to derive the redundant symbols similar to (3.66), but without reducing the number of data subcarriers. In [73] this is achieved by allowing some systematic noise in the guard interval instead of generating a pure zero word. The following part establishes a general framework, where the approach from [73] can be derived as a special case. As a starting point serves the equation

$$\mathbf{x} = \mathbf{F}_N^{-1}\mathbf{B}\mathbf{P}\begin{bmatrix}\mathbf{d}+\boldsymbol{\Delta}_d\\ \mathbf{r}\end{bmatrix} = \begin{bmatrix}\mathbf{x}_d\\ \mathbf{0}+\boldsymbol{\Delta}_{x_u}\end{bmatrix}. \tag{3.69}$$

This equation corresponds to the conventional approach described in (2.58) (with $\mathbf{A} = \mathbf{P}$ in $\mathbf{G}$), except for the vectors $\boldsymbol{\Delta}_d \in \mathbb{C}^{N_d\times 1}$ and $\boldsymbol{\Delta}_{x_u} \in \mathbb{C}^{N_u\times 1}$. The introduction of these additive deviations will, while neglecting the details at the moment, help to relax the zero word constraint and consequently reduce the redundant energy. In order to achieve an even more general framework, (3.69) is expanded towards

$$\mathbf{x} = \mathbf{F}_N^{-1}\begin{bmatrix}\mathbf{B}\mathbf{P} & \mathbf{B}_z\end{bmatrix}\begin{bmatrix}\mathbf{d}+\boldsymbol{\Delta}_d\\ \mathbf{r}\\ \mathbf{0}+\boldsymbol{\Delta}_z\end{bmatrix} = \begin{bmatrix}\mathbf{x}_d\\ \mathbf{0}+\boldsymbol{\Delta}_{x_u}\end{bmatrix}. \tag{3.70}$$

Here, $\boldsymbol{\Delta}_z \in \mathbb{C}^{N_z\times 1}$ represents an additive deviation on the zero subcarriers and $\mathbf{B}_z \in \mathbb{C}^{N\times N_z}$ a matrix that places the deviation on the appropriate zero subcarrier positions. Hence, the basic idea is to allow some *noise* at the data and zero subcarriers as well as in the guard interval to achieve the goal of reducing the redundant energy. As a next step let

$$\mathbf{F}_N^{-1}\begin{bmatrix}\mathbf{B}\mathbf{P} & \mathbf{B}_z\end{bmatrix} = \begin{bmatrix}\mathbf{F}_N^{-1}\mathbf{B}\mathbf{P} & \mathbf{F}_N^{-1}\mathbf{B}_z\end{bmatrix} = \begin{bmatrix}\mathbf{M} & \mathbf{M}'\end{bmatrix}, \tag{3.71}$$

whereas

$$\mathbf{M} = \begin{bmatrix}\mathbf{M}_{11} & \mathbf{M}_{12}\\ \mathbf{M}_{21} & \mathbf{M}_{22}\end{bmatrix} \tag{3.72}$$

$$\mathbf{M}' = \begin{bmatrix}\mathbf{M}_{13}\\ \mathbf{M}_{23}\end{bmatrix}. \tag{3.73}$$

Note that $\mathbf{M}$ and its submatrices correspond exactly to the definitions in (2.44). The matrices $\mathbf{M}' \in \mathbb{C}^{N\times N_z}$, $\mathbf{M}_{13} \in \mathbb{C}^{(N-N_u)\times N_z}$ and $\mathbf{M}_{23} \in \mathbb{C}^{N_u\times N_z}$ only originate from the fact that the 'zero subcarriers' are in fact not exactly zero anymore. In case systematic noise at the zero subcarriers is not allowed, $\mathbf{M}'$ vanishes and

$\begin{bmatrix}\mathbf{M} & \mathbf{M}'\end{bmatrix}$ collapses to the usual case as in (2.44). Inserting (3.71)-(3.73) into (3.70) yields

$$\begin{bmatrix}\mathbf{M}_{11} & \mathbf{M}_{12} & \mathbf{M}_{13}\\ \mathbf{M}_{21} & \mathbf{M}_{22} & \mathbf{M}_{23}\end{bmatrix}\begin{bmatrix}\mathbf{d}+\boldsymbol{\Delta}_d\\ \mathbf{r}\\ \mathbf{0}+\boldsymbol{\Delta}_z\end{bmatrix}=\begin{bmatrix}\mathbf{x}_d\\ \boldsymbol{\Delta}_{x_u}\end{bmatrix}, \tag{3.74}$$

and extracting the lower row delivers

$$\mathbf{M}_{21}\left(\mathbf{d}+\boldsymbol{\Delta}_d\right)+\mathbf{M}_{22}\mathbf{r}+\mathbf{M}_{23}\boldsymbol{\Delta}_z=\boldsymbol{\Delta}_{x_u}. \tag{3.75}$$

This leads to an underdetermined system of equations for $\boldsymbol{\Delta}_d$, $\mathbf{r}$, $\boldsymbol{\Delta}_z$ and $\boldsymbol{\Delta}_{x_u}$ that can be written as

$$\underbrace{\begin{bmatrix}\mathbf{M}_{22} & -\mathbf{I} & \mathbf{M}_{21} & \mathbf{M}_{23}\end{bmatrix}}_{\mathbf{A}}\underbrace{\begin{bmatrix}\mathbf{r}\\ \boldsymbol{\Delta}_{\mathbf{x}_u}\\ \boldsymbol{\Delta}_d\\ \boldsymbol{\Delta}_z\end{bmatrix}}_{z}=\underbrace{-\mathbf{M}_{21}\mathbf{d}}_{\mathbf{b}}, \tag{3.76}$$

with $\mathbf{A}\in\mathbb{C}^{N_u\times(N_r+N_u+N_d+N_z)}$, $\mathbf{z}\in\mathbb{C}^{(N_r+N_u+N_d+N_z)\times 1}$ and $\mathbf{b}\in\mathbb{C}^{N_u\times 1}$. Defining a weighting matrix $\mathbf{W}_n$ for the systematic noise distribution

$$\mathbf{W}_n=\operatorname{diag}\left(\begin{bmatrix}\mathbf{w}_r\\ \mathbf{w}_{x_u}\\ \mathbf{w}_d\\ \mathbf{w}_z\end{bmatrix}\right) \tag{3.77}$$

with the weighting vectors $\mathbf{w}_r\in\mathbb{R}_+^{N_r\times 1}$, $\mathbf{w}_{x_u}\in\mathbb{R}_+^{N_u\times 1}$, $\mathbf{w}_d\in\mathbb{R}_+^{N_d\times 1}$ and $\mathbf{w}_z\in\mathbb{R}_+^{N_z\times 1}$ consisting of non-negative elements, the problem translates to finding $\mathbf{z}$ that solves

$$\mathbf{A}\mathbf{z}=\mathbf{b}, \tag{3.78}$$

and that minimizes the cost function

$$g(\mathbf{z})=\mathbf{z}^H\mathbf{W}_n\mathbf{z}, \tag{3.79}$$

leading to the final optimization problem

$$\min\{\mathbf{z}^H\mathbf{W}_n\mathbf{z}\}\qquad\text{s.t. }\ \mathbf{A}\mathbf{z}=\mathbf{b}. \tag{3.80}$$

The solution of this optimization problem follows to

$$\begin{aligned}\mathbf{z}&=\mathbf{W}_n^{-1}\mathbf{A}^H(\mathbf{A}\mathbf{W}_n^{-1}\mathbf{A}^H)^{-1}\mathbf{b} &(3.81)\\ &=\mathbf{W}_n^{-1}\mathbf{A}^H(\mathbf{A}\mathbf{W}_n^{-1}\mathbf{A}^H)^{-1}\left(-\mathbf{M}_{21}\right)\mathbf{d} &(3.82)\\ &=\mathbf{T}'\mathbf{d}, &(3.83)\end{aligned}$$

cf. [42], with $\mathbf{T}'\in\mathbb{C}^{(N_r+N_u+N_d+N_z)\times N_d}$. Note that due to the additional degrees of freedom, this approach even allows $N_r<N_u$ [73].

For the following considerations, the weighting matrix is treated as given for the moment, a more detailed analysis of the impact will be provided later in this section. From $\mathbf{z}$ defined in (3.76), the additive noise vectors can easily be extracted by appropriate selection matrices:

$$\mathbf{r} = \begin{bmatrix}\mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0}\end{bmatrix}\mathbf{z} = \begin{bmatrix}\mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0}\end{bmatrix}\mathbf{T}'\mathbf{d} = \mathbf{T}_r\mathbf{d} \tag{3.84}$$

$$\boldsymbol{\Delta}_{x_u} = \begin{bmatrix}\mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0}\end{bmatrix}\mathbf{z} = \begin{bmatrix}\mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0}\end{bmatrix}\mathbf{T}'\mathbf{d} = \mathbf{T}_{x_u}\mathbf{d} \tag{3.85}$$

$$\boldsymbol{\Delta}_{d} = \begin{bmatrix}\mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0}\end{bmatrix}\mathbf{z} = \begin{bmatrix}\mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0}\end{bmatrix}\mathbf{T}'\mathbf{d} = \mathbf{T}_{d}\mathbf{d} \tag{3.86}$$

$$\boldsymbol{\Delta}_{z} = \begin{bmatrix}\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}\end{bmatrix}\mathbf{z} = \begin{bmatrix}\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}\end{bmatrix}\mathbf{T}'\mathbf{d} = \mathbf{T}_{z}\mathbf{d}. \tag{3.87}$$

These noise vectors depend on the actual data set and will thus vary from OFDM symbol to OFDM symbol. With (3.84)-(3.87) and (3.70) the frequency domain transmit symbol follows to

$$\tilde{\mathbf{x}} = \begin{bmatrix}\mathbf{BP} & \mathbf{B}_z\end{bmatrix}\begin{bmatrix}\mathbf{I}+\mathbf{T}_d \\ \mathbf{T}_r \\ \mathbf{T}_z\end{bmatrix}\mathbf{d} = \left(\mathbf{BP}\begin{bmatrix}\mathbf{I}+\mathbf{T}_d \\ \mathbf{T}_r\end{bmatrix} + \mathbf{B}_z\mathbf{T}_z\right)\mathbf{d} \tag{3.88}$$

$$= \mathbf{BGd} + \mathbf{G}_z\mathbf{d}. \tag{3.89}$$

Except for the additive term $\mathbf{G}_z\mathbf{d}$, the UW-OFDM frequency domain symbol $\tilde{\mathbf{x}}$ in (3.89) formally coincides with the one used in Sec. 2.2. There are now two cases to consider: In case of $\Delta_z = \mathbf{0}$ corresponding to an exclusion of the zero subcarriers from the optimization problem in (3.80), the additive term $\mathbf{G}_z\mathbf{d}$ vanishes and the same system model and thus receiver concepts as in Sec. 2.3 apply. In case of $\Delta_z \neq \mathbf{0}$, subtracting $\mathbf{G}_z\mathbf{d}$ will lead again to the same system model, however, the redundancy on the zero subcarriers cannot be used. In order to fully exploit the inherent redundancy, let us introduce the generator matrix

$$\mathbf{G}' = \mathbf{BG} + \mathbf{G}_z \qquad \mathbf{G}' \in \mathbb{C}^{N\times N_d}. \tag{3.90}$$

Replacing $\mathbf{BG}$ with $\mathbf{G}'$ in (2.60) and additionally not excluding entries corresponding to zero subcarriers in (2.62), the original formalism holds and the very same receiver concepts can be employed again.

So far, the weighting matrix has been considered as given and not been detailed yet. In fact, $\mathbf{W}_n$ provides a means to control the distribution of the systematic noise, i.e., the higher the weighting coefficient the lower the average power of the noise placed on a specific subcarrier/sample. The determining factor for the resulting distribution is the relative value of a weighting coefficient compared to the others rather than its absolute value. The presented approach provides a very general framework with a variety of degrees of freedom, but does not consider its applicability to practical scenarios. For instance, it might not be beneficial to put additive noise on zero subcarriers and thus destroy the spectral properties, especially when considering the superior out-of-band-emissions of UW-OFDM (Fig. 3.17). Any

subcarrier or guard interval sample can easily be excluded from the optimization by setting the appropriate weighting coefficient w_k to

$$w_k = [\mathbf{W}_n]_{k,k} = \infty. \tag{3.91}$$

While analytically well defined, for a numerical optimization w_k might be initialized with the maximum representable value of a given data type. Alternatively, an exclusion can also be implemented by introducing an appropriate selection matrix $\mathbf{S} \in \{0,1\}^{(N_r+N_u+N_d+N_z-N_e)\times(N_r+N_u+N_d+N_z)}$ to avoid solving an ill-conditioned system of equations numerically. N_e determines the number of excluded elements and $\mathbf{S}$ is derived from an identity matrix by deleting the rows corresponding to the indices of the elements in $\mathbf{z}$ and $\mathbf{w}$ to be excluded. In order to utilize this approach, one has to replace $\mathbf{z}$ by $\mathbf{z}' = \mathbf{S}\mathbf{z}$ and substitute $\mathbf{W}_n$ by $\mathbf{W}'_n = \mathbf{S}\mathbf{W}_n\mathbf{S}^T$ and $\mathbf{A}$ by $\mathbf{A}' = \mathbf{A}\mathbf{S}^T$ in (3.76)-(3.83).

3.7.1. Systematic Noise in the Guard Interval

In [73] only the guard interval deviation Δ_{x_u} has been used to further reduce the redundant energy. The motivation originates from the fact that additive noise is always present in any system. Perfect cyclic data structure (to transform the linear into a cyclic convolution) is not achievable anyway in a noisy environment and a well-designed second noise component does not significantly further degrade this property. The weighing coefficients are chosen as $\mathbf{w}_r = \mathbf{1}$ where $\mathbf{1}$ denotes a column vector with all entries being 1. The weighting vector for the guard interval $\mathbf{w}_{x_u}$ consists of elements

$$w_{x_u,k} = Ae^{k/\tau} \qquad k = 0, 1, \ldots, N_u - 1, \tag{3.92}$$

with varying values for A and $\tau = 2$ determined after several experimental runs. The concept of weighting the samples differently instead of equally originates from the idea that the cyclic convolution property might be less destroyed in the first case. In case of (3.92), the resulting average systematic noise power should exponentially decrease along the UW samples, i.e., in an inverse manner compared to the weighting coefficients. These coefficients are a result of the exponentially decaying taps of the channel model utilized in this work [46]. Note that different noise variances along the UW samples may applicable only in certain cases like UW-OFDM systems with zero UWs. For systems with non-zero UWs used for synchronization and estimation tasks, additional noise shall (if tolerable at all) probably be equally distributed.

Tab. 3.8 summarizes the setups investigated in the following, which all originate from setup A in Tab. 2.2. For a clear distinction, these subsetups are now labeled numerically instead of alphabetically. The first three follow the weighting approach in (3.92) with different values for A, the fourth one weights all samples equally and achieves the same redundant energy E_r as setup 3. Due to the different weighting

Table 3.8.: Setups for UW-OFDM systems with systematic noise in the guard interval, derived from setup A in Tab. 2.2.

	Setup 1	Setup 2	Setup 3	Setup 4
$w_{x_u,k}$	$1e^{k/2}$	$2e^{k/2}$	$8e^{k/2}$	453.75
$E_r N$	6.19	8.71	14.22	14.22
$E_{\mathbf{\Delta}_{x_u}} N$	5.38	4.63	3.13	1.11

concepts, the same amount of redundant energy is distributed slightly different as demonstrated in Fig. 3.24.

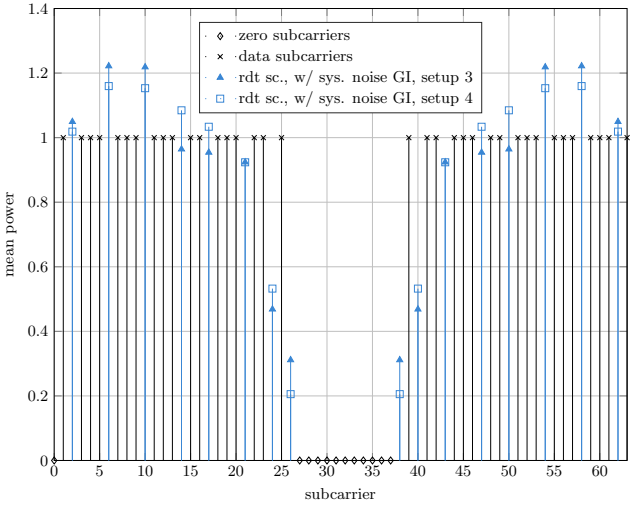


Figure 3.24.: Mean power of individual subcarrier symbols for setup 3 and setup 4 (Tab. 3.8) of an UW-OFDM system with systematic noise in the guard interval.

Fig. 3.25 illustrates the performance in an AWGN scenario. Note that the curve labeled as 'UW-OFDM (w/o sys. noise)' corresponds to the conventional systematically encoded UW-OFDM system with LMMSE estimation shown in Fig. 3.18. The same applies analogously for the subsequent plots visualizing different channel scenarios. As expected, the performance improves with decreasing redundant energy at the price of increasing noise power in the guard interval, cf. Tab. 3.8. Since the cyclic structure property is not required for equalization of a frequency

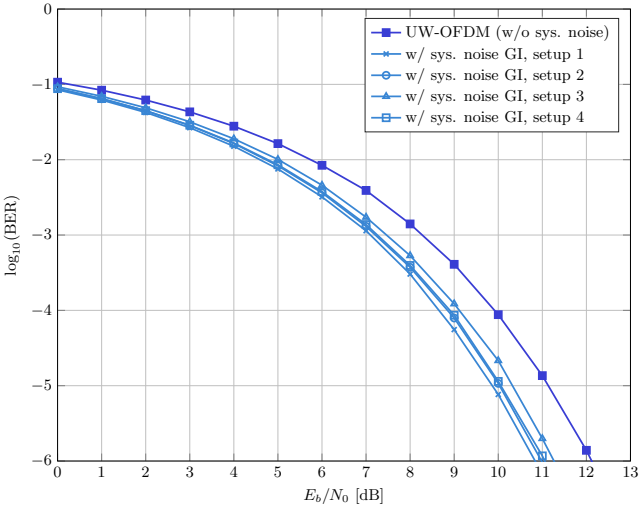


Figure 3.25.: BER comparison of UW-OFDM systems with systematic noise in the guard interval for various setups (Tab. 3.8) and uncoded transmission in AWGN.

flat channel, the power level of the systematic noise does not degrade the BER performance. Although setup 3 and 4 feature the same value for E_r, the BER values slightly differ. This can easily be explained by evaluating the error variances on the subcarriers after LMMSE detection according to (2.84). The sum as well as the maximum single error variance on a subcarrier will always be higher for setup 3 irrespective of the additive noise variance σ_n^2.

Uncoded multipath transmission as illustrated in Fig. 3.26 reveals the price of systematic noise in the guard interval. At a certain E_b/N_0 value, the systematic noise dominates the additive noise component and an error floor emerges. Consequently, the lower the redundant energy and thus the higher the systematic noise energy, the earlier the error floor arises. However, comparing setup 3 and setup 4 reveals that an intelligent distribution of the systematic noise along the guard interval indeed destroys the cyclic property to a lesser extent. Fig. 3.27 illustrates the average systematic noise power on the guard interval samples for both setups. Since the average total energy $E_{\boldsymbol{\Delta}_{x_u}}$ is clearly lower for setup 4, cf. Tab. 3.8, the average power of the systematic noise samples at the end of the guard interval have to cause the earlier error floor of setup 4.

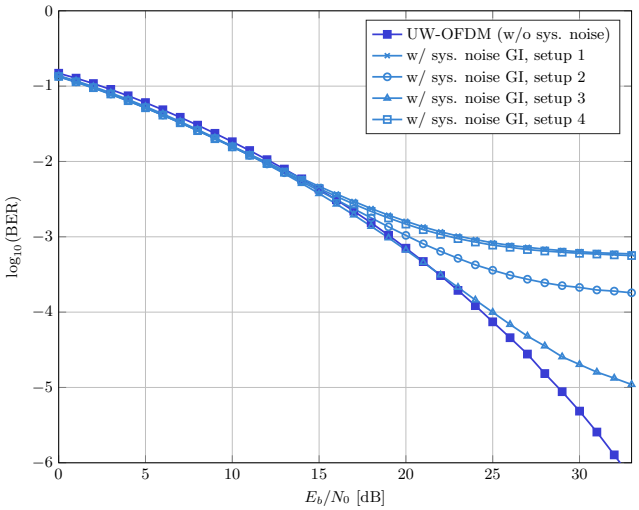


Figure 3.26.: BER comparison of UW-OFDM systems with systematic noise in the guard interval for various setups (Tab. 3.8) and uncoded transmission in a multipath environment with $\tau_{\mathrm{RMS}} = 100\,\mathrm{ns}$.

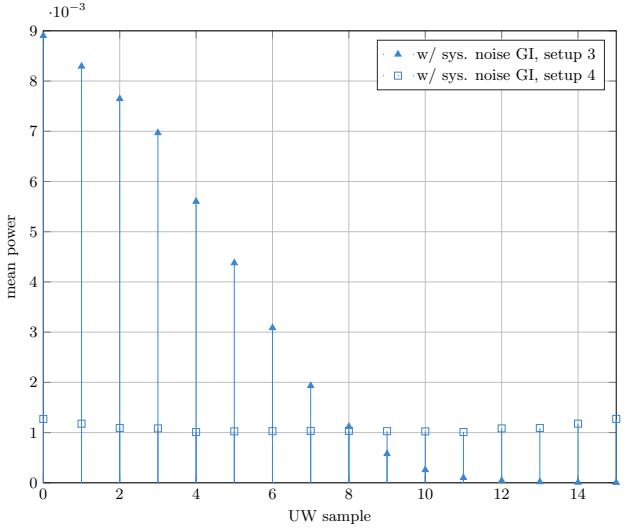


Figure 3.27.: Comparison of mean power of systematic noise samples in the guard interval for setup 3 and setup 4 (Tab. 3.8).

Despite the error floor, this concept still offers benefits for coded transmission as observable from Fig. 3.28. A gain up to 0.8 dB at a BER of 10^{-6} for $r = 1/2$ and 0.6 dB for $r = 3/4$ is achieved by the tested setups in comparison to conventional systematically encoded UW-OFDM. Due to the limited E_b/N_0 region in consideration, a weighting approach leading to an early error floor (in the uncoded case) can still provide the biggest gain within this scenario.

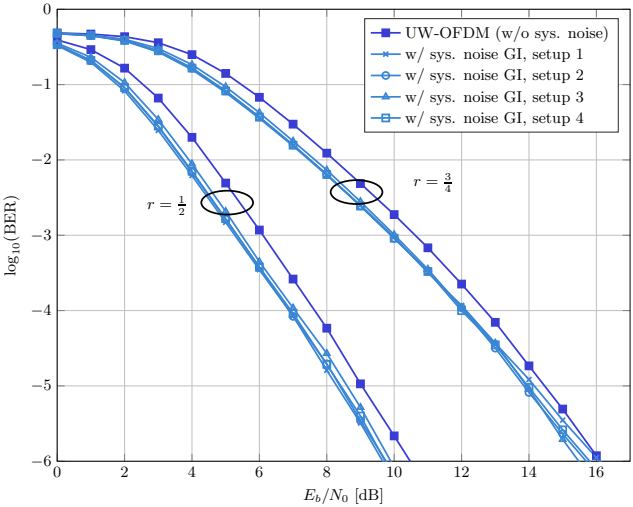


Figure 3.28.: BER comparison of UW-OFDM systems with systematic noise in the guard interval for various setups (Tab. 3.8) and coded transmission ($r = 1/2$ and $r = 3/4$) in a multipath environment with $\tau_{\mathrm{RMS}} = 100$ ns.

3.7.2. Systematic Noise on the Data Subcarriers

Next, systematic noise on the data subcarriers is studied instead. From (3.69) and (3.88) it follows that

$$\mathbf{x} = \mathbf{F}_N^{-1}\mathbf{B}\mathbf{P}\begin{bmatrix}\mathbf{d}+\Delta_d\\ \mathbf{r}\end{bmatrix} = \mathbf{F}_N^{-1}\mathbf{B}\mathbf{P}\begin{bmatrix}\mathbf{d}+\mathbf{T}_\mathbf{d}\mathbf{d}\\ \mathbf{r}\end{bmatrix} \tag{3.93}$$

$$= \mathbf{F}_N^{-1}\mathbf{B}\mathbf{P}\begin{bmatrix}\mathbf{I}+\mathbf{T}_\mathbf{d}\\ \mathbf{T}_r\end{bmatrix}\mathbf{d} = \mathbf{F}_N^{-1}\mathbf{B}\mathbf{G}^{(\mathrm{nd})}\mathbf{d}, \tag{3.94}$$

with the generator matrix

$$\mathbf{G}^{(\mathrm{nd})} = \mathbf{P}\begin{bmatrix}\mathbf{I}+\mathbf{T}_d\\ \mathbf{T}_r\end{bmatrix} \tag{3.95}$$

distributing redundancy on the data as well as on the redundant subcarriers. Consequently, a codeword of the original form $\mathbf{c} = \mathbf{P}\begin{bmatrix}\mathbf{d}^T & \mathbf{r}^T\end{bmatrix}^T$ in (3.9) is also extended to

$$\mathbf{c} = \mathbf{P}\begin{bmatrix}\mathbf{d}+\Delta_d\\ \mathbf{r}\end{bmatrix}. \tag{3.96}$$

Fig. 3.29 displays the resulting generator matrix for $\mathbf{w}_r = \mathbf{1}$ and $\mathbf{w}_d = 3\cdot\mathbf{1}$, the latter was found by experiments. $\mathbf{G}^{(\mathrm{nd})}$ stills shows a similar diagonal-like structure as $\mathbf{G}$ in Fig. 3.3, however, now each codeword symbol depends on several data symbols. Fig. 3.30 visualizes the BER performance of a system with $\mathbf{G}^{(\mathrm{nd})}$ in

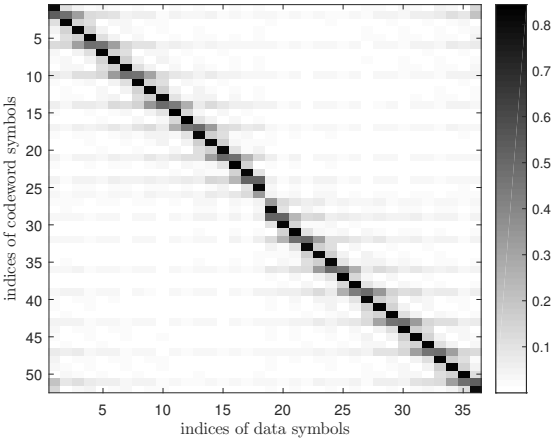


Figure 3.29.: $|\mathbf{G}^{(\mathrm{nd})}|$, magnitude of entries of generator matrix $\mathbf{G}^{(\mathrm{nd})}$ that generates systematic noise on the data subcarriers.

an AWGN scenario. Besides the conventional systematically encoded UW-OFDM, the best setup from Tab. 3.8 serves as an additional benchmark in the following. This is either setup 1 or setup 3, depending on the channel and the coding rate. For uncoded AWGN transmission the new approach outperforms the best UW-OFDM system with systematic noise by another 0.3 dB, leading to a total performance gain of 1.3 dB over the conventional system. For uncoded transmission in a multipath environment, the gain even becomes 1.4 dB. Furthermore, Fig. 3.31 clearly indicates that systematic noise on the data subcarriers instead of in the guard interval does not end up in an error floor. For coded transmission, the performance gain increases

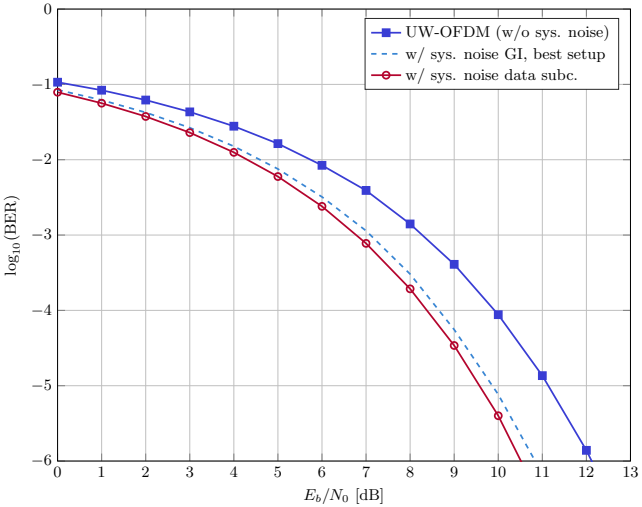


Figure 3.30.: BER comparison of an UW-OFDM system with systematic noise either in the guard interval or on the data subcarriers for uncoded transmission in AWGN.

by 0.4 to 1.2 dB for $r = 1/2$ and by 0.8 to a total of 1.4 dB for $r = 3/4$ over systematically encoded UW-OFDM without systematic noise. Studying the power distribution over the subcarriers in Fig. 3.32 shows that except for the subcarriers close to the band edges, the average power on the data and redundant subcarriers is now almost identical.

In conclusion, the latter concept can in a certain way be interpreted as giving up the clear separation between data and redundant subcarrier symbols. The original idea of putting *noise* and therefore disturbance on the data subcarriers might not be an appropriate interpretation anymore. In fact, these promising results motivate a transition from a systematically encoded to a non-systematically encoded UW-OFDM concept. This will be the focus of the following chapter.

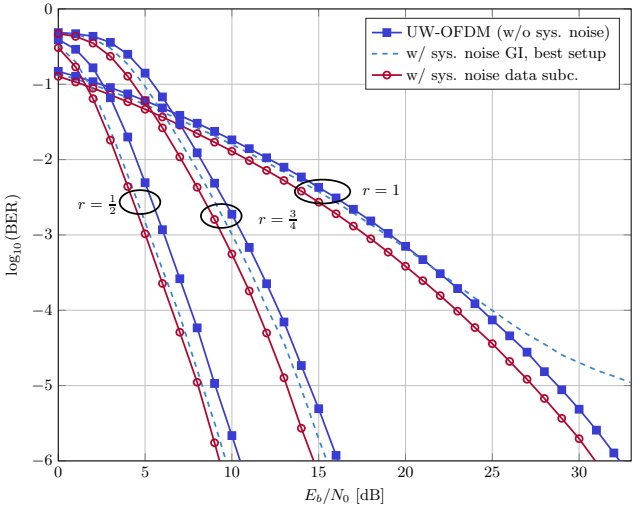


Figure 3.31.: BER comparison of an UW-OFDM system with systematic noise either in the guard interval or on the data subcarriers for uncoded ($r = 1$) and coded transmission ($r = 1/2$ and $r = 3/4$) in a multipath environment with $\tau_{\mathrm{RMS}} = 100\,\mathrm{ns}$.

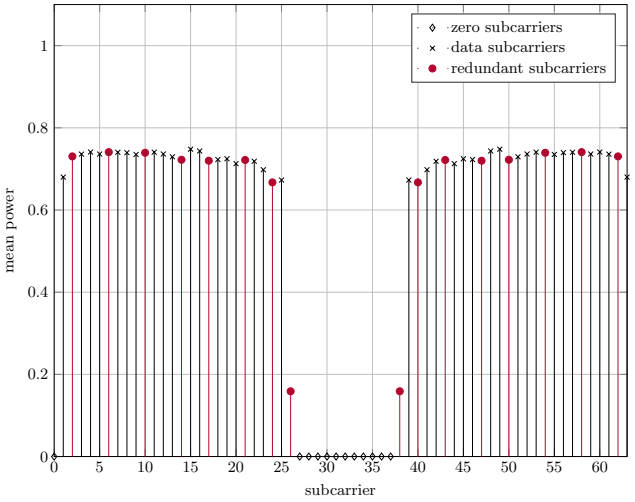


Figure 3.32.: Mean power of individual subcarrier symbols of an UW-OFDM system with systematic noise on the data subcarriers.

4. Non-Systematically Encoded Unique Word OFDM

The concept of systematically encoded UW-OFDM presented in the previous chapter has already proven its potential over conventional CP-OFDM. The results demonstrated a predominance in terms of spectral properties and BER performance. It turned out that minimizing the redundant subcarrier energy is the key to achieve a practically applicable system. Despite this optimization, the energy of the redundant subcarrier symbols is still considerably higher than that of the data symbols (Fig. 3.4), demanding further improvement. Consequently, possibilities for an additional minimization have been investigated:

- In Sec. 3.6 the redundant energy is reduced by increasing the number of redundant subcarriers. Unfortunately, this performance enhancement comes at the price of a decreased bandwidth efficiency.
- In Sec. 3.7 the introduction of systematic noise in the guard interval decreases the redundant energy which manifests in a BER performance improvement. The drawback of this method is an inevitable error floor, arising as soon as the systematic noise dominates the AWGN component.
- Additionally, Sec. 3.7 presents the concept of systematic noise on the data subcarriers. This method delivers a performance gain without the disadvantages of the first two approaches.

Note that the latter does not correspond to a classical systematically encoded UW-OFDM approach anymore. A redundant subcarrier still serves the same purpose, but a data subcarrier is now loaded with a data symbol and systematic noise arising from a weighted superposition of all other data symbols. Hence, data and some sort of redundancy are present on the former data subcarrier. The data symbols are not explicitly visible anymore, as it is supposed to for a systematic encoding approach. However, the impressive performance of this approach motivates to completely abandon the idea of dedicated redundant/data subcarriers and instead to introduce the idea of non-systematically encoded UW-OFDM in the following.

The interpretation of generating code words according to

$$\mathbf{c} = \mathbf{G}\mathbf{d} \tag{4.1}$$

still holds, but in contrast to systematically encoded UW-OFDM, $\mathbf{A}$ in

$$\mathbf{G} = \mathbf{A}\begin{bmatrix}\mathbf{I}\\ \mathbf{T}\end{bmatrix} \tag{4.2}$$

is not restricted to a permutation matrix anymore, cf. (3.3). Instead, the very general definition in (2.42) applies, where $\mathbf{A}$ can be any nonsingular real-valued[1] matrix. Furthermore, a non-systematic approach does not know dedicated redundant subcarriers anymore, hence minimizing the redundant energy cannot serve as a cost function. Alternatively, the same cost functions J_{BLUE} and J_{LMMSE} as in Sec. 3.4 are utilized which measure the sum of the error variances after BLUE and LMMSE detection for a fixed SNR ratio, respectively. Since both cost functions will slightly differ from the systematic approach case due to the different properties of $\mathbf{A}$, they are stated again in the following. Analogously to (3.43), the sum of the error variances $J'(\mathbf{A})$ for the BLUE in an AWGN scenario is

$$J'(\mathbf{A}) = \mathrm{tr}\left(\mathbf{C}_{ee}\right) = N\sigma_n^2 \mathrm{tr}\left(\left(\mathbf{G}^H\mathbf{G}\right)^{-1}\right). \tag{4.3}$$

Fixing the ratio $\gamma = E_s/\sigma_n^2$, where

$$E_s = \frac{E_x}{N_d} \tag{4.4}$$

denotes the mean data symbol energy and

$$\begin{aligned} E_x &= \frac{1}{N}\mathrm{E}\left\{\mathrm{tr}\left(\mathbf{G}^H\mathbf{G}\mathbf{d}\mathbf{d}^H\right)\right\} \\ &= \frac{1}{N}\mathrm{tr}\left(\mathrm{E}\left\{\mathbf{G}^H\mathbf{G}\mathbf{d}\mathbf{d}^H\right\}\right) \\ &= \frac{\sigma_d^2}{N}\mathrm{tr}\left(\mathbf{G}^H\mathbf{G}\right) \end{aligned} \tag{4.5}$$

the mean OFDM symbol energy, the variance of the noise is given by

$$\sigma_n^2 = \frac{E_s}{\gamma} = \frac{\sigma_d^2 \mathrm{tr}\left(\mathbf{G}^H\mathbf{G}\right)}{\gamma N N_d}. \tag{4.6}$$

With (4.6) the cost function follows to

$$J_{\mathrm{BLUE}}(\mathbf{A}) = \frac{\sigma_d^2}{\gamma N_d}\mathrm{tr}\left(\mathbf{G}^H\mathbf{G}\right)\mathrm{tr}\left(\left(\mathbf{G}^H\mathbf{G}\right)^{-1}\right). \tag{4.7}$$

In case of an LMMSE estimator, $\mathrm{tr}\left(\mathbf{C}_{ee}\right)$ becomes

$$J'(\mathbf{A}) = N\sigma_n^2 \mathrm{tr}\left(\left(\mathbf{G}^H\mathbf{G} + \frac{N\sigma_n^2}{\sigma_d^2}\mathbf{I}\right)^{-1}\right). \tag{4.8}$$

[1] In order to increase the number of available degrees of freedom, $\mathbf{A}$ could also be chosen to be a complex-valued instead of a real-valued matrix. However, investigations showed that the global minimum presented later in Sec. 4.1 is already achieved with a real-valued matrix, hence a definition over the real numbers is sufficient.

Inserting (4.6) into (4.8) delivers

$$J_{\mathrm{LMMSE}}(\mathbf{A}) = \frac{\sigma_d^2}{\gamma N_d}\mathrm{tr}\left(\mathbf{G}^H\mathbf{G}\right)\mathrm{tr}\left(\left(\mathbf{G}^H\mathbf{G} + \frac{\mathrm{tr}\left(\mathbf{G}^H\mathbf{G}\right)}{\gamma N_d}\mathbf{I}\right)^{-1}\right) \tag{4.9}$$

$$= \sigma_d^2\mathrm{tr}\left(\left(\frac{\gamma N_d}{\mathrm{tr}\left(\mathbf{G}^H\mathbf{G}\right)}\mathbf{G}^H\mathbf{G} + \mathbf{I}\right)^{-1}\right). \tag{4.10}$$

Finally, the constrained optimization problem to find $\mathbf{G}$ reads

$$\check{\mathbf{A}} = \underset{\mathbf{A}}{\mathrm{argmin}}\{J_{\mathrm{BLUE}} \vee J_{\mathrm{LMMSE}}\} \quad \text{s.t. } \mathbf{F}_N^{-1}\mathbf{B}\mathbf{G} = \begin{bmatrix}\boldsymbol{\Xi}\\ \mathbf{0}\end{bmatrix} \wedge \mathbf{G} = \mathbf{A}\begin{bmatrix}\mathbf{I}\\ \mathbf{T}\end{bmatrix}. \tag{4.11}$$

The solutions to the optimization problem will lead to code generator matrices $\mathbf{G}$ matched to the BLUE or LMMSE detection procedure. For sufficiently large γ, J_{LMMSE} approaches J_{BLUE} and the particular choice of γ is irrelevant for the optimization procedure.

The resulting code generator matrices are explicitly optimized for the AWGN case with $\mathbf{H} = \mathbf{I}$. Otherwise, a generator matrix may be perfectly suitable only for one specific frequency selective transmission channel, but may fail in other situations. Of course, selecting the generator matrix depending on the transmission conditions could enhance the performance further, but would also require channel state information (CSI) at the transmitter for a proper selection. This work considers only CSI free transmitter design by intention, results towards a generator matrix design taking into account the transmission channel can be found in [74, 75].

4.1. Signal Properties

This section delivers an overview over the signal properties of non-systematically encoded UW-OFDM symbols. As in Sec. 3.1 for systematically encoded UW-OFDM, the level of detailing increases successively. The first level denotes thus the highest degree of abstraction and focuses on the overall generation process. Insights are based on the whole OFDM symbol. The second level presents results based on the properties of single subcarrier symbols.

Sec. 4.2 will later present a numerical way to solve the optimization problem in (4.11) by a steepest descent algorithm. The solution is ambiguous and depends on the initialization of a matrix $\mathbf{A}^{(0)}$ in the algorithm, leading to completely different generator matrices. Until explicitly specified otherwise, the following results are valid for all optimum generator matrix instances implementing non-systematically encoded UW-OFDM according to (4.11).

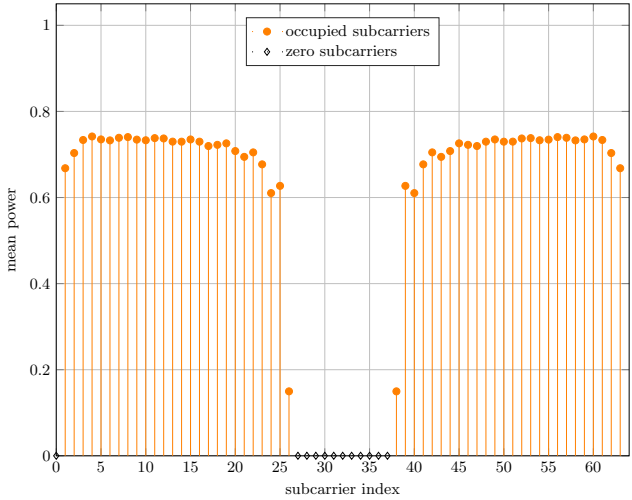


Figure 4.1.: Mean power of individual subcarrier symbols for non-systematically encoded UW-OFDM with $\mathbf{G}'$. The particular generator matrix $\mathbf{G}'$ will be explained in detail later in this chapter.

Properties drawn from an OFDM symbol. Optimum generator matrices have been found by minimizing the cost functions J_{BLUE} or J_{LMMSE}. Fig. 4.1 shows the mean power of the subcarriers. Contrary to *conventional* systematically encoded UW-OFDM (Fig. 3.4), all non-zero subcarriers experience almost the same mean power value. However, the results are very similar to systematically encoded UW-OFDM with systematic noise on the data subcarriers (Fig. 3.32). These values can be analytically expressed as $\sigma_d^2 \mathrm{diag}\left(\mathbf{G}\mathbf{G}^H\right)$, cf. (3.11).

The following results refer to the BLUE and are then extended to the LMMSE case. It turns out that all found local minima for the BLUE based numerical optimization (based on the algorithm presented in Sec. 4.2) feature the very same value $J_{\mathrm{BLUE,min}}$ for the cost function, independent of the initialization of a matrix $\mathbf{A}^{(0)}$. Furthermore, all corresponding code generator matrices feature the property

$$\mathbf{G}^H\mathbf{G} = \alpha\mathbf{I} \tag{4.12}$$

with some real constant $\alpha > 0$ (which varies dependent on the results of the optimization process). This property[2] has a number of important implications.

[2] The property $\mathbf{G}^H\mathbf{G} = \alpha\mathbf{I}$ might mislead to the impression of $\mathbf{G}$ being a unitary matrix. Note, however, that the generator matrix $\mathbf{G}$ is a rectangular matrix, while a unitary matrix is quadratic by definition.

First, inserting (4.12) into the cost function (4.7) leads to

$$J_{\mathrm{BLUE,min}} = \frac{\sigma_d^2}{\gamma N_d}(N_d\alpha)(N_d\alpha^{-1}) = \frac{\sigma_d^2 N_d}{\gamma}, \tag{4.13}$$

which coincides with the numerically found local minima. Hence, every $\mathbf{G}$ fulfilling

$$\mathbf{G}^H\mathbf{G} = \alpha\mathbf{I} \quad \text{and} \quad \mathbf{F}_N^{-1}\mathbf{B}\mathbf{G} = \begin{bmatrix}\boldsymbol{\Xi}\\ \mathbf{0}\end{bmatrix} \tag{4.14}$$

with any α delivers the same value $J_{\mathrm{BLUE,min}}$ of the cost function and produces a zero UW in time domain.

Second, each code generator matrix satisfying (4.14) induces an error covariance matrix after the data estimation in the AWGN channel given by

$$\mathbf{C}_{ee,\mathrm{BLUE}} = \frac{\sigma_d^2}{\gamma}\mathbf{I}. \tag{4.15}$$

This simply follows from (4.12), (4.6) and the definition of the error covariance matrix in (4.3). Consequently, the noise at the output of the BLUE is uncorrelated under AWGN conditions, which is clearly in contrast to systematically encoded UW-OFDM, where $\mathbf{C}_{ee}$ has non-zero off-diagonal entries also in the AWGN channel case.

And third, (4.12) implies that all singular values of $\mathbf{G}$ are identical. To prove this, consider a singular value decomposition (SVD)

$$\mathbf{G} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H, \tag{4.16}$$

with unitary matrices $\mathbf{U}$ and $\mathbf{V}$, and with the matrix $\boldsymbol{\Sigma} = \begin{bmatrix}\mathbf{D} & \mathbf{0}\end{bmatrix}^T$, where $\mathbf{D}$ is a real diagonal matrix having the singular values $s_1, s_2, ..., s_{N_d}$ of $\mathbf{G}$ at its main diagonal. With (4.12) this leads to

$$\begin{aligned}\alpha\mathbf{I} = \mathbf{G}^H\mathbf{G} &= \mathbf{V}\boldsymbol{\Sigma}^H\mathbf{U}^H\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H = \mathbf{V}\mathbf{D}^2\mathbf{V}^H\\ \Leftrightarrow \alpha\mathbf{I} = \mathbf{D}^2 &= \mathrm{diag}\left(s_1^2, s_2^2, ..., s_{N_d}^2\right).\end{aligned} \tag{4.17}$$

Therefore, $\mathbf{G}^H\mathbf{G} = \alpha\mathbf{I}$ implies $\alpha = s_1^2 = s_2^2 = \cdots s_{N_d}^2 := s^2$. The property in (4.12) can therefore also be written as

$$\mathbf{G}^H\mathbf{G} = s^2\mathbf{I}. \tag{4.18}$$

The argumentation can also be done the other way round: If all singular values of $\mathbf{G}$ are identical then $\mathbf{G}^H\mathbf{G} = \alpha\mathbf{I}$ with $\alpha = s^2$.

An open question is still whether $J_{\mathrm{BLUE,min}}$ in (4.13) corresponding to the numerically found local minima also depicts the global minimum of the constrained optimization problem in (4.11). To answer this question the general formulation

of the cost function J_{BLUE} in (4.7) serves as a starting point and the constraint in (4.11) is disregarded for a moment. Let $\mathbf{s} = \begin{bmatrix} s_1 & s_2 & \cdots & s_{N_d} \end{bmatrix}^T$ with $s_i > 0$ be the vector of singular values of $\mathbf{G}$. Appendix B will prove that $\partial J_{\mathrm{BLUE}}/\partial \mathbf{s} = \mathbf{0}$ if and only if all singular values of $\mathbf{G}$ are identical. Consequently, every possible candidate $\mathbf{G}$ for a local minimum of the unconstrained problem satisfies $\mathbf{G}^H\mathbf{G} = s^2\mathbf{I}$ (cf. (4.17) and its implications). Inserting $\mathbf{G}^H\mathbf{G} = s^2\mathbf{I}$ into the cost function (4.7) leads to the same expression as in (4.13). Hence, every $\mathbf{G}$ fulfilling $\mathbf{G}^H\mathbf{G} = s^2\mathbf{I}$ results in the same (and minimum) value

$$J_{\mathrm{BLUE,min}} = \sigma_d^2 N_d/\gamma, \tag{4.19}$$

which therefore constitutes the global minimum of the cost function (see also [76] for an analytical proof).

So far, only the link to the constrained problem in (4.11) is missing. The numerical solutions showed the existence of matrices, that firstly satisfy $\mathbf{G}^H\mathbf{G} = s^2\mathbf{I}$ and therefore result in the global minimum of the cost function J_{BLUE}, and that secondly fulfill the constraint $\mathbf{F}_N^{-1}\mathbf{B}\mathbf{G} = \begin{bmatrix} \boldsymbol{\Xi} \\ \mathbf{0} \end{bmatrix}$. These considerations finally yield the following important proposition:

Optimum code generator matrices: *A code generator matrix* $\mathbf{G}$ *is optimum, i.e., leads to a global minimum of the constrained optimization problem in* (4.11)*, if and only if* $\mathbf{G}$ *satisfies*

$$\mathbf{G}^H\mathbf{G} = s^2\mathbf{I} \quad \textit{and} \tag{4.20}$$

$$\mathbf{F}_N^{-1}\mathbf{B}\mathbf{G} = \begin{bmatrix} \boldsymbol{\Xi} \\ \mathbf{0} \end{bmatrix}, \tag{4.21}$$

where $s := s_1 = s_2 = \cdots = s_{N_d} > 0$ *are the (all identical) singular values of* $\mathbf{G}$*. The global minimum of the cost function is given by* (4.13)*, and the error covariance matrix after data estimation in the AWGN channel is the scaled identity matrix as given in* (4.15)*.*

According to (4.13) the particular value of $\alpha = s^2$ does not influence the resulting cost function value. Hence, all found code generator matrices are normalized in the following such that $\alpha = s^2 = 1$ or $\mathbf{G}^H\mathbf{G} = \mathbf{I}$. Note that the columns of any normalized optimum code generator matrix $\mathbf{G}$ form an orthonormal basis of an N_d-dimensional subspace of $\mathbb{C}^{(N_d+N_r)\times 1}$. As a further consequence of $s^2 = 1$, the operation $\mathbf{c} = \mathbf{G}\mathbf{d}$ becomes energy-invariant, leading for one particular realization of $\mathbf{c}$ to a transmit symbol energy of

$$E_x^{(r)} = \frac{1}{N}\mathbf{c}^H\gamma = \frac{1}{N}\mathbf{d}^H\mathbf{G}^H\mathbf{G}\mathbf{d} = \frac{1}{N}\mathbf{d}^H\mathbf{d}. \tag{4.22}$$

Without loss of generality a zero UW is assumed in this case. Consequently, $E_x^{(r)} = \frac{1}{N}\mathrm{tr}\left(\mathbf{d}\mathbf{d}^H\right)$, and the mean symbol energy becomes $E_x = \frac{1}{N}\mathrm{E}\left\{\mathrm{tr}\left(\mathbf{d}\mathbf{d}^H\right)\right\} = \frac{N_d\sigma_d^2}{N}$. Hence, for a data symbol alphabet with elements having all the same power like QPSK, the transmit energy appears to be constant for every OFDM symbol. This is in clear contrast to systematically encoded UW-OFDM, cf. (3.17).

All conclusions drawn so far hold for the LMMSE estimator based transceiver optimization as well, only the particular expressions $J_{\mathrm{LMMSE,min}}$ and $\mathbf{C}_{ee,\mathrm{LMMSE}}$ differ slightly. With (4.12) and (4.9) it immediately follows that

$$J_{\mathrm{LMMSE,min}} = \frac{\sigma_d^2 N_d}{\gamma + 1}, \tag{4.23}$$

$$\mathbf{C}_{ee,\mathrm{LMMSE}} = \frac{\sigma_d^2}{\gamma + 1}\mathbf{I}. \tag{4.24}$$

In summary, a code generator matrix which is optimum for the BLUE based 'decoding' procedure is automatically also optimum for the LMMSE based data estimation (and vice versa). Appendix C proves analytically that also $\partial J_{\mathrm{LMMSE}}/\partial\mathbf{s} = \mathbf{0}$, if and only if all singular values of $\mathbf{G}$ are identical.

Up till now all presented properties hold for any generator matrix realizing a non-systematically encoded UW-OFDM system. In the subsequent paragraphs though, the focus is laid on the diverse properties of generator matrices originating from different initializations $\mathbf{A}^{(0)}$ in the optimization procedure. Out of many possibilities, two representative initializations are picked in the following to demonstrate the potential of the non-systematic approach:

1. The first initialization is chosen as

$$\mathbf{A}^{(0)} = \mathbf{P}, \tag{4.25}$$

   which implies $\mathbf{T}^{(0)} = \mathbf{T}$ and

$$\mathbf{G}^{(0)} = \mathbf{P}\begin{bmatrix}\mathbf{I} & \mathbf{T}^T\end{bmatrix}^T = \mathbf{G}. \tag{4.26}$$

   The iterative optimization process consequently starts with the code generator matrix $\mathbf{G}$ of the systematically encoded UW-OFDM concept, which can definitely be assumed to be a good initial starting point. The resulting optimum code generator matrix (found after convergence of the algorithm) is denoted with $\mathbf{G}'$.

2. For the second initialization each element of $\mathbf{A}^{(0)}$ is a realization of a Gaussian random variable with zero mean and variance one, i.e.,

$$\left[\mathbf{A}^{(0)}\right]_{ij} \sim \mathcal{N}(0,1). \tag{4.27}$$

   The resulting code generator matrix is labeled as $\mathbf{G}''$.

For both J_{BLUE} and J_{LMMSE}, the steepest descent algorithm converges at least one order of magnitude faster when initialized with (4.25) instead of (4.27). For the random initialization approach the resulting code generator matrix generally varies from trial to trial.

The columns of $\mathbf{G}'$ (Fig. 4.2) reveal that the energy of one data symbol is mainly (however, not exclusively) spread locally. $\mathbf{G}'$ can actually be regarded as the natural perfecting of $\mathbf{G}$. In fact, it also possesses the same symmetry property as $\mathbf{G}$ with

$$\mathbf{G}' = [\mathbf{g}'_0 \cdots \mathbf{g}'_{N_d/2-1} \operatorname{flip}\{(\mathbf{g}'_{N_d/2-1})^*\} \cdots \operatorname{flip}\{(\mathbf{g}'_0)^*\}]. \tag{4.28}$$

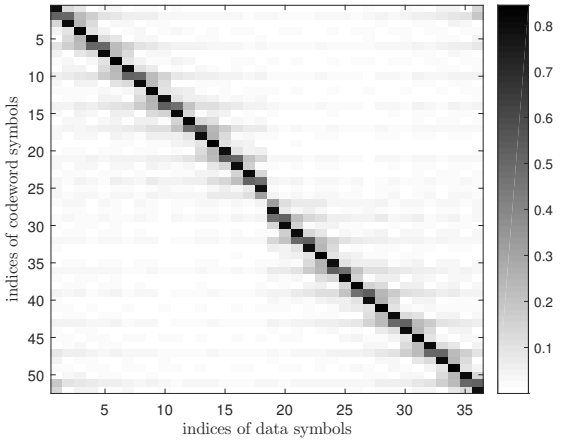


Figure 4.2.: $|\mathbf{G}'|$, magnitude of entries of generator matrix $\mathbf{G}'$.

Properties drawn from single subcarrier symbols. Furthermore, studying the rows corresponding to code word symbols c_i (from $\mathbf{c} = \mathbf{G}'\mathbf{d}$) that have originally been dedicated data symbols in the systematic code ($\mathbf{G}$), reveals that the weighted sum of N_d iid data symbols yielding a particular c_i is dominated by one particular data symbol (see the black entries of the matrix plot in Fig. 4.2). A condition for the CLT to hold is that no single term in the sum dominates [55], which is clearly violated in these cases. This is exemplarily demonstrated in the middle plot of Fig. 4.3, where the PDF of the real part of c_0 – estimated by a distribution of the relative frequencies – is not Gaussian anymore, but instead shows a bimodal behavior. In contrast, the CLT condition is fulfilled for (the real and imaginary parts of) all code word symbols c_i that originally corresponded to redundant symbols in the

systematically encoded case. Here, no single term in the weighted sum dominates. The PDFs of these c_i's may be well approximated by a complex Gaussian PDF $\mathcal{CN}(0, \sigma_{c_i}^2)$ with $\sigma_{c_i}^2 = \left[\mathrm{E}\left\{\mathbf{c}\mathbf{c}^H\right\}\right]_{ii} = [\mathbf{C}_{cc}]_{ii}$, whereas the real and imaginary parts of c_i can be approximated by a PDF of the form $\mathcal{N}(0, \sigma_{c_i}^2/2)$. Exemplarily,

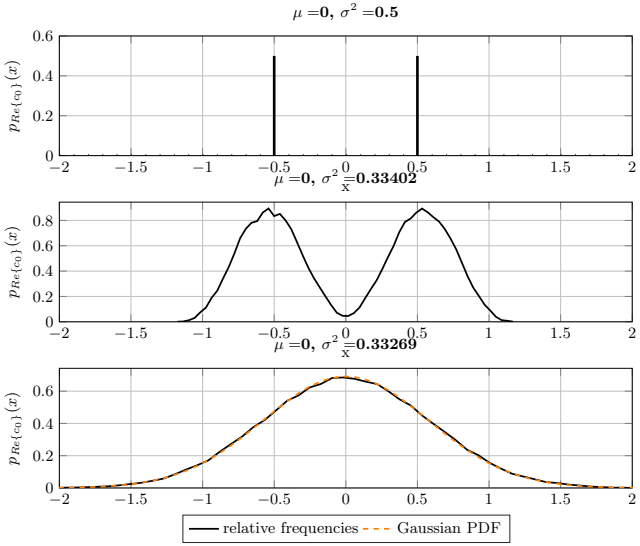


Figure 4.3.: Distribution of relative frequencies to estimate the PMF/PDFs of the real part of c_0 for $\mathbf{G}$, $\mathbf{G}'$, and $\mathbf{G}''$, respectively. A fitted Gaussian PDF serves as reference in the lower plot.

the middle plot of Fig. 4.4 provides an estimated PDF of the real part of c_1 by displaying a distribution of relative frequencies, and compares it with a PDF drawn from a fitted Gaussian distribution.

The situation is different for $\mathbf{G}''$ (see Fig. 4.5), since $\mathbf{G}''$ spreads the energy of each data symbol over the whole code word $\mathbf{c}$. The CLT holds for (the real and imaginary parts of) all code word symbols c_i, consequently the PDF of each c_i is approximated well by a complex Gaussian PDF $\mathcal{CN}(0, \sigma_{c_i}^2)$ with $\sigma_{c_i}^2 = [\mathbf{C}_{cc}]_{ii}$ (see the lower parts of Fig. 4.4 and Fig. 4.3).

4.2. Optimization Algorithm

In this section, a realization of the steepest descent algorithm is presented. This realization allows for numerically finding generator matrices that implement a non-

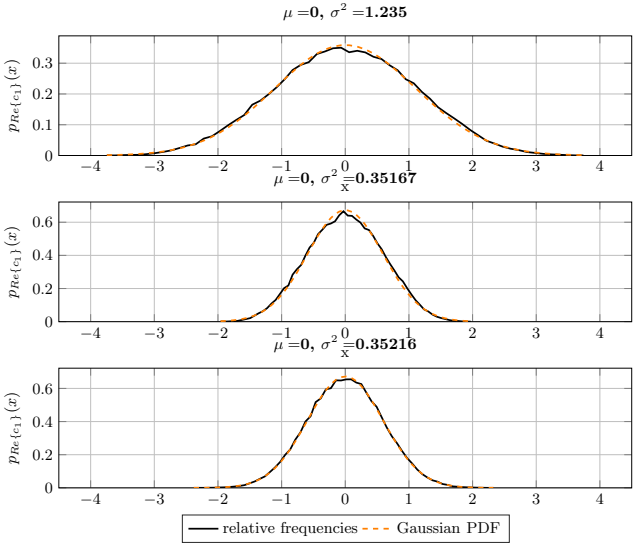


Figure 4.4.: Distribution of relative frequencies to estimate the PDFs of the real part of c_1 for $\mathbf{G}$, $\mathbf{G}'$, and $\mathbf{G}''$, respectively. A fitted Gaussian PDF serves as reference.

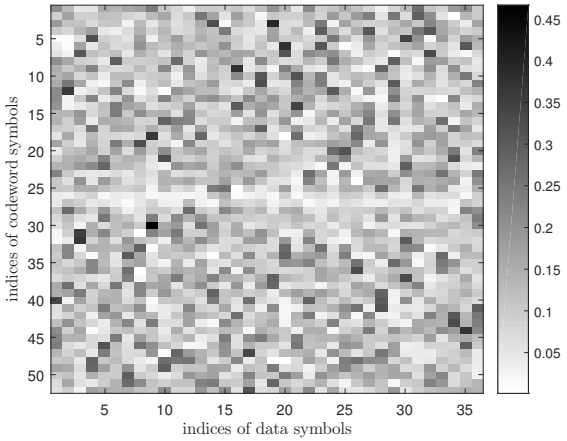


Figure 4.5.: $|\mathbf{G}''|$, magnitude of entries of generator matrix $\mathbf{G}''$.

systematically encoded UW-OFDM system. More in detail, the presented algorithm solves the optimization problem in (4.11). Any cost function J can be utilized within this algorithm, in this work, however, J is chosen out of $\{J_{\text{BLUE}}, J_{\text{LMMSE}}\}$. The found generator matrices are optimum w.r.t. these cost functions and fulfill the properties analytically derived in Sec. 4.1.

Note that a steepest descent algorithm is in general only able to find a local minimum, but it is not clear, whether the local minimum coincides with the global minimum. However, the analytical derivations in Sec. 4.1 proved the found generator matrices to be globally optimum as well.

Algorithm 3 `steepest_descent_algorithm` starts with initializing $\mathbf{A}$. Note that the approach from (2.42) of choosing $\mathbf{G} = \mathbf{A}\left[\begin{smallmatrix}\mathbf{I}\\ \mathbf{T}\end{smallmatrix}\right]$ is essential for successfully applying this optimization algorithm, as it can only deal with unconstrained optimization problems. For the exemplary generator matrices $\mathbf{G}'$ and $\mathbf{G}''$ considered in this work, the specific initialization approaches $\mathbf{A}^{(0)} = \mathbf{P}$ and $[\mathbf{A}^{(0)}]_{ij} \sim \mathcal{N}(0,1)$ have been used, respectively. Based on the resulting generator matrix, an initial value for J is calculated. Next, the function `calculate_gradient` is called to evaluate the gradient matrix $\frac{\partial J}{\partial \mathbf{A}} \in \mathbb{C}^{(N_d+N_r)\times(N_d+N_r)}$ defined as

$$\frac{\partial J}{\partial \mathbf{A}} = \begin{bmatrix} \frac{\partial J}{\partial [\mathbf{A}]_{11}} & \frac{\partial J}{\partial [\mathbf{A}]_{12}} & \cdots & \frac{\partial J}{\partial [\mathbf{A}]_{1(N_d+N_r)}} \\ \frac{\partial J}{\partial [\mathbf{A}]_{21}} & \frac{\partial J}{\partial [\mathbf{A}]_{22}} & \cdots & \frac{\partial J}{\partial [\mathbf{A}]_{2(N_d+N_r)}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial J}{\partial [\mathbf{A}]_{(N_d+N_r)1}} & \frac{\partial J}{\partial [\mathbf{A}]_{(N_d+N_r)2}} & \cdots & \frac{\partial J}{\partial [\mathbf{A}]_{(N_d+N_r)(N_d+N_r)}} \end{bmatrix}. \tag{4.29}$$

The partial derivatives $\partial J / \partial [A]_{ij}$ are approximated by

$$\frac{\partial J}{\partial [A]_{ij}} \approx \frac{J([A]_{ij} + \epsilon_1) - J([A]_{ij} - \epsilon_1)}{2\epsilon_1}, \tag{4.30}$$

with a very small ϵ_1.

After obtaining $\mathbf{J}_{\text{grad}}$, a new instance of $\mathbf{A}$ is derived by going along the gradient for a distance of ϵ_2. Subsequently, $\mathbf{G}$ and then J are updated. In case the new value of J is higher than the initial value, the distance ϵ_2 is cut in half. This is repeatedly executed until a smaller value for J is obtained. In case the new small value of J is below the previous one by at least ϵ_3, the optimization process continues. In the other case, it is assumed that the (local) minimum is already found and the optimization procedure stops. The values for ϵ_1, ϵ_2 and ϵ_3 are subject to a tradeoff between accuracy of results and simulation time. In this work, the values have been set to $\epsilon_1 = 10^{-7}$, $\epsilon_2 = 2 \cdot 10^{-2}$ and $\epsilon_3 = 10^{-6}$.

The solution of the optimization problem in (4.11) is ambiguous, which in turn allows to choose those generator matrices that meet other criteria as well, e.g., peak to power ratio (PAPR) or spectral shaping requirements.

Algorithm 3 `steepest_descent_algorithm`

initialize $\mathbf{A}$, ϵ_1, ϵ_2, ϵ_3
calculate $\mathbf{T}$ and $\mathbf{G}$ using $\mathbf{A}$ (cf. (2.44), (2.47) and (2.42))
calculate cost function J
$J_{\text{old}} \leftarrow J$
stop $\leftarrow$ false
while not stop **do**
    $(\mathbf{J}_{\text{grad}}) \leftarrow$ CALCULATE_GRADIENT$(\mathbf{A}, \epsilon_1)$
    $\mathbf{A}_{\text{new}} \leftarrow \mathbf{A} - \epsilon_2 \frac{\mathbf{J}_{\text{grad}}}{\left\|\mathbf{J}_{\text{grad}}\right\|_2}$
    update $\mathbf{G}, J$
    while $J > J_{\text{old}}$ **do**
        $\epsilon_2 \leftarrow \frac{\epsilon_2}{2}$
        $\mathbf{A}_{\text{new}} \leftarrow \mathbf{A} - \epsilon_2 \frac{\mathbf{J}_{\text{grad}}}{\left\|\mathbf{J}_{\text{grad}}\right\|_2}$
        update $\mathbf{G}, J$
    end while
    reinitialize ϵ_2
    if $J < J_{\text{old}} - \epsilon_3$ **then**
        $J_{\text{old}} \leftarrow J$
        $\mathbf{A} \leftarrow \mathbf{A}_{\text{new}}$
    else
        stop $\leftarrow$ true
    end if
end while
$J_{\text{opt}} \leftarrow J_{\text{old}}$
$\mathbf{A}_{\text{opt}} \leftarrow \mathbf{A}$

Algorithm 4 `calculate_gradient`

function CALCULATE_GRADIENT$(\mathbf{A}, \epsilon_1)$
    for $k = 0, 1, ..., (N_d + N_r - 1)$ **do**
        for $l = 0, 1, ..., (N_d + N_r - 1)$ **do**
            $\mathbf{A}[k, l] \leftarrow \mathbf{A}[k, l] + \epsilon_1$
            calculate $\mathbf{G}, J$
            $J' \leftarrow J$
            $\mathbf{A}[k, l] \leftarrow \mathbf{A}[k, l] - 2\epsilon_1$
            update $\mathbf{G}, J$
            $\mathbf{A}[k, l] \leftarrow \mathbf{A}[k, l] + \epsilon_1$
            $\mathbf{J}_{\text{grad}}[k, l] \leftarrow \frac{J' - J}{2\epsilon_1}$
        end for
    end for
    return $(\mathbf{J}_{\text{grad}})$
end function

4.3. Performance Evaluation

The following section evaluates the performance of non-systematically encoded UW-OFDM in terms of spectral properties and BER behavior. CP-OFDM will again serve as reference system. Since the non-systematic approach shall annihilate the flaws of the systematic concept, the latter is also considered for comparison.

4.3.1. Power Spectral Density

The Welch power spectral densities (PSDs) presented in the following are based on simulated UW-OFDM and CP-OFDM bursts, whereas the parameters have been chosen identically to the PSD analysis in Sec. 3.5.2. A burst is composed of a preamble (in all cases the IEEE 802.11a preamble [36]) and a data part comprising 8000 information bits. An outer channel code with coding rate $r = 1/2$ is employed. Note that no additional filters are applied for spectral shaping. For a better comparison, the PSDs have been normalized such that the passband of each spectral mask is centered at 0 dB. Fig. 4.6 highlights the significantly better sidelobe suppression of UW-OFDM compared to the CP-OFDM spectrum. The

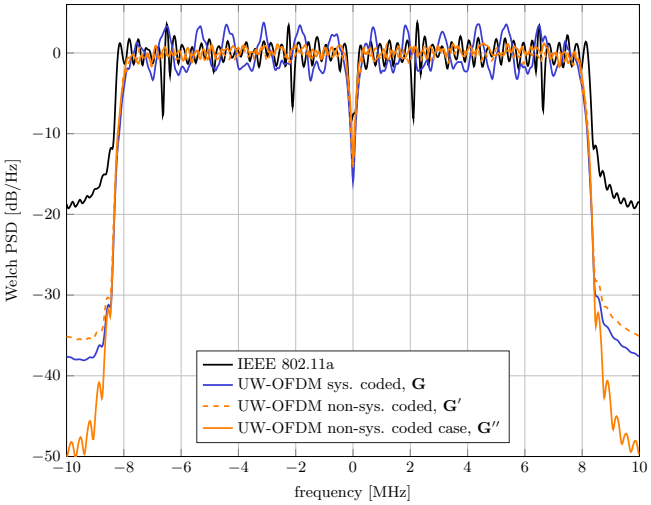


Figure 4.6.: Comparison of normalized Welch power spectral densities: CP-OFDM against UW-OFDM with $\mathbf{G}$, $\mathbf{G}'$ and $\mathbf{G}''$.

out-of-band emissions generated by $\mathbf{G}$ and $\mathbf{G}'$ are more than 15 dB below the

emissions of the CP-OFDM system. The emissions are even notably lower for $\mathbf{G}''$. These properties make UW-OFDM also an attractive candidate for cognitive radio networks [77].

Note that the spectra for $\mathbf{G}'$ and $\mathbf{G}''$ feature an extremely flat passband region compared to systematically encoded UW-OFDM with $\mathbf{G}$. This can be explained by the fact, that for the latter the mean power strongly varies between data and redundant subcarriers (Fig. 3.4), while all subcarriers (except the ones at the band edges) show almost equal power in the non-systematic case (Fig. 4.1).

4.3.2. Bit Error Ratio Simulations

Clearly, OFDM is designed for data transmission in dispersive channels, e.g., transmitting in a frequency selective environment. Nevertheless, the BER investigations start with results in AWGN, since the non-systematic code generator matrices are optimized for that particular case. In Fig. 4.7 the BER behavior of the IEEE 802.11a CP-OFDM based standard, and of both, the systematically encoded ($\mathbf{G}$) and the non-systematically encoded UW-OFDM approach are compared (note that $\mathbf{G}'$ and $\mathbf{G}''$ perform identical in AWGN). No outer code is used for these simulations. Results are provided for a BLUE and an LMMSE data estimator. As

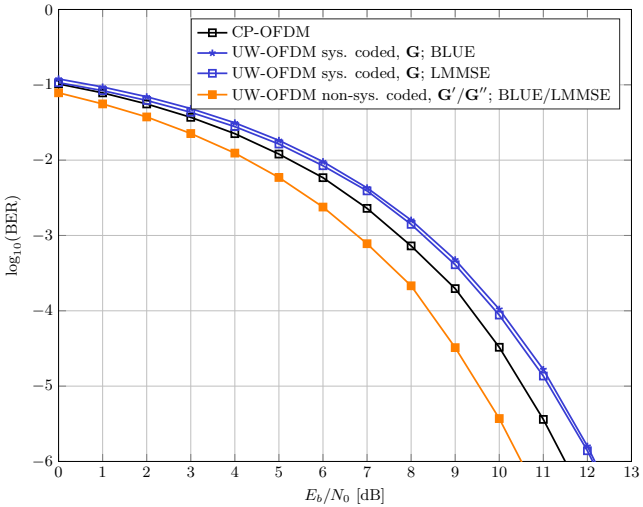


Figure 4.7.: BER comparison of UW-OFDM ($\mathbf{G}$, $\mathbf{G}'$ and $\mathbf{G}''$) and CP-OFDM for uncoded transmission in AWGN.

expected, the LMMSE estimator slightly outperforms the BLUE in case of systematically encoded UW-OFDM in the low E_b/N_0 region. For the non-systematically encoded approach though, both estimators perform completely identical. This is somewhat surprising, since the error variances after data estimation are not identical, cf. (4.15) and (4.24). However, it turns out that in the AWGN case the QPSK symbol estimates of the BLUE and of the LMMSE data estimator always lie in the same decision region of the constellation diagram (Fig. 4.8), and the difference in the error variances does not translate into a difference in the BER performance [78]. This originates from the LMMSE estimator (2.83) corresponding to a scaled version of the BLUE (2.75)

$$\hat{\mathbf{d}}_{\mathrm{BLUE}} = \mathbf{G}^H \tilde{\mathbf{y}} \tag{4.31}$$

$$\hat{\mathbf{d}}_{\mathrm{LMMSE}} = \frac{\sigma_d^2}{\sigma_d^2 + N\sigma_n^2} \mathbf{G}^H \tilde{\mathbf{y}} = \alpha_1 \hat{\mathbf{d}}_{\mathrm{BLUE}} \tag{4.32}$$

in case of $\mathbf{H} = \mathbf{I}$, $\mathbf{G}^H\mathbf{G} = \alpha\mathbf{I}$ and with $\alpha_1 < 1$. As a side effect, this example additionally illustrates the biasedness property of an LMMSE estimator discussed in Sec. 2.3.2. Systematically encoded UW-OFDM performs slightly worse compared

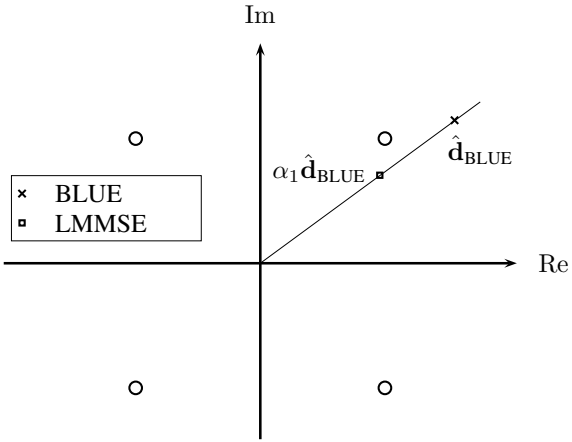


Figure 4.8.: Relationships between the BLUE and LMMSE estimates in the AWGN channel for non-systematically encoded UW-OFDM.

to the CP-OFDM reference system, the non-systematically encoded UW-OFDM system tops CP-OFDM by 1 dB and the systematic approach (with LMMSE data estimation) by 1.6 dB, confirming a remarkable performance enhancement by the non-systematic concept.

The remaining part of this section concentrates on results in a frequency selective environment, which is a more realistic field of application for OFDM than AWGN.

Analogously to Sec. 3.5.3, simulation results are obtained by averaging over the same fixed set of 10^4 CIR realizations. These CIR realizations were derived from the channel model in Sec. 2.4.2, with the channel delay spread τ_{RMS} as input parameter.

In the following, all considerations of the non-systematic approach focus on $\mathbf{G}'$ only, results for $\mathbf{G}''$ are provided at the end of this section. Fig. 4.9 illustrates the BER performance in case no outer channel code is applied. The simulations

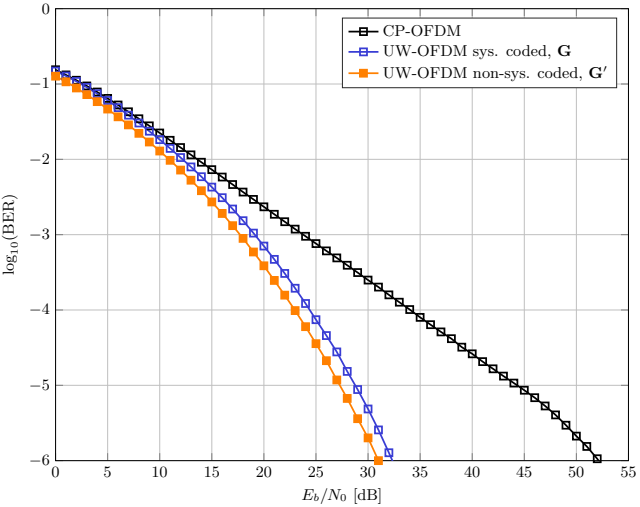


Figure 4.9.: BER comparison of UW-OFDM ($\mathbf{G}$ and $\mathbf{G}'$) and CP-OFDM for uncoded transmission in a multipath environment with $\tau_{\mathrm{RMS}} = 100\,\mathrm{ns}$.

reveal that already the *systematically* encoded approach provides a diversity gain over the CP-OFDM concept. This diversity gain translates to different slopes of the BER curves. *Non-systematically* encoded UW-OFDM delivers an additional coding gain of 1.3 dB (visible as a horizontal shift of the BER curve to the left), ending up in a total gain of 21.2 dB over CP-OFDM. Fig. 4.10 displays the results in case an additional outer channel code is applied. For $r = 3/4$, the non-systematic approach provides an additional gain of 1.3 dB over the systematic one, leading to an outperformance of 2.4 dB over CP-OFDM. Interestingly, the whole gain of 1.3 dB between both UW-OFDM approaches obtained in the uncoded case is also transferred to the coded scenario. In case of $r = 1/2$ the gain increases by 1.0 dB to a total of 1.5 dB over CP-OFDM.

To provide results for higher order modulation as well, Fig. 4.11 plots the uncoded ($r = 1$) and coded ($r = 1/2$ and $r = 3/4$) performance in case of 16-QAM. The

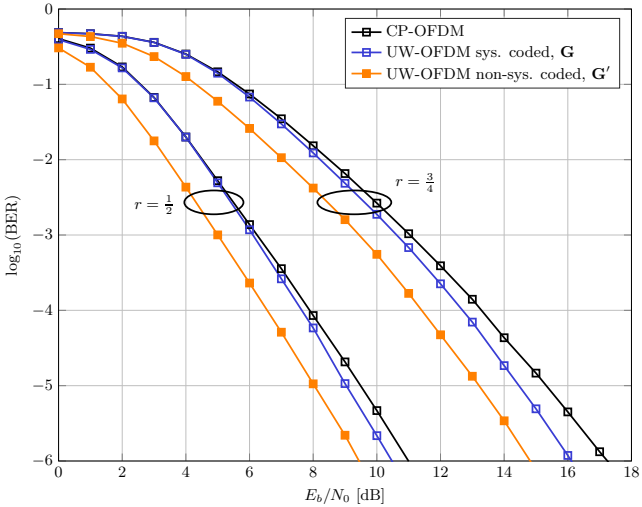


Figure 4.10.: BER comparison of UW-OFDM ($\mathbf{G}$ and $\mathbf{G}'$) and CP-OFDM for coded transmission ($r = 1/2$ and $r = 3/4$) in a multipath environment with $\tau_{\mathrm{RMS}} = 100\,\mathrm{ns}$.

same tendencies as for QPSK can be observed. UW-OFDM surpasses CP-OFDM by 1.6 dB for $r = 1/2$ and 1.7 dB in case of $r = 3/4$.

The next part provides a performance comparison between the two exemplary derived code generator matrices $\mathbf{G}'$ and $\mathbf{G}''$. Clearly, both present exactly the same performance in AWGN, since every optimum code generator matrix features the same error covariance matrix (4.24) in this scenario. Nevertheless, $\mathbf{G}'$ and $\mathbf{G}''$ show quite a different behavior in dispersive channels such as frequency selective channels. Sec. 4.1 already discussed the different structures of $\mathbf{G}'$ and the particularly chosen $\mathbf{G}''$. As a reminder, $\mathbf{G}'$ spreads the energy of one data symbol mainly locally (see Fig. 4.2) and can be regarded as the natural perfecting of $\mathbf{G}$. In contrast, $\mathbf{G}''$ displayed in Fig. 4.5 spreads the energy of each data symbol approximately uniformly over the code word $\mathbf{c}$ and thus over the whole bandwidth. From this point of view a system with $\mathbf{G}''$ behaves comparable to a single-carrier system, where the energy of each individual data symbol is distributed uniformly over the whole bandwidth. A system with $\mathbf{G}'$, however, rather shares similarities with classical OFDM, where a subcarrier is loaded with exactly one data symbol. This condition does not hold true for $\mathbf{G}'$, but a subcarrier is still dominated by a single data symbol.

Fig. 4.12 shows that $\mathbf{G}''$ features extremely good results without an outer code,

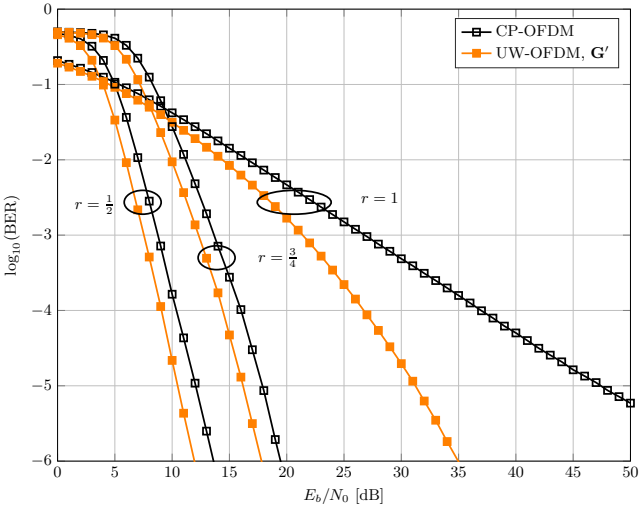


Figure 4.11.: BER comparison of UW-OFDM with $\mathbf{G}'$ and CP-OFDM for uncoded ($r = 1$) and coded transmission ($r = 1/2$ and $r = 3/4$) in a multi-path environment with $\tau_{\mathrm{RMS}} = 100\,\mathrm{ns}$ and 16-QAM as modulation alphabet.

exceeding $\mathbf{G}'$ significantly in that case. For $r = 3/4$, $\mathbf{G}''$ is still slightly better than $\mathbf{G}'$, while for $r = 1/2$, $\mathbf{G}'$ clearly outperforms $\mathbf{G}''$. The coding gain achieved by a strong outer code in a frequency selective channel is high for $\mathbf{G}'$ as it might be expected for a system rather related to classical OFDM, while it is comparably low for $\mathbf{G}''$ with its precoding properties making the transceiver rather behave like a single-carrier system. In other words, a severe fading hole will most likely erase one data symbol completely in case of $\mathbf{G}'$. However, channel codes normally handle single errors well and are able to recover those data symbols. In case of $\mathbf{G}''$ and its almost uniform spreading of the data symbols over the whole bandwidth, all are degraded by a certain extent, which usually determines a less optimum prerequisite for channel codes. From an information theory point of view, this behavior can also be considered as a consequence of memory increasing mutual information [79].

Furthermore, Fig. 4.13 extends the analysis to BER performance results in channels exceeding the guard interval length. This is clearly given here with a channel delay spread of $\tau_{\mathrm{RMS}} = 200\,\mathrm{ns}$, thus explaining the emerging error floor in uncoded transmission. In coded transmission, the best UW-OFDM system outplays CP-OFDM by 1.4 dB for $r = 1/2$ ($\mathbf{G}'$) and 3.5 dB for $r = 3/4$ ($\mathbf{G}''$). Note that in general, it is much easier in UW-OFDM to adapt to different channel conditions than in

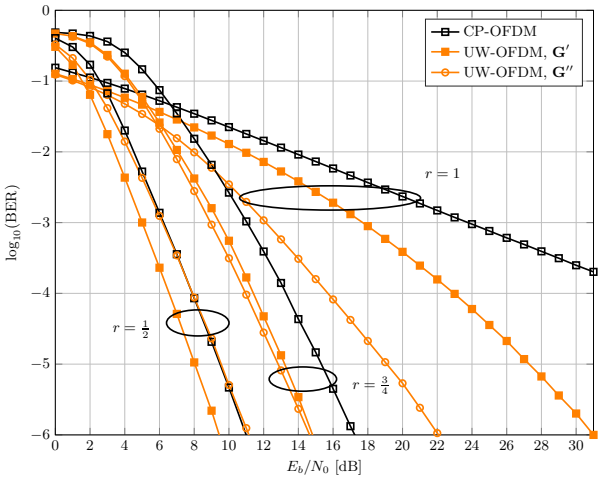


Figure 4.12.: BER comparison of UW-OFDM between $\mathbf{G}'$ and $\mathbf{G}''$ for uncoded ($r = 1$) and coded transmission ($r = 1/2$ and $r = 3/4$) in a multipath environment with $\tau_{\mathrm{RMS}} = 100\,\mathrm{ns}$.

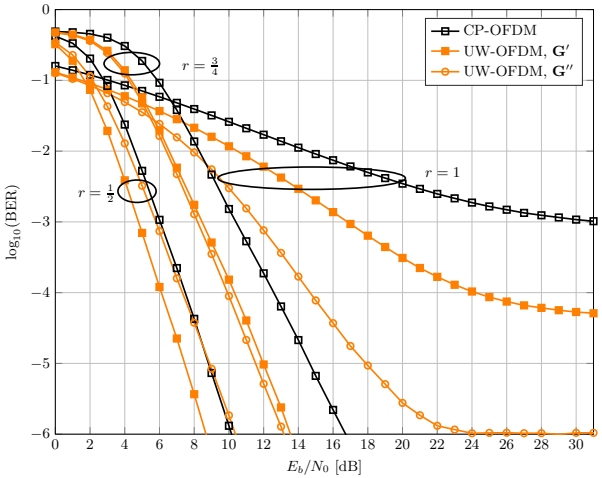


Figure 4.13.: BER comparison of UW-OFDM ($\mathbf{G}'$ and $\mathbf{G}''$) and CP-OFDM for uncoded ($r = 1$) and coded transmission ($r = 1/2$ and $r = 3/4$) in a multipath environment with $\tau_{\mathrm{RMS}} = 200\,\mathrm{ns}$.

any other signaling schemes. Based on the dispersity properties of the channel, the guard interval length N_u can be adapted without changing the DFT length and thus the framing structure. Especially w.r.t. providing flexible communication setups, this is a strong argument for UW-OFDM.

There are several insights that can be derived from this section:

- UW-OFDM in general offers a superior sidelobe suppression compared to CP-OFDM. It depends on the specific instance of the generator matrix, if the non-systematic or the systematic approach provides better results. However, these differences are on a high level anyway. In the passband though, non-systematically encoded UW-OFDM always outperforms the systematic concept due to its relatively flat spectrum.
- In terms of BER performance, non-systematically encoded UW-OFDM provides an additional coding gain over the systematic approach.
- In contrast to the introduction of additional redundant subcarriers or systematic noise, cf. Sec. 3.6 and 3.7 , the presented concept offers a performance gain without sacrificing bandwidth efficiency or introducing an error floor.
- There are infinitely many different generator matrices that meet the conditions of the non-systematically encoding idea. The matrix can thus be optimized towards a specific setup, e.g., the outer channel encoder, to achieve the best BER performance.
- Although not explicitly shown, the degree of freedom offered in designing the generator matrix can also be utilized to optimize towards any other criteria. Reasonable criteria might be the PAPR or peak to minimum power ratio (PMR) performance [56, 80], dedicated spectral properties or particular channel instances [74, 75].

In conclusion, non-systematically encoded UW-OFDM is the most beneficial concept out of all presented ones to implement the idea of UW-OFDM.

4.3.3. Impact of Setup and UW

As already stressed several times within this work, CP-OFDM is currently the most popular multi-carrier technique and thus denotes the gold standard test for all other signaling schemes. When evaluating two different concepts against each other, there are always several ways of how to compare them. For most of the simulations so far, the focus has been laid on practical scenarios and thus a CP-OFDM setup based on the parameters of the IEEE 802.11a standard has been considered [36]. For a fair comparison it is crucial to choose a UW-OFDM setup as similar as possible to the reference CP-OFDM setup. The utilized UW-OFDM setup A shares thus the same

- DFT length,
- sampling frequency,
- number of subcarriers,
- subcarrier spacing, and also the same
- guard interval length.

Nevertheless, both setups differ in the

- OFDM symbol duration and the
- number of data subcarriers per OFDM symbol.

The first follows from the guard interval already being part of the DFT output, the latter is due to the necessity of redundancy in UW-OFDM to generate the desired UW in the time domain. The CP-OFDM system based on setup A provides thus a 6.67 % higher theoretical data rate (additional overhead like e.g., a preamble for estimation and synchronization tasks are not considered here) than the corresponding UW-OFDM system. The UW-OFDM achieves a lower data rate on one hand, but it also requires less bandwidth on the other hand, cf. Fig. 4.6, both systems show thus almost the same bandwidth efficiency in bits/s/Hz. Hence, this seems to be a reasonable and fair comparison considering that bandwidth is an expensive resource, especially w.r.t. real-world communication scenarios.

In this section, setup C (see Tab. 2.2 and 2.3 in Sec. 2.4) is used now for both concepts, leading to the same theoretical data rate of the CP-OFDM and UW-OFDM system. For a DFT size of $N = 64$, CP-OFDM allows to transmit 64 data symbols within an OFDM symbol duration of $4\,\mu$s. In order to achieve the same data rate for the UW-OFDM concept, the DFT size is enlarged to $N = 80$ and the DFT period (and thus also the OFDM symbol duration) to $4\,\mu$s. Hence, both systems provide the same theoretical data rate, but now differ in the

- DFT size,
- DFT period and consequently also in
- subcarrier spacing.

This section will demonstrate that UW-OFDM surpasses CP-OFDM regardless of the particular configuration. The considerations start with uncoded transmission in the AWGN channel, followed by an analysis of multipath communications. Furthermore, the effect of different UW configurations is investigated.

Applying the linear model in (2.68) to CP-OFDM, i.e., $\mathbf{G} = \mathbf{I}$, and taking into account the AWGN condition $\tilde{\mathbf{H}} = \mathbf{I}$ yields

$$\tilde{\mathbf{y}} = \mathbf{d} + \mathbf{v}, \tag{4.33}$$

with the noise vector $\mathbf{v} \sim \mathcal{CN}(\mathbf{0}, \sigma_v^2\mathbf{I})$ and $\sigma_v^2 = N\sigma_n^2$. Eq. (4.33) denotes the final model for data estimation, possible zero or pilot subcarriers are consequently already excluded. Restricting the estimator to be linear, $\hat{\mathbf{d}} = \tilde{\mathbf{y}}$ determines already the best estimate. The error $\mathbf{e} = \hat{\mathbf{d}} - \mathbf{d}$ can be characterized by the error covariance matrix

$$\mathbf{C}_{ee,\mathrm{cp}} = \mathrm{E}\left\{\mathbf{e}\mathbf{e}^H\right\} = \sigma_v^2\mathbf{I}. \tag{4.34}$$

The error covariance matrix serves later as a means to assess the BER performance. As a first step towards this assessment, the average energy of a CP-OFDM symbol calculates as

$$E_{x,\mathrm{cp}} = (E'_d + E'_p)\frac{(N' + N'_g)}{N'} = \left(\frac{\sigma_d^2 N'_d}{N'} + E'_p\right)\frac{N' + N'_g}{N'}, \tag{4.35}$$

where N' determines the DFT size, N'_g the guard interval size, E'_d the mean energy of the data symbols and E'_p the mean energy of the pilot symbols. Note that in this context, the notation ' serves as a means to distinguish CP-OFDM variables from UW-OFDM ones. With (4.35) and assuming now any given E_b/N_0 value, $\sigma_{v,\mathrm{cp}}^2$ follows as

$$\begin{aligned}
\sigma_{v,\mathrm{cp}}^2 &= N'\sigma_{n,\mathrm{cp}}^2 && (4.36)\\
&= N'\frac{E_{x,\mathrm{cp}}}{2b\frac{E_b}{N_0}} && (4.37)\\
&= N'\frac{1}{2qN'_d\frac{E_b}{N_0}}\frac{(\sigma_d^2 N'_d + N'E'_p)(N' + N'_g)}{N'^2} && (4.38)\\
&= \frac{1}{2q\frac{E_b}{N_0}}\left(\frac{\sigma_d^2 N'_d(N' + N'_g)}{N'N'_d} + \frac{N'E'_p(N' + N'_g)}{N'N'_d}\right) && (4.39)\\
&= \frac{1}{2q\frac{E_b}{N_0}}\left(\sigma_d^2\left(1 + \frac{N'_g}{N'}\right) + E'_p\frac{N' + N'_g}{N'_d}\right). && (4.40)
\end{aligned}$$

Here, $b = qN'_d$ denotes the number of bits per OFDM symbol and q the number of bits per symbol of the modulation alphabet, e.g., $q = 2$ in case of QPSK. It is immediately apparent from (4.40) that for $E'_p = 0$, both used CP-OFDM setups A and B (see Tab. 2.3) have the same error covariance matrix and thus BER performance in AWGN.

According to the derivations in Sec. 4.1, the mean symbol energy of a non-systematically encoded UW-OFDM symbol with zero word is

$$E_x = \frac{\sigma_d^2 \mathrm{tr}\left(\mathbf{G}^H\mathbf{G}\right)}{N} = \frac{\sigma_d^2 N_d}{N}. \tag{4.41}$$

Adding a UW yields

$$E_{x'} = E_x + E_{x_u} = E_x(1 + \beta). \tag{4.42}$$

The factor β depends on the amount of energy spent for the UW and is detailed later in this section. With (2.85), (4.41), (4.42) and a fixed E_b/N_0 value, the noise variance is given by

$$\begin{aligned}\sigma_{v,\mathrm{uw}}^2 &= N\sigma_{n,\mathrm{uw}}^2 && (4.43)\\ &= N\frac{1}{2qN_d\frac{E_b}{N_0}}E_x\,(1+\beta) && (4.44)\\ &= \frac{1}{2q\frac{E_b}{N_0}}\frac{N}{N_d}\frac{\sigma_d^2 N_d}{N}\,(1+\beta) && (4.45)\\ &= \frac{1}{2q\frac{E_b}{N_0}}\sigma_d^2\,(1+\beta)\,. && (4.46)\end{aligned}$$

In (4.15) and (4.24) the error covariance matrices for a BLUE and an LMMSE estimator are derived. Since the LMMSE estimator collapses into the BLUE for high E_b/N_0 values anyway, and for reasons of simplicity, the focus is on the error covariance matrix of the BLUE defined as

$$\mathbf{C}_{ee,\mathrm{uw}} = \sigma_{v,\mathrm{uw}}^2\mathbf{I}, \tag{4.47}$$

derived from (4.6) with $\mathrm{tr}\left(\mathbf{G}^H\mathbf{G}\right) = N_d$, inserting into (4.15) and taking into account $N\sigma_n^2 = \sigma_v^2 \stackrel{!}{=} \sigma_{v,\mathrm{uw}}^2$. An unanswered question is still the energy E_{x_u} spent on the UW and its effect on the performance. In principle, there are three configurations suitable for comparison:

1. Zero word as UW. In this case no energy is spent on pilot symbols in UW-OFDM as well as in CP-OFDM. This is the default configuration in this work.

2. Same relative amount of energy for UW and pilot subcarriers. In this case the UW is scaled such that the ratio between UW energy and energy of the rest of the OFDM symbol is the same as the ratio between pilot and data symbols in CP-OFDM.

3. Same average power as the rest of the OFDM symbol. Here, the UW is scaled such that its samples experience the same average power as the remaining $N - N_u$ samples of the OFDM symbol.

The different values of β in (4.42) based on the chosen UW configuration are summarized in Tab. 4.1. The first and the last scaling factor are immediately

apparent, the second one follows from

$$E_{x'} = E_x + \frac{E'_p}{E'_d + E'_p} E_{x'} \tag{4.48}$$

$$E_{x'} \left(1 - \frac{E'_p}{E'_d + E'_p}\right) = E_x \tag{4.49}$$

$$E_{x'} = E_x \frac{1}{1 - \frac{E'_p}{E'_d + E'_p}} \tag{4.50}$$

$$= E_x \frac{1}{\frac{E'_d + E'_p - E'_p}{E'_d + E'_p}} \tag{4.51}$$

$$= E_x \frac{E'_d + E'_p}{E'_d} \tag{4.52}$$

$$= E_x \left(1 + \frac{E'_p}{E'_d}\right). \tag{4.53}$$

Table 4.1.: Summary of β values to determine the UW energy E_{x_u} in (4.42).

configuration	1	2	3
β	0	$\frac{E'_p}{E'_d}$	$\frac{N_u}{N-N_u}$
UW scaling	zero UW	partial scale	full scale

To conclude from $\sigma^2_{v,\mathrm{uw}}$ in (4.46) and Tab. 4.1, all UW-OFDM setups with a zero word provide the same AWGN performance, independent of the specific values for DFT size N, UW length N_u, number of data subcarriers N_d or redundancy N_r. The same holds true for UW-OFDM setups with a UW scaling based on the second configuration. Note that performance equality of the different setups does not hold in multipath communications. For the third configuration in Tab. 4.1, the specific setup influences the BER behavior in multipath as well as AWGN.

Since $\mathbf{C}_{ee,\mathrm{cp}}$ as well as $\mathbf{C}_{ee,\mathrm{uw}}$ correspond to scaled identity matrices, the BER performance gain of UW-OFDM over CP-OFDM in uncoded AWGN transmission can simply be calculated as

$$G = 10\log_{10}\left(\frac{\sigma^2_{v,\mathrm{cp}}}{\sigma^2_{v,\mathrm{uw}}}\right) \text{ dB}. \tag{4.54}$$

For the default configuration with zero UW and zero pilot symbols, it follows that $\beta = 0$ and $E_p' = 0$ and the gain gets

$$G = 10\log_{10}\left(1 + \frac{N_g'}{N'}\right) = 0.97\,\text{dB}. \tag{4.55}$$

Since in case of a zero UW $\sigma^2_{v,\text{uw}}$ does not depend on the parameters of the specific setup, the performance gain between UW-OFDM and CP-OFDM stays the same for setup A as well as setup C. The same gain of $G = 0.97\,\text{dB}$ as for the zero energy approach is achieved when evaluating (4.54) for the second UW scaling approach in Tab. 4.1, setup A and $E_p' = 4/N'$, the latter value originating from the pilot symbols $\{1, -1, 1, 1\}$ defined in [36]. Both cases are thus equivalent when considering the BER performance difference between UW-OFDM and CP-OFDM. The gain vanishes to 0.07 dB for a fully scaled UW (third column in

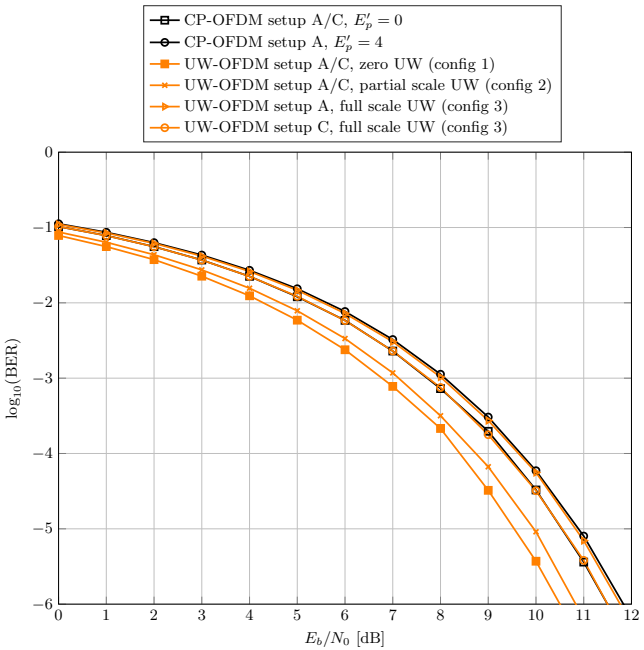


Figure 4.14.: BER comparison of UW-OFDM and CP-OFDM for different UW energy configurations according to Tab. 4.1 in case of uncoded transmission in AWGN, with setup A and C according to Tab. 2.2 and 2.3.

Tab. 4.1). CP-OFDM and UW-OFDM perform even identically in case of setup C

and a fully scaled UW. All results are also confirmed by means of simulations in Fig. 4.14.

The last two examples raise the impression that the performance gain of UW-OFDM only originates from an energy argument, and that there is no advantage over CP-OFDM without it. In this case, however, UW-OFDM would then offer the benefit of (more) pilot symbols and the spent energy E_{x_u} could serve as a means to balance the gain between pilot symbols and BER performance. Contrary to UW-OFDM, a CP-OFDM system does not offer the possibility to trade the energy spent for the cyclic prefix against an improved BER behavior. Furthermore and most important, the BER performance of UW-OFDM and CP-OFDM is only equivalent in case of AWGN, which is definitely not the scenario an OFDM signaling scheme is designed for.

Fig. 4.15 illustrates the uncoded performance in a multipath environment for setup C and zero UW/pilots. As expected, a similar gain as for setup A is achieved. Using a non-zero UW would now result in the same shift of the BER curve as in AWGN, however, the diversity order (visible as the slope of the BER curve) would stay unaffected.

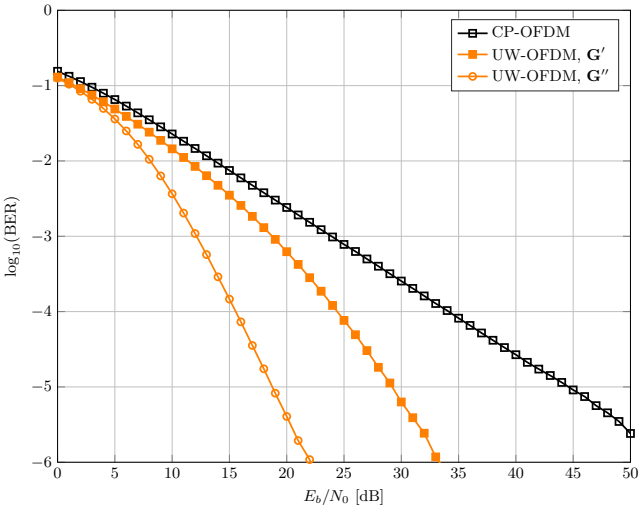


Figure 4.15.: BER comparison of UW-OFDM ($\mathbf{G}'$ and $\mathbf{G}''$) and CP-OFDM for uncoded transmission in a multipath environment with $\tau_{\mathrm{RMS}} = 100\,\mathrm{ns}$, setup C according to Tab. 2.2.

Fig. 4.16 compares the coded BER performance for setup C, again with zero UW/pilots. Note that for both (outer) coding rates, the two UW-OFDM ap-

proaches always beat the CP-OFDM system. Whereas the UW-OFDM system utilizing $\mathbf{G}''$ as code generator matrix achieves only a marginal gain of 0.1 dB in case of $r = 1/2$, this gain increases to 1.1 dB when $\mathbf{G}'$ is applied. For an (outer) coding rate of $r = 3/4$, both systems outperform the conventional CP-OFDM concept remarkably, namely by 3 dB and 2 dB, respectively. Using a non-zero UW would also lead to a shift of the BER curves. A general quantification of this shift is rather difficult for coded transmission, the remaining gain over CP-OFDM depends on the utilized (outer) channel code and its interaction with the general setup.

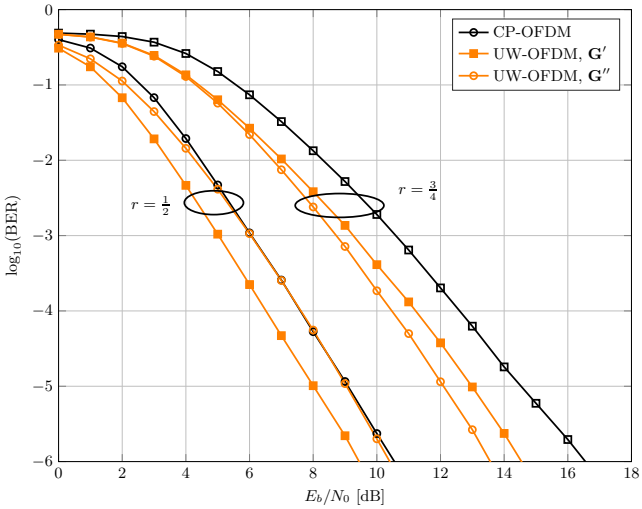


Figure 4.16.: BER comparison of UW-OFDM ($\mathbf{G}'$ and $\mathbf{G}''$) and CP-OFDM for coded transmission ($r = 1/2$ and $r = 3/4$) in a multipath environment with $\tau_{\mathrm{RMS}} = 100$ ns, setup C according to Tab. 2.2.

Fig. 4.17 compares the coded performance for setup C when transmitting over multipath channels featuring a channel delay spread of 200 ns. Again the same tendencies are observed. The UW-OFDM approaches surpass the CP-OFDM system by 0.2 and 1.3 dB in case of $r = 1/2$ and by 3.3 dB and 2.1 dB in case of $r = 3/4$.

4.3.4. Impact of Imperfect Channel Knowledge

The performance results presented so far assumed perfect channel knowledge at the receiver, hence the effect of channel estimation errors on the BER performance is in-

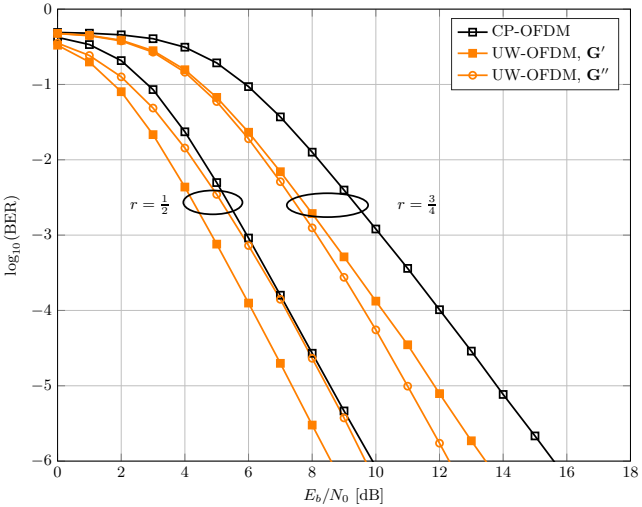


Figure 4.17.: BER comparison of UW-OFDM ($\mathbf{G}'$ and $\mathbf{G}''$) and CP-OFDM for coded transmission ($r = 1/2$ and $r = 3/4$) in a multipath environment with $\tau_{\mathrm{RMS}} = 200\,\mathrm{ns}$, setup C according to Tab. 2.2.

vestigated next. Since in UW-OFDM the channel $\tilde{\mathbf{H}}$ (and consequently its estimate $\hat{\tilde{\mathbf{H}}}$) is incorporated differently into the receiver processing than in CP-OFDM, it is not immediately obvious, whether channel estimation errors will degrade the BER performance in the same scale. One difference originates from the data estimators. In CP-OFDM the best linear estimator is simply given by the channel inverter $\mathbf{E} = \tilde{\mathbf{H}}^{-1}$, in UW-OFDM more advanced concepts like BLUE or LMMSE estimator can be applied. Furthermore, in UW-OFDM channel estimation errors have an additional impact, namely in the processing step where the influence of the UW is subtracted from the received symbol, cf. (2.66)-(2.68).

A preamble based channel estimation procedure described briefly in the following is applied to investigate the influence of channel estimation errors. The UWs are intentionally left out for this estimation task to ensure a fair comparison with the CP-OFDM reference system by using the same algorithm. The IEEE 802.11a preamble defined in [36] is utilized, which contains two identical BPSK (binary phase shift keying) modulated OFDM symbols (preceded by a guard interval) denoted by $\mathbf{x}_p = \mathbf{x}_{p_1} = \mathbf{x}_{p_2} \in \mathbb{C}^{N\times 1}$. Note that the downsized frequency domain version $\tilde{\mathbf{x}}_{p,d} = \mathbf{B}^T\mathbf{F}_N\mathbf{x}_p$ corresponds to $\tilde{\mathbf{x}}_{p,d} \in \{-1,1\}^{(N_d+N_r)\times 1}$. Let $\mathbf{y}_{p_1}$ and $\mathbf{y}_{p_2}$ be the received noisy preamble symbols in time domain, and let $\tilde{\mathbf{y}}_{p,d} = \mathbf{B}^T\mathbf{F}_N(\mathbf{y}_{p_1} + \mathbf{y}_{p_2})/2$. Then a first course unbiased estimate of the vector of

channel frequency response coefficients is obtained as

$$\hat{\tilde{\mathbf{h}}}_1[k] = \frac{\tilde{\tilde{\mathbf{y}}}_{p,d}[k]}{\tilde{\mathbf{x}}_{p,d}[k]} = \tilde{\tilde{\mathbf{y}}}_{p,d}[k]\tilde{\mathbf{x}}_{p,d}[k] \tag{4.56}$$

for $k = 0, ..., (N_d + N_r - 1)$. The latter step follows from $\tilde{\mathbf{x}}_{p,d}[k] \in \{-1, 1\}$. This course channel estimate can be significantly improved or rather noise reduced by exploiting the usually valid assumption of the channel impulse response not exceeding the guard duration N_u. With the vector of channel impulse response coefficients $\mathbf{h} \in \mathbb{C}^{N_u \times 1}$ and its zero padded version $\mathbf{h}_{zp} \in \mathbb{C}^{N \times 1}$, this assumption is incorporated by modelling the course channel estimate as

$$\begin{aligned} \hat{\tilde{\mathbf{h}}}_1 &= \mathbf{B}^T \mathbf{F}_N \mathbf{h}_{zp} + \mathbf{B}^T \mathbf{F}_N \mathbf{n} \\ &= \mathbf{B}^T \mathbf{F}_N \begin{bmatrix} \mathbf{h} \\ \mathbf{0} \end{bmatrix} + \mathbf{v}, \end{aligned} \tag{4.57}$$

where $\mathbf{v} \in \mathbb{C}^{(N_d+N_r)\times 1}$ represents a zero-mean Gaussian noise vector with $\mathbf{v} \sim \mathcal{CN}(\mathbf{0}, N\sigma_n^2\mathbf{I})$. Decomposing the DFT matrix into $\mathbf{F}_N = \begin{bmatrix} \mathbf{M}_1 & \mathbf{M}_2 \end{bmatrix}$ with $\mathbf{M}_1 \in \mathbb{C}^{N \times N_u}$ and $\mathbf{M}_2 \in \mathbb{C}^{N \times (N-N_u)}$, (4.57) is developed to

$$\hat{\tilde{\mathbf{h}}}_1 = \mathbf{B}^T \mathbf{M}_1 \mathbf{h} + \mathbf{v}. \tag{4.58}$$

Based on the linear model in (4.58) the MVU estimator [42] of the channel impulse response is given as

$$\hat{\mathbf{h}} = \left(\mathbf{M}_1^H \mathbf{B}\mathbf{B}^T \mathbf{M}_1 \right)^{-1} \mathbf{M}_1^H \mathbf{B} \hat{\tilde{\mathbf{h}}}_1. \tag{4.59}$$

Returning to frequency domain and again excluding the zero subcarriers from further operation delivers the final and highly noise reduced frequency domain channel estimate

$$\begin{aligned} \hat{\tilde{\mathbf{h}}}_2 &= \mathbf{B}^T \mathbf{F}_N \begin{bmatrix} \hat{\mathbf{h}} \\ \mathbf{0} \end{bmatrix} \\ &= \underbrace{\mathbf{B}^T \mathbf{M}_1 \left(\mathbf{M}_1^H \mathbf{B}\mathbf{B}^T \mathbf{M}_1 \right)^{-1} \mathbf{M}_1^H \mathbf{B}}_{\mathbf{W}} \hat{\tilde{\mathbf{h}}}_1. \end{aligned} \tag{4.60}$$

Note that the smoothing matrix $\mathbf{W} \in \mathbb{C}^{(N_d+N_r)\times(N_d+N_r)}$ does not depend on the channel and has to be calculated only once during system design. The preamble based estimate of the channel matrix is therefore given by $\hat{\tilde{\mathbf{H}}} = \mathrm{diag}\left(\hat{\tilde{\mathbf{h}}}_2\right)$.

Fig. 4.18 compares the performance loss of CP-OFDM and non-systematically encoded UW-OFDM in case of imperfect channel estimation based on (4.60). For the latter a CAZAC sequence [65] as a representative of a non-zero UW instead of the usual zero UW case has been applied to cover the effect of an imperfect subtraction of the UW offset as well. For reasons of fair comparison, the UW has been scaled

such that the ratio between UW energy and energy of the rest of the OFDM symbol is the same as the ratio between pilot and data symbols in CP-OFDM. Details can be found in Sec. 4.3.3, the corresponding scaling factor in Tab. 4.1 is labeled as 'partial scaling'. It turns out that an UW-OFDM system with $\mathbf{G}'$ loses about 0.7 dB independent of the outer coding rate, a system with $\mathbf{G}''$ approximately 1.2 dB for $r = 1/2$ and 0.7 dB for $r = 3/4$. For CP-OFDM and $r = 1/2$, the performance degrades about 1.2 dB and in case of $r = 3/4$ around 0.6 dB compared to perfect channel knowledge. The results confirm that UW-OFDM is at least equivalently robust to channel estimation errors as CP-OFDM. Furthermore, keeping in mind that a combination of $\mathbf{G}''$ and $r = 1/2$ will not be used in practice, as $\mathbf{G}'$ and $r = 1/2$ is simply better suited (see Fig. 4.18), UW-OFDM shows for this coding rate a significantly better robustness than CP-OFDM. This outcome becomes even more meaningful considering that in contrast to CP-OFDM, UW-OFDM suffers from channel estimation errors in a twofold manner, namely imperfect data estimation and additionally imperfect UW subtraction.

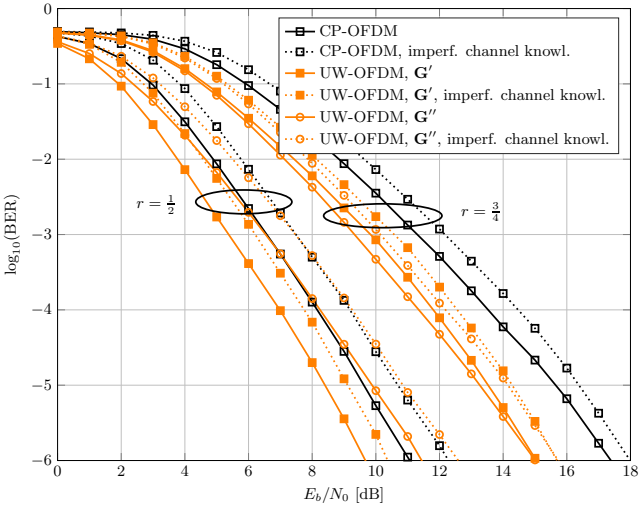


Figure 4.18.: Impact of imperfect channel estimation on the BER performance of UW-OFDM ($\mathbf{G}'$ and $\mathbf{G}''$) with a CAZAC sequence as UW in combination with 'partial scaling' (Tab. 4.1) and CP-OFDM for coded transmission ($r = 1/2$ and $r = 3/4$) in a multipath environment with $\tau_{\mathrm{RMS}} = 100$ ns.

Finally, Fig. 4.19 shows the performance degradation in case the UW samples are scaled to have the same average power as the remaining $N - N_u$ samples of the OFDM symbol, cf. Tab. 4.1 and 'full scaling'. CP-OFDM is missing in this graph,

as there is no equivalent setup defined in this work with appropriate pilot symbol energy E_p' to ensure a fair comparison. The performance of UW-OFDM degrades for $\mathbf{G}'$ by 0.8 dB and 0.7 dB and for $\mathbf{G}''$ by 1.2 dB and 0.9 dB in case of $r = 1/2$ and $r = 3/4$, respectively. In comparison to the results in Fig. 4.18, the relative performance loss due to imperfect channel estimation increases by at most 0.2 dB, therefore determining only a moderate additional loss.

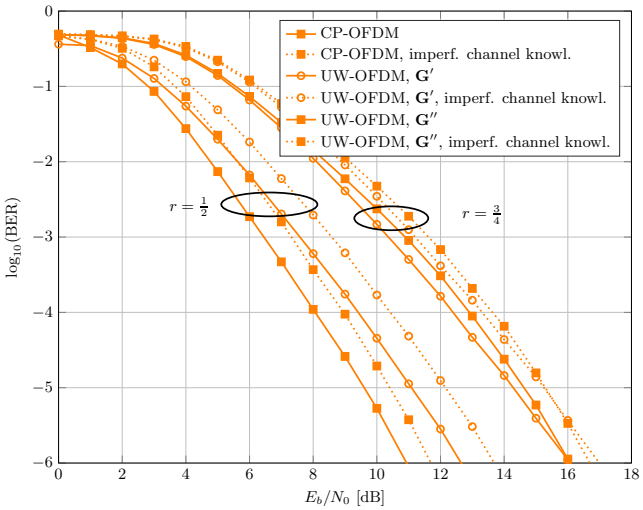


Figure 4.19.: Impact of imperfect channel estimation on the BER performance of UW-OFDM ($\mathbf{G}'$ and $\mathbf{G}''$) with a CAZAC sequence as UW in combination with 'full scaling' (Tab. 4.1) for coded transmission ($r = 1/2$ and $r = 3/4$) in a multipath environment with $\tau_{\mathrm{RMS}} = 100$ ns.

5. Inclusion of Pilot Tones in Unique Word OFDM

Pilot tones are a popular means – especially in multi-carrier techniques – to estimate various system parameters at the receiver. This chapter extends the UW-OFDM concept in a way that it also enables the inclusion of deterministic pilot symbols $\mathbf{p} \in \mathbb{C}^{N_p \times 1}$ at dedicated subcarriers in the frequency domain. Since the UW already constitutes a collection of deterministic pilot symbols in the time domain, the necessity of additional pilot tones might be questioned at first hand. Nevertheless, this work shall provide the prerequisites, the decision whether parameter estimation is based solely on UWs, pilot tones or a combination of both is decided best from situation to situation.

The presented framework is based on partitioning the transmit signal into two additive terms. The first term maps the data symbols on the signal by using a generator matrix $\mathbf{G}_d \in \mathbb{C}^{(N_d+N_r+N_p)\times N_d}$, and the second term incorporates the pilot symbols by using $\mathbf{G}_p \in \mathbb{C}^{(N_d+N_r+N_p)\times N_p}$, yielding

$$\mathbf{x} = \mathbf{F}_N^{-1}\mathbf{B}\underbrace{\mathbf{B}_p\breve{\mathbf{G}}_d}_{\mathbf{G}_d}\mathbf{d} + \mathbf{F}_N^{-1}\mathbf{B}\mathbf{G}_p\mathbf{p} \tag{5.1}$$

$$= \mathbf{F}_N^{-1}\mathbf{B}\mathbf{G}_d\mathbf{d} + \mathbf{F}_N^{-1}\mathbf{B}\mathbf{G}_p\mathbf{p} = \begin{bmatrix}\mathbf{x}_d \\ \mathbf{0}\end{bmatrix}. \tag{5.2}$$

$\mathbf{B} \in \mathbb{C}^{N\times(N_d+N_r+N_p)}$ models the insertion of zero subcarriers and exactly coincides in design and size with matrix $\mathbf{B}$ from the previous chapters, only the definition of the size itself slightly varies due to the additional parameter N_p[1]. $\mathbf{B}_p \in \mathbb{C}^{(N_d+N_r+N_p)\times(N_d+N_r)}$ takes care of inserting N_p additional zeros at the positions of the pilot subcarriers. Consequently, the data part $\mathbf{G}_d\mathbf{d}$ will not influence the pilot subcarriers, in contrast however, the pilot part with $\mathbf{G}_p\mathbf{p}$ is allowed to overlay the data part.

Note that the definition of $\mathbf{G}_d = \mathbf{B}_p\breve{\mathbf{G}}_d$ ($\breve{\mathbf{G}}_d \in \mathbb{C}^{(N_d+N_r)\times N_d}$) has been introduced to simplify notational matters in terms of taking care of only one zero subcarrier insertion matrix. This will especially be useful in the next chapter. Independent of the specific notation though, the partition of the transmit signal in (5.2) into two additive and independent terms shows that the generator matrices can

[1] Contrary to the previous chapters, in this one it holds that $N - N_z = N_d + N_r + N_p$ instead of $N - N_z = N_d + N_r$.

be optimized independently from each other as part of two distinct optimization problems with distinct cost functions. These problems will be tackled in the following.

5.1. Optimization of Data Generator Matrix

The optimization problem for the data generator matrix is given as

$$\check{\mathbf{G}}_d = \underset{\mathbf{G}_d}{\operatorname{argmin}} \{J_d\} \quad \text{s.t. } \mathbf{F}_N^{-1}\mathbf{B}\mathbf{G}_d = \begin{bmatrix} \boldsymbol{\Xi} \\ \mathbf{0} \end{bmatrix}, \tag{5.3}$$

with J_d denoting a freely selectable cost function. In order to achieve an unconstrained optimization problem as in the previous chapters, let us choose the approach

$$\mathbf{G}_d = \mathbf{B}_p\mathbf{A}\begin{bmatrix} \mathbf{I} \\ \mathbf{T}_1 \end{bmatrix}, \tag{5.4}$$

whereas $\mathbf{A} \in \mathbb{R}^{(N_d+N_r)\times(N_d+N_r)}$ and $\mathbf{T}_1 \in \mathbb{C}^{N_r\times N_d}$ with $\mathbf{T}_1 = {\mathbf{M}'_{22}}^{-1}\mathbf{M}'_{21}$. The submatrices $\mathbf{M}'_{21} \in \mathbb{C}^{N_u\times N_d}$ and $\mathbf{M}'_{22} \in \mathbb{C}^{N_u\times N_r}$ with $N_u = N_r$ follow from

$$\mathbf{F}_N^{-1}\mathbf{B}\mathbf{B}_p\mathbf{A} = \begin{bmatrix} \mathbf{M}'_{11} & \mathbf{M}'_{12} \\ \mathbf{M}'_{21} & \mathbf{M}'_{22} \end{bmatrix}. \tag{5.5}$$

For pilotless UW-OFDM systems, minimizing the sum of the error covariances after data estimation at a fixed SNR turned out to deliver generator matrices with very good BER performance (see Sec. 3.4 and chapter 4). This motivates to apply the same optimization criterion in case of UW-OFDM symbols with pilot subcarriers. With (2.37) and (5.2), the transmit signal $\mathbf{x}'$ can be modelled as

$$\mathbf{x}' = \mathbf{F}_N^{-1}\left(\mathbf{B}\mathbf{G}_d\mathbf{d} + \mathbf{B}\mathbf{G}_p\mathbf{p} + \tilde{\mathbf{x}}_u\right), \tag{5.6}$$

where $\tilde{\mathbf{x}}_u = \mathbf{F}_N\begin{bmatrix}\mathbf{0}^T & \mathbf{x}_u^T\end{bmatrix}^T$ represents the frequency domain version of the UW. Similar to (2.65), the frequency domain signal at the receiver (the zero subcarriers are already excluded) corresponds to an affine model

$$\tilde{\mathbf{y}}_d = \tilde{\mathbf{H}}\mathbf{G}_d\mathbf{d} + \tilde{\mathbf{H}}\mathbf{G}_p\mathbf{p} + \tilde{\mathbf{H}}\mathbf{B}^T\tilde{\mathbf{x}}_u + \mathbf{B}^T\mathbf{F}_N\mathbf{n}. \tag{5.7}$$

As a preparatory step, the known signal parts (assuming that $\tilde{\mathbf{H}}$ or at least an estimate of it is available) caused by the UW and the pilots are subtracted from $\tilde{\mathbf{y}}_d$, yielding the linear model

$$\tilde{\mathbf{y}} = \tilde{\mathbf{y}}_d - \tilde{\mathbf{H}}\mathbf{B}^T\tilde{\mathbf{x}}_u - \tilde{\mathbf{H}}\mathbf{G}_p\mathbf{p} \tag{5.8}$$

$$= \tilde{\mathbf{H}}\mathbf{G}_d\mathbf{d} + \mathbf{v}. \tag{5.9}$$

This model is equivalent to (2.68) and confirms that every UW-OFDM approach – regardless of the presence of pilot symbols in the frequency domain – leads to the very same transmission model. The only difference between an UW-OFDM system with and without pilot tones is an additional subtraction of the pilot induced offset on top of the UW part to yield the linear model. Based on (5.9), the following conclusions can already be drawn:

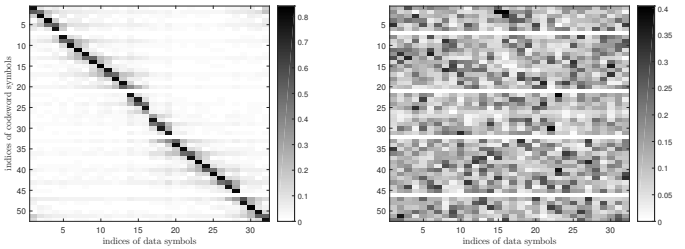
- The same linear transmission model enables the deployment of the same estimator concepts as in the case without pilots.
- In the previous chapters, the generator matrix has been optimized w.r.t. the whole transceiver performance, e.g., by minimizing the sum of the error variances at a given SNR in case of an LMMSE estimator. Due to the same linear transmission model, the same optimization criteria can be followed, leading to $J_d = J_{\text{LMMSE}}$ or $J_d = J_{\text{BLUE}}$, cf. (4.11).
- Pilots are introduced for system parameter estimation purposes and then simply subtracted before the data estimation process, as they do not contribute any information to the latter. In this sense, pilots denote a certain overhead to the primary task of data estimation in a receiver that should be kept as small as possible. From that perspective, $\mathbf{G}_p$ shall be designed such that it enables the insertion of pilot symbols and fulfills the constraint in (5.2), while keeping the required energy at a minimum. The design of $\mathbf{G}_p$ is the topic of Sec. 5.2.

Within this notation, $\mathbf{G}_d$ as defined in (5.4) takes over the tasks of $\mathbf{G}$ from the pilotless case. Consequently, the same cost functions and the very same steepest descent algorithm from Sec. 4.2 can be used. Fig. 5.1 illustrates the generator matrices $\mathbf{G}_d'$ and $\mathbf{G}_d''$, whereas the steepest descent algorithm has been initialized once with $\mathbf{A}^{(0)} = \mathbf{P}'$ and once with $\left[\mathbf{A}^{(0)}\right]_{ij} \sim \mathcal{N}(0,1)$. Note that $\mathbf{P}'$ generally varies from $\mathbf{P}$ in (4.25), as the indices of the pilot subcarriers are excluded as optimization parameters for the permutation matrix. Still, the similarities with $\mathbf{G}'$ and $\mathbf{G}''$ from the pilotless case in chapter 4 are highly visible. The pilot based UW-OFDM system is based on setup F in Tab. 2.3. This setup is a modified version of setup A with additional $N_p = 4$ pilot subcarriers at indices $\mathcal{I} = \{7, 21, 43, 57\}$, the number of data symbols reduces therefore to $N_d = 32$.

5.2. Optimization of Pilot Generator Matrix

The optimization problem for the pilot generator matrix is given as

$$\check{\mathbf{G}}_p = \underset{\mathbf{G}_p}{\operatorname{argmin}} \{J_p\} \quad \text{s.t. } \mathbf{F}_N^{-1}\mathbf{B}\mathbf{G}_p = \begin{bmatrix} \boldsymbol{\Xi} \\ \mathbf{0} \end{bmatrix}, \tag{5.10}$$



(a) $|\mathbf{G}'_d|$ (b) $|\mathbf{G}''_d|$

Figure 5.1.: Magnitude of entries of data generator matrices $\mathbf{G}_d$.

with J_p denoting a freely selectable cost function. This generator matrix shall place the pilot symbols on dedicated subcarriers and load the remaining subcarriers such that the zero word constraint is fulfilled. Further, the linear model in (5.9) demonstrates that the pilots cannot be used for the primary task of data estimation, thus suggesting an optimization of $\mathbf{G}_p$ w.r.t. an energy minimization. In fact, this optimization problem shows strong similarities with the one for finding generator matrices of a systematically encoded UW-OFDM system. Consequently, let us choose the approach

$$\mathbf{G}_p = \mathbf{P}_p \begin{bmatrix} \mathbf{I} \\ \mathbf{T}_2 \end{bmatrix}, \tag{5.11}$$

whereas $\mathbf{P}_p \in \{0,1\}^{(N_d+N_r+N_p)\times(N_d+N_r+N_p)}$ is a permutation matrix that places the pilot symbols at the corresponding subcarrier positions, and $\mathbf{T}_2 \in \mathbb{C}^{(N_d+N_r)\times N_p}$ is responsible for loading the other subcarriers such that the zero word constraint is fulfilled. $\mathbf{T}_2$ is calculated as $\mathbf{T}_2 = -{\mathbf{M}''_{22}}^{\dagger}\mathbf{M}''_{21}$ with the submatrices $\mathbf{M}''_{21} \in \mathbb{C}^{N_u\times N_p}$ and $\mathbf{M}''_{22} \in \mathbb{C}^{N_u\times(N_d+N_r)}$ with $N_u = N_r$ derived from

$$\mathbf{F}_N^{-1}\mathbf{B}\mathbf{P}_p = \begin{bmatrix} \mathbf{M}''_{11} & \mathbf{M}''_{12} \\ \mathbf{M}''_{21} & \mathbf{M}''_{22} \end{bmatrix}. \tag{5.12}$$

The energy induced by the pilots is given by the cost function[2]

$$J_p = E_p^{(r)} N = \mathbf{p}^H \mathbf{G}_p^H \mathbf{G}_p \mathbf{p}. \tag{5.13}$$

The two different factors influencing the resulting energy $E_p^{(r)}$ are

- the positions of the pilot subcarriers determined by $\mathbf{P}_p$, and
- the values of the pilot symbols.

[2] A scaling of the energy term by N is introduced to omit the prefactor $\frac{1}{N}$ and simplify the resulting cost function J_p.

A close to uniform distribution of the pilot subcarriers over the available bandwidth is beneficial for various system parameter estimation tasks, cf. [81]. Keeping in mind the main purpose of pilot symbols, the positions are thus usually ruled out as a degree of freedom to minimize (5.13). This leaves the pilot symbols $\mathbf{p}$ as optimization parameters. In the following, the pilot symbols are constrained to have unit energy with $|p_i| = 1$, leading to the problem formulation

$$\check{\mathbf{p}} = \underset{\mathbf{p}}{\operatorname{argmin}} \{J_p\} \quad \text{s.t.} \quad \mathbf{F}_N^{-1}\mathbf{B}\mathbf{G}_p = \begin{bmatrix} \boldsymbol{\Xi} \\ \mathbf{0} \end{bmatrix} \ \wedge \ |p_i| = 1. \tag{5.14}$$

The optimization problem has been solved by an exhaustive search, whereas a pilot symbol is drawn from an alphabet

$$\mathcal{A} = \left\{ e^{j\frac{2\pi k}{|\mathcal{A}|}} \right\} \qquad k = 0 \ldots |\mathcal{A}| - 1 \tag{5.15}$$

with cardinality $|\mathcal{A}|$, resulting in $|\mathcal{A}|^{N_p}$ different combinations in case of N_p pilot subcarriers. Tab. 5.1 shows the resulting minimum value for $E_p^{(r)}$ as a function of the cardinality of the pilot symbol alphabet. Note that a scaling factor N

Table 5.1.: Minimum pilot induced energy $E_p^{(r)}$ as a function of the cardinality of $\mathcal{A}$.

$\lvert\mathcal{A}\rvert$	2	4	6	10	20
$N \cdot E_p^{(r)}$	5.4633	5.2423	5.1969	5.1864	5.1783

originating from the DFT has been introduced to easier link the resulting energies with the number of pilots N_p. Since the performance does not significantly vary with the cardinality $|\mathcal{A}|$, the granularity of the alphabet can thus kept low. Fig. 5.2 illustrates the realization of $\mathbf{G}_p$ achieving minimum energy $E_p^{(r)}$ for $|\mathcal{A}| = 20$. The corresponding pilot symbols p_i are derived from (5.15) with $k = k_i$ and k_i denoting the ith element of the vector $\mathbf{k} = [17, 14, 3, 0]$.

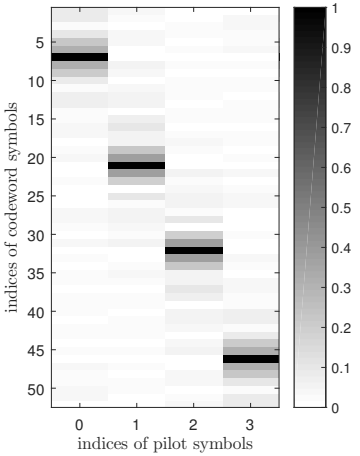


Figure 5.2.: Magnitude of entries of pilot generator matrix $\mathbf{G}_p$ when optimized w.r.t. $E_p^{(r)}$.

6. Carrier Frequency Offset

A well-known critical issue for OFDM systems is the presence of a carrier frequency offset (CFO). A CFO is the result of an oscillator mismatch between transmitter and receiver, or of a Doppler effect due to a movement of at least one of the two. The CFO experienced by a real communication system can be quite high, i.e., in the order of multiples of the subcarrier spacing. Due to the high sensitivity of OFDM, an accurate estimation and compensation of the CFO is essential. Covering a large range while at the same time providing high accuracy are, in general, contradictory requirements on an estimation task. Therefore, estimation and compensation is usually split up in two phases, the acquisition and the tracking phase [82]. This allows the deployment of two distinct algorithms, where one is optimized to offer a large range while the other provides high accuracy. In the acquisition phase, a rough estimate is obtained which is usually valid for a whole burst. The acquisition is often based on a preamble that is sent at the beginning of each burst, a typical example would be the IEEE 802.11a standard [36]. The estimate obtained in this phase should compensate for the integer part of the CFO (measured in integer units of the subcarrier spacing), but to a certain extent also for the fractional part (measured as a fraction of the subcarrier spacing). In the subsequent tracking phase, the remaining CFO can be safely assumed not to exceed 10% of the subcarrier spacing, thus enabling a finer estimation. This estimation and compensation is usually performed on an OFDM symbol by symbol basis.

In general, the acquisition phase in UW-OFDM does not differ from that of other signaling schemes as e.g., CP-OFDM. Utilizing the same preamble and algorithms will deliver the same results, therefore no additional insights can be gained. Hence, only the tracking phase will be considered in the following. For all investigations carried out in the following, the CFO is modelled to be static during transmission. Furthermore, ideal timing synchronization with a perfect detection of the beginning of a burst is assumed.

This chapter starts with introducing a mathematical description of the CFO and incorporating it in the present UW-OFDM model[1]. Subsequently, the model is expanded to the receiver side as well (Sec. 6.2). In Sec. 6.3 the effect of imperfect UW and pilot tone subtraction due to nonideal CFO estimation is elaborated. Sec. 6.4 shows the potential of different CFO compensation techniques under the assumption that the additive parts of UW and pilot tones have been perfectly

[1] Note that parts of the CFO model have already been developed in [83].

subtracted. Next, Sec. 6.5 details a CFO estimation technique based on pilot subcarriers, of which the results have already been implicitly used in the precedent sections. Of course, the UW may also be used for CFO estimation, hence additional pilots in the frequency domain may seem unnecessary. However, the main goal of this CFO analysis is to investigate the impact of a CFO on the performance of UW-OFDM and to compare it against CP-OFDM. In order to ensure a fair comparison, the same estimation method based on pilot tones is thus employed, an intentional utilization of the UW is avoided. Finally, Sec. 6.6 concludes this chapter by investigating the degrading effects of a CFO on the BER performance of UW-OFDM and CP-OFDM.

Considerations of non-zero UWs in this chapter are restricted to investigations, whether they influence the performance of the system in the presence of a CFO. In this case, the samples of a non-zero UW have been scaled such that the average power of an UW sample corresponds to the average power of the non-zero samples of an UW-OFDM symbol with zero word (cf. Sec. 4.3.3). Therefore, all samples of the resulting UW-OFDM symbol have the same average power.

Two different classes of generator matrices are considered in this chapter. For visualizing the effects of a CFO in Sec. 6.1, the matrices $\mathbf{G}'$ and $\mathbf{G}''$ already known from chapter 4 are investigated. Since these matrices do not incorporate pilot tones for CFO estimation, $\mathbf{G}'_d$ or $\mathbf{G}''_d$ in combination with $\mathbf{G}_p$ derived in chapter 5 are taken from Sec. 6.2 on. Note that $\mathbf{G}'_d$ and $\mathbf{G}'$ as well as $\mathbf{G}''_d$ and $\mathbf{G}''$ share the same optimization criterion and initialization of the optimization algorithm, hence also show similar properties and CFO behavior. The decision to use the pilotless generator matrices in Sec. 6.1 anyway is due to the reader's high familiarity with $\mathbf{G}'$ and $\mathbf{G}''$ from the detailed analysis in chapter 4. This should ensure a better understanding of the CFO effects specific to UW-OFDM and different to other signaling schemes. Although many findings apply to the systematic UW-OFDM approach as well, the focus is laid on non-systematically encoded UW-OFDM in this chapter. This is due to the current research status suggesting non-systematically encoded UW-OFDM to be the more attractive approach.

In order to provide a thorough analytical framework, the following derivations incorporate an additive white Gaussian noise term. However, for all simulations presented in this chapter except for Sec. 6.6, the additive noise is always chosen to be zero to clearly identify and elaborate the error only caused by CFO effects.

6.1. CFO Model

Assuming a carrier frequency offset f_{CFO} present in a system, the time domain samples in the complex baseband experience an incremental phase offset of

$$y(uT_s) = \mathrm{e}^{j2\pi f_{\mathrm{CFO}} uT_s} \mathrm{e}^{j\phi_0} x(uTs), \tag{6.1}$$

where u denotes the discrete time variable, T_s the sampling time and ϕ_0 denotes an arbitrary phase offset. Perfect timing synchronization is expected to take care of ϕ_0, hence it is set to $\phi_0 = 0$ and discarded in the following. The impact of f_{CFO} on the OFDM system performance depends on the relative proportion to the subcarrier spacing Δ_f rather than on the absolute value, thus motivating to introduce a relative carrier frequency offset

$$\epsilon = \frac{f_{\mathrm{CFO}}}{\Delta_f} = \frac{f_{\mathrm{CFO}}}{f_s/N} = \frac{f_{\mathrm{CFO}} N}{f_s} = f_{\mathrm{CFO}} N T_s. \tag{6.2}$$

With (6.2) the incremental phase offset in (6.1) translates to $\mathrm{e}^{j\frac{2\pi\epsilon u}{N}}$ (where $u = 0, 1, \ldots N-1$ when considering only one OFDM symbol) and in matrix notation to

$$\mathbf{y} = \mathbf{\Lambda}' \mathbf{x} \tag{6.3}$$

with $\mathbf{\Lambda}' \in \mathbb{C}^{N \times N}$ given as

$$\mathbf{\Lambda}' = \mathrm{diag}\left(\begin{bmatrix} 1 & \mathrm{e}^{j\frac{2\pi\epsilon}{N}} & \cdots & \mathrm{e}^{j\frac{2\pi\epsilon(N-1)}{N}} \end{bmatrix}^T \right). \tag{6.4}$$

In order to take into account the phase accumulated by previous OFDM symbols and the additional UW in front of the burst (which ensures the cyclic structure for the first OFDM symbol, cf. Fig. 2.8), a diagonal matrix $\mathbf{\Lambda}'^{(l)} \in \mathbb{C}^{N \times N}$ with

$$\mathbf{\Lambda}'^{(l)} = \mathrm{e}^{j\psi_l} \mathbf{\Lambda}' = \mathrm{e}^{j\frac{2\pi\epsilon(Nl+N_u)}{N}} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & \mathrm{e}^{j\frac{2\pi\epsilon}{N}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \mathrm{e}^{j\frac{2\pi\epsilon(N-1)}{N}} \end{bmatrix} \tag{6.5}$$

is introduced, whereas $l \in \{0, 1, \ldots L-1\}$, L denotes the number of OFDM symbols per burst, and ψ_l is a phase offset defined as

$$\psi_l = \frac{2\pi\epsilon\,(Nl + N_u)}{N}. \tag{6.6}$$

In the following, matrices with the notation $'$ as e.g., $\mathbf{\Lambda}'$ encompass the whole frequency range including the zero subcarriers. The counterpart $\mathbf{\Lambda}$ and similar matrices introduced later will only comprise non-zero subcarriers.

Starting with the transmit signal in (5.6) and taking into account (6.5), the lth OFDM time domain symbol at the receiver can be modelled as

$$\mathbf{y}_r^{(l)} = \mathbf{\Lambda}'^{(l)}\mathbf{H}_c\mathbf{F}_N^{-1}\left(\mathbf{B}\mathbf{G}_d\mathbf{d} + \mathbf{B}\mathbf{G}_p\mathbf{p} + \tilde{\mathbf{x}}_u\right) + \mathbf{n} \tag{6.7}$$

$$= \mathbf{\Lambda}'^{(l)}\mathbf{F}_N^{-1}\mathbf{F}_N\mathbf{H}_c\mathbf{F}_N^{-1}\left(\tilde{\mathbf{x}}_d^{(l)} + \tilde{\mathbf{x}}_p + \tilde{\mathbf{x}}_u\right) + \mathbf{n} \tag{6.8}$$

$$= \mathbf{\Lambda}'^{(l)}\mathbf{F}_N^{-1}\tilde{\mathbf{H}}'\tilde{\mathbf{x}}''^{(l)} + \mathbf{n}, \tag{6.9}$$

with $\tilde{\mathbf{x}}''^{(l)} = \tilde{\mathbf{x}}_d^{(l)} + \tilde{\mathbf{x}}_p + \tilde{\mathbf{x}}_u \in \mathbb{C}^{N\times 1}$ summarizing the effects of data, pilots and the UW in one frequency domain vector. Let us highlight again at this point, that in time domain a CFO leads to an accumulation of a phase offset from sample to sample of an OFDM symbol. This insight is required later on when detailing the effects of a CFO on UW-OFDM. In order to fully explain these effects, an analysis of a CFO in the frequency domain is also necessary. Applying the DFT yields the frequency domain OFDM symbol

$$\tilde{\mathbf{y}}_r^{(l)} = \mathbf{F}_N\mathbf{y}_r^{(l)} \tag{6.10}$$

$$= \mathbf{F}_N\mathbf{\Lambda}'^{(l)}\mathbf{F}_N^{-1}\tilde{\mathbf{H}}'\tilde{\mathbf{x}}''^{(l)} + \mathbf{F}_N\mathbf{n} \tag{6.11}$$

$$= \tilde{\mathbf{\Lambda}}'^{(l)}\tilde{\mathbf{H}}'\tilde{\mathbf{x}}''^{(l)} + \mathbf{v}', \tag{6.12}$$

with a noise vector $\mathbf{v}' \in \mathbb{C}^{N\times 1}$ defined as

$$\mathbf{v}' = \mathbf{F}_N\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, N\sigma_n^2\mathbf{I}), \tag{6.13}$$

which has a simple relationship with the noise vector $\mathbf{v}$ used in the previous chapters so far (see e.g., (2.68))

$$\mathbf{v} = \mathbf{B}^T\mathbf{F}_N\mathbf{n} = \mathbf{B}^T\mathbf{v}'. \tag{6.14}$$

Since $\mathbf{\Lambda}'^{(l)}$ is diagonal, multiplying with $\mathbf{F}_N$ and $\mathbf{F}_N^{-1}$ results in a circulant matrix $\tilde{\mathbf{\Lambda}}'^{(l)}$. In order to provide a better insight on $\tilde{\mathbf{\Lambda}}'^{(l)}$ and its relationship to ϵ, let us start with the definition of the kth element of vector $\tilde{\mathbf{y}}_r^{(l)}$

$$\tilde{y}_r^{(l)}[k] = \left[\mathbf{F}_N\right]_{k,*}\mathbf{y}_r^{(l)} = \sum_{u=0}^{N-1} \mathrm{e}^{-j\frac{2\pi ku}{N}} y_r^{(l)}[u]. \tag{6.15}$$

According to (6.9), the uth element is expressed as

$$y_r^{(l)}[u] = \left[\mathbf{\Lambda}'^{(l)}\right]_{u,*}\mathbf{F}_N^{-1}\tilde{\mathbf{H}}'\tilde{\mathbf{x}}''^{(l)} + n[u] \tag{6.16}$$

$$= \mathrm{e}^{j\psi_l}\mathrm{e}^{j\frac{2\pi\epsilon u}{N}}\left[\mathbf{F}_N^{-1}\right]_{u,*}\tilde{\mathbf{H}}'\tilde{\mathbf{x}}''^{(l)} + n[u] \tag{6.17}$$

$$= \mathrm{e}^{j\psi_l}\mathrm{e}^{j\frac{2\pi\epsilon u}{N}}\frac{1}{N}\sum_{m=0}^{N-1}\mathrm{e}^{j\frac{2\pi mu}{N}}\left[\tilde{\mathbf{H}}'\right]_{m,m}\tilde{x}''^{(l)}[m] + n[u], \tag{6.18}$$

where $[\cdot]_{u,*}$ represents all elements of row number u. Note that (6.17) follows from (6.16) by considering that only the uth element of the vector $\left[\mathbf{\Lambda}'^{(l)}\right]_{u,*}$ is non-zero. Plugging (6.18) into (6.15) and rearranging yields

$$\begin{aligned}
\tilde{y}_r^{(l)}[k] &= \sum_{u=0}^{N-1} \mathrm{e}^{-j\frac{2\pi ku}{N}} \mathrm{e}^{j\psi_l} \mathrm{e}^{j\frac{2\pi\epsilon u}{N}} \\
&\quad \cdot \frac{1}{N} \sum_{m=0}^{N-1} \mathrm{e}^{j\frac{2\pi mu}{N}} \left[\tilde{\mathbf{H}}'\right]_{m,m} \tilde{x}'^{(l)}[m] + v'[k] \\
&= \mathrm{e}^{j\psi_l} \frac{1}{N} \sum_{u=0}^{N-1} \sum_{m=0}^{N-1} \mathrm{e}^{j\frac{2\pi(m+\epsilon-k)u}{N}} \left[\tilde{\mathbf{H}}'\right]_{m,m} \tilde{x}'^{(l)}[m] + v'[k] \\
&= \mathrm{e}^{j\psi_l} \frac{1}{N} \sum_{m=0}^{N-1} \left[\tilde{\mathbf{H}}'\right]_{m,m} \tilde{x}'^{(l)}[m] \sum_{u=0}^{N-1} \mathrm{e}^{j\frac{2\pi(m+\epsilon-k)u}{N}} + v'[k].
\end{aligned} \tag{6.19}$$

The relationship between $\tilde{\mathbf{y}}_r^{(l)}$ and $\tilde{\mathbf{x}}'^{(l)}$ is fully determined by $\tilde{\mathbf{\Lambda}}'^{(l)}\tilde{\mathbf{H}}'$, cf. (6.12). Applying this knowledge on (6.19) together with $\tilde{\mathbf{\Lambda}}'^{(l)} = \mathrm{e}^{j\psi_l}\tilde{\mathbf{\Lambda}}'$ leads to the definition

$$\left[\tilde{\mathbf{\Lambda}}'\right]_{k,m} = \frac{1}{N} \sum_{u=0}^{N-1} \mathrm{e}^{j\frac{2\pi}{N}(m+\epsilon-k)u} \qquad k = 0 \ldots N-1;\ m = 0 \ldots N-1. \tag{6.20}$$

(6.21)

These formulas allow a compact notation of $\tilde{\mathbf{\Lambda}}'^{(l)}$, however, an immediate interpretation of the CFO impact in the frequency domain is rather difficult. This will thus be provided in the following. Let $\tilde{x}_h^{(l)}[m] = \left[\tilde{\mathbf{H}}'\right]_{m,m} \tilde{x}'^{(l)}[m]$ for reasons of compactness, and separate the impact of the subcarrier in consideration (indicated with index k) from all others, then (6.19) can be rewritten as

$$\begin{aligned}
\tilde{y}_r^{(l)}[k] &= \frac{1}{N}\mathrm{e}^{j\psi_l} \tilde{x}_h^{(l)}[k] \sum_{u=0}^{N-1} \mathrm{e}^{j\frac{2\pi\epsilon u}{N}} \\
&\quad + \frac{1}{N}\mathrm{e}^{j\psi_l} \sum_{m=0, m\neq k}^{N-1} \tilde{x}_h^{(l)}[m] \sum_{u=0}^{N-1} \mathrm{e}^{j\frac{2\pi(m+\epsilon-k)u}{N}} + v'[k] \qquad (6.22) \\
&= \frac{1}{N}\mathrm{e}^{j\psi_l} \tilde{x}_h^{(l)}[k] \frac{1-\mathrm{e}^{j2\pi\epsilon}}{1-\mathrm{e}^{j\frac{2\pi\epsilon}{N}}} \\
&\quad + \frac{1}{N}\mathrm{e}^{j\psi_l} \sum_{m=0, m\neq k}^{N-1} \tilde{x}_h^{(l)}[m] \frac{1-\mathrm{e}^{j2\pi(m+\epsilon-k)}}{1-\mathrm{e}^{j\frac{2\pi(m+\epsilon-k)}{N}}} + v'[k] \qquad (6.23) \\
&= \frac{1}{N}\mathrm{e}^{j\psi_l} \frac{\mathrm{e}^{j\pi\epsilon}\left(\mathrm{e}^{-j\pi\epsilon} - \mathrm{e}^{j\pi\epsilon}\right)}{\mathrm{e}^{j\frac{\pi\epsilon}{N}}\left(\mathrm{e}^{-j\frac{\pi\epsilon}{N}} - \mathrm{e}^{j\frac{\pi\epsilon}{N}}\right)} \tilde{x}_h^{(l)}[k] + \frac{1}{N}\mathrm{e}^{j\psi_l}
\end{aligned}$$

$$\cdot \sum_{m=0,m\neq k}^{N-1} \tilde{x}_h^{(l)}[m] \frac{\mathrm{e}^{j\pi(m+\epsilon-k)} \left(\mathrm{e}^{-j\pi(m+\epsilon-k)} - \mathrm{e}^{j\pi(m+\epsilon-k)}\right)}{\mathrm{e}^{j\frac{\pi(m+\epsilon-k)}{N}} \left(\mathrm{e}^{-j\frac{\pi(m+\epsilon-k)}{N}} - \mathrm{e}^{j\frac{\pi(m+\epsilon-k)}{N}}\right)} + v'[k] \quad (6.24)$$

$$= \mathrm{e}^{j\psi_l} \mathrm{e}^{j\frac{\pi\epsilon(N-1)}{N}} \frac{\sin(\pi\epsilon)}{N\sin(\frac{\pi\epsilon}{N})} \tilde{x}_h^{(l)}[k] + \mathrm{e}^{j\psi_l} \mathrm{e}^{j\frac{\pi\epsilon(N-1)}{N}}$$

$$\cdot \sum_{m=0,m\neq k}^{N-1} \mathrm{e}^{j\frac{\pi(m-k)(N-1)}{N}} \frac{\sin\left(\pi(m+\epsilon-k)\right)}{N\sin\left(\frac{\pi(m+\epsilon-k)}{N}\right)} \tilde{x}_h^{(l)}[m] + v'[k]. \quad (6.25)$$

Note that (6.23) follows from (6.22) by applying the formula for the sum of a geometric series $S_n = \sum_{p=0}^{P-1} r^p = \frac{1-r^P}{1-r}$, and (6.24) is a preparation step to apply $\left(\mathrm{e}^{-ja} - \mathrm{e}^{ja}\right) = 2j\sin(a)$. Finally, the frequency domain receive signal corrupted by CFO is

$$\tilde{y}_r^{(l)}[k] = \mathrm{e}^{j\psi_l} \mathrm{e}^{j\frac{2\pi}{N}\epsilon\left(\frac{N-1}{2}\right)} \frac{\sin(\pi\epsilon)}{N\sin(\frac{\pi\epsilon}{N})} \left[\tilde{\mathbf{H}}'\right]_{k,k} \tilde{x}''^{(l)}[k] + i^{(l)}[k] + v'[k] \quad (6.26)$$

$$= \mathrm{e}^{j\varphi_l} \frac{\sin(\pi\epsilon)}{N\sin(\frac{\pi\epsilon}{N})} \left[\tilde{\mathbf{H}}'\right]_{k,k} \tilde{x}''^{(l)}[k] + i^{(l)}[k] + v'[k], \quad (6.27)$$

with an intercarrier interference (ICI) $i^{(l)}[k]$ given as

$$i^{(l)}[k] = \mathrm{e}^{j\varphi_l} \sum_{m=0,m\neq k}^{N-1} \mathrm{e}^{j\frac{\pi(m-k)(N-1)}{N}} \frac{\sin\left(\pi(m+\epsilon-k)\right)}{N\sin\left(\frac{\pi(m+\epsilon-k)}{N}\right)} \left[\tilde{\mathbf{H}}'\right]_{m,m} \tilde{x}''^{(l)}[m], \quad (6.28)$$

and a phase offset

$$\varphi_l = \psi_l + \frac{2\pi}{N}\epsilon\left(\frac{N-1}{2}\right) = \frac{2\pi}{N}\epsilon\left(Nl + N_u + \frac{N-1}{2}\right). \quad (6.29)$$

In order to simplify notation further for subsequent derivations, let us additionally introduce

$$\psi = \frac{2\pi}{N}\epsilon\left(\frac{N-1}{2}\right) \quad (6.30)$$

to yield

$$\varphi_l = \psi_l + \psi. \quad (6.31)$$

There are three effects of a CFO on a subcarrier symbol $\tilde{x}''^{(l)}[k]$:

- A phase offset by φ_l,
- an attenuation of $\frac{\sin(\pi\epsilon)}{N\sin(\pi\epsilon/N)}$, and

- an intercarrier interference term $i^{(l)}[k]$ with similar properties as additive noise.

It will be shown later in this section that φ_l causes a rotation of the data symbols in the underlying constellation diagram. Therefore, and also to distinguish from other phase offsets presented in the following, the phase offset caused by φ_l will from now on be referred to as *phase rotation*.

In matrix notation, (6.27) translates to

$$\begin{aligned}\tilde{\mathbf{y}}_r^{(l)} &= \tilde{\mathbf{\Lambda}}'^{(l)}\tilde{\mathbf{H}}'\tilde{\mathbf{x}}''^{(l)} + \mathbf{v}' \\ &= \mathrm{e}^{j\psi_l}\tilde{\mathbf{\Lambda}}'\tilde{\mathbf{H}}'\tilde{\mathbf{x}}''^{(l)} + \mathbf{v}' \\ &= \mathrm{e}^{j\psi_l}\mathrm{e}^{j\psi}\tilde{\mathbf{\Lambda}}'_{\mathrm{stat}}\tilde{\mathbf{H}}'\tilde{\mathbf{x}}''^{(l)} + \mathbf{v}' \\ &= \mathrm{e}^{j\varphi_l}\tilde{\mathbf{\Lambda}}'_{\mathrm{stat}}\tilde{\mathbf{H}}'\tilde{\mathbf{x}}''^{(l)} + \mathbf{v}'. \end{aligned} \tag{6.32}$$

Note that $\tilde{\mathbf{\Lambda}}'_{\mathrm{stat}} \in \mathbb{C}^{N\times N}$ with

$$\tilde{\mathbf{\Lambda}}'_{\mathrm{stat}} = \mathrm{e}^{-j\psi}\tilde{\mathbf{\Lambda}}', \tag{6.33}$$

$$\left[\tilde{\mathbf{\Lambda}}'_{\mathrm{stat}}\right]_{k,m} = \frac{\sin\left(\pi(m+\epsilon-k)\right)}{N\sin\left(\frac{\pi(m+\epsilon-k)}{N}\right)}\mathrm{e}^{j\frac{\pi(m-k)(N-1)}{N}}. \tag{6.34}$$

For the main diagonal entries with $k = m$, (6.34) collapses to $\frac{\sin(\pi\epsilon)}{N\sin(\frac{\pi\epsilon}{N})}$. In other words, the corresponding frequency domain symbol $\tilde{x}''^{(l)}[k]$ experiences only an attenuation but no phase rotation. Furthermore, the main diagonal entries are independent of the subcarrier index k and the OFDM symbol number l, thus motivating the nomenclature static.

The derivations so far characterize a CFO effect in the time as well as in the frequency domain. Nevertheless, the impact on an actual UW-OFDM transmission system is still not clear. The reason is that for non-systematically encoded UW-OFDM the data symbols *lie somewhere between* these two domains, which will be elaborated subsequently. Of course, this approach of *lying between domains* is only a conceptual idea to make the occurring CFO effects more comprehensible. As a preparatory step, let us extend (6.32) such that the CFO effects on the data symbols can be modelled as

$$\begin{aligned}\mathbf{d}_{\mathrm{CFO}}^{(l)} &= \mathbf{E}\tilde{\mathbf{y}}_r^{(l)} \\ &= \mathbf{E}\mathrm{e}^{j\varphi_l}\tilde{\mathbf{\Lambda}}'_{\mathrm{stat}}\tilde{\mathbf{H}}'\tilde{\mathbf{x}}''^{(l)} + \mathbf{E}\mathbf{v}' \\ &= \mathbf{E}\mathrm{e}^{j\varphi_l}\tilde{\mathbf{\Lambda}}'_{\mathrm{stat}}\mathbf{I}\mathbf{B}\mathbf{G}\mathbf{d}^{(l)} + \mathbf{E}\mathbf{v}' \\ &= \mathrm{e}^{j\varphi_l}\mathbf{E}\tilde{\mathbf{\Lambda}}'_{\mathrm{stat}}\breve{\mathbf{G}}\mathbf{d}^{(l)} + \mathbf{E}\mathbf{v}'. \end{aligned} \tag{6.35}$$

AWGN conditions with $\tilde{\mathbf{H}} = \mathbf{I}$ as well as $\tilde{\mathbf{x}}_p = \mathbf{0}$ and $\tilde{\mathbf{x}}_u = \mathbf{0}$ are assumed for better isolating the impact of a CFO. $\mathbf{E}$ determines any potential linear estimator,

$\mathbf{G}$ corresponds to a generator matrix of an UW-OFDM system and $\mathbf{B}$ models an optional insertion of zero subcarriers. $\mathbf{G}$ and $\mathbf{B}$ are merged to $\breve{\mathbf{G}} = \mathbf{BG}$ to span the whole frequency range. This notation does not change any properties, but avoids taking care of $\mathbf{B}$ separately on the one hand, and additionally eases a comparison of UW-OFDM with other signaling schemes on the other hand. Alternatively, the impact on the kth data symbol corresponds to

$$\begin{aligned} d_{\mathrm{CFO}}^{(l)}[k] &= \mathrm{e}^{j\varphi_l}\,[\mathbf{E}]_{k,*}\,\tilde{\mathbf{\Lambda}}'_{\mathrm{stat}}\breve{\mathbf{G}}\mathbf{d}^{(l)} + [\mathbf{E}]_{k,*}\,\mathbf{v}' \\ &= \mathrm{e}^{j\varphi_l}\,[\mathbf{E}]_{k,*}\,\tilde{\mathbf{\Lambda}}'_{\mathrm{stat}}\left[\breve{\mathbf{G}}\right]_{*,k} d^{(l)}[k] \\ &\quad + \mathrm{e}^{j\varphi_l}\,[\mathbf{E}]_{k,*}\,\tilde{\mathbf{\Lambda}}'_{\mathrm{stat}}\sum_{m=0,m\neq k}^{N_d-1}\left[\breve{\mathbf{G}}\right]_{*,m} d^{(l)}[m] + [\mathbf{E}]_{k,*}\,\mathbf{v}' \\ &= \underbrace{\mathrm{e}^{j\varphi_l}\mathbf{e}_k^T\tilde{\mathbf{\Lambda}}'_{\mathrm{stat}}\breve{\mathbf{g}}_k d^{(l)}[k]}_{\text{corrupted data symbol}} + \underbrace{\mathrm{e}^{j\varphi_l}\mathbf{e}_k^T\tilde{\mathbf{\Lambda}}'_{\mathrm{stat}}\sum_{m=0,m\neq k}^{N_d-1}\breve{\mathbf{g}}_m d^{(l)}[m]}_{\text{interdata interference}} + \underbrace{[\mathbf{E}]_{k,*}\,\mathbf{v}'}_{\text{additive noise}}\,. \end{aligned} \tag{6.36}$$

Both, the elementwise and matrix notation are used in the following, depending on which notation provides the best prerequisites for interpretation. There are now three terms in (6.36) determining $d_{\mathrm{CFO}}^{(l)}[k]$, namely a *corrupted* version of the actually transmitted data symbol $d^{(l)}[k]$, an interference term caused by the other data symbols denoted as *interdata interference*, and an *additive noise* term, the latter without any connection to a CFO. It seems important to stress at this point, that a CFO still introduces the degrading effects phase rotation, attenuation and intercarrier interference as stated in (6.27). Nevertheless, these effects have a different impact depending on the specific signaling scheme, which will be shown in the subsequent paragraphs.

For a better understanding it is beneficial to classify all signaling schemes based on the domain in which data symbols are specified. Hence, on one side there are so called single-carrier systems (preferably implemented as single-carrier/frequency domain equalization (SC/FDE)), and on the other side multi-carrier systems (in this case with special focus on OFDM systems). Characteristic for the first group is the consecutive placement of data symbols next to each other in time domain. As a consequence, the symbols are orthogonal in time and thus are also detected in the very same domain. In contrast, data symbols in OFDM systems are orthogonal to each other and detected in frequency domain. According to this classification, all other schemes like e.g., precoded OFDM or UW-OFDM as well, can be interpreted as a mixture of these two concepts, meaning that the data symbols "lie somewhere between" time and frequency domain. In fact, the following derivations demonstrate that the CFO effects on UW-OFDM data symbols highly depend on the design of the generator matrix $\breve{\mathbf{G}}$.

Starting with single-carrier systems[2], let $\breve{\mathbf{G}} = \mathbf{F}_N$ and $\mathbf{E} = \mathbf{F}_N^{-1}$ (and of course with appropriately sized vector $\mathbf{d}^{(l)}$), then

$$\begin{aligned}\mathbf{d}_{\mathrm{CFO}}^{(l)} &= \mathrm{e}^{j\varphi_l}\mathbf{F}_N^{-1}\tilde{\mathbf{\Lambda}}'_{\mathrm{stat}}\mathbf{F}_N\mathbf{d}^{(l)} + \mathbf{F}_N^{-1}\mathbf{v}' \\ &= \mathrm{e}^{j\varphi_l}\mathrm{e}^{-j\psi}\mathbf{F}_N^{-1}\tilde{\mathbf{\Lambda}}'\mathbf{F}_N\mathbf{d}^{(l)} + \mathbf{n} \\ &= \mathrm{e}^{j\varphi_l}\mathrm{e}^{-j\psi}\mathbf{\Lambda}'\mathbf{d}^{(l)} + \mathbf{n}. \end{aligned} \tag{6.37}$$

Eq. (6.37) follows from $\tilde{\mathbf{\Lambda}}' = \mathrm{e}^{j\psi}\tilde{\mathbf{\Lambda}}'_{\mathrm{stat}}$ defined in (6.33) and the fact that multiplying with $\mathbf{F}_N^{-1}$ and $\mathbf{F}_N$ diagonalizes the circulant matrix $\tilde{\mathbf{\Lambda}}'$. W.r.t. the notation in (6.36), this diagonalization simply means that the second term modelling intercarrier interference vanishes. Assuming now that each data symbol is drawn from a QPSK alphabet with unit variance, Fig. 6.1 exemplarily illustrates the effects of a CFO on the data symbols. As already indicated before, each time domain sample and therefore each data symbol experiences a phase rotation, whereas the phase rotation increases with each sample. Consequently, the rotation is different for all data symbols.

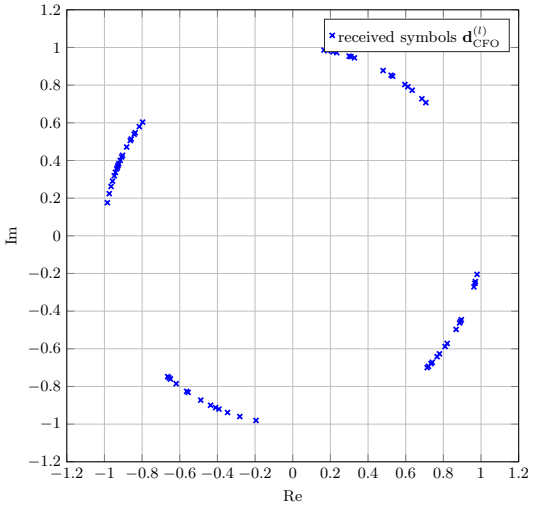


Figure 6.1.: Constellation diagram of one single data symbol vector instance $\mathbf{d}_{\mathrm{CFO}}^{(l)}$ with $d^{(l)}[k] \in 1/\sqrt{2}\{-1-j, -1+j, 1-j, 1+j\}$ for $\breve{\mathbf{G}} = \mathbf{F}_N$, $\mathbf{E} = \mathbf{F}_N^{-1}$ (i.e., single-carrier), $\epsilon = 0.1$, $l = 0$ and $\sigma_v^2 = 0$.

[2]Note that pulse shaping filters, which are usually part of communication systems, are omitted for simplicity.

An OFDM system is modelled as $\breve{\mathbf{G}} = \mathbf{I}_N$ and $\mathbf{E} = \mathbf{I}_N$, yielding

$$\mathbf{d}_{\mathrm{CFO}}^{(l)} = \mathrm{e}^{j\varphi_l}\mathbf{I}\tilde{\mathbf{\Lambda}}'_{\mathrm{stat}}\mathbf{I}\mathbf{d}^{(l)} + \mathbf{v}' \tag{6.38}$$

$$= \mathrm{e}^{j\varphi_l}\tilde{\mathbf{\Lambda}}'_{\mathrm{stat}}\mathbf{d}^{(l)} + \mathbf{v}'. \tag{6.39}$$

Again, the size of $\mathbf{d}^{(l)}$ is adjusted accordingly. W.r.t. one single element, the second term in (6.36) remains and $\mathbf{e}_k^T\tilde{\mathbf{\Lambda}}'_{\mathrm{stat}}\breve{\mathbf{g}}_k$ in the first term collapses to a constant scaling factor corresponding to the normalized sinc function

$$\mathrm{sinc}(\epsilon) = \frac{\sin(\pi\epsilon)}{N\sin(\frac{\pi\epsilon}{N})}. \tag{6.40}$$

The CFO effects on the data symbols differ substantially, cf. Fig. 6.2. First, the

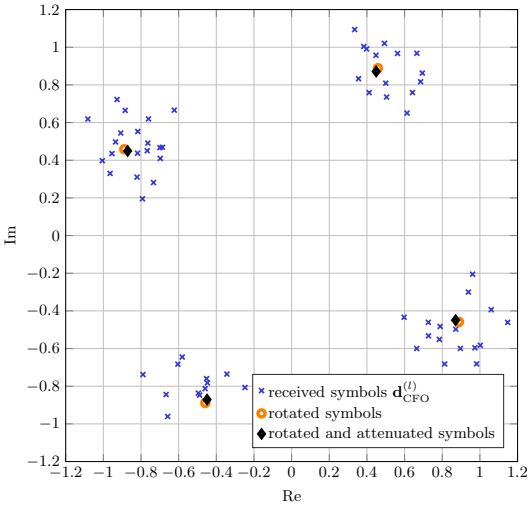


Figure 6.2.: Constellation diagram of one single data symbol vector instance $\mathbf{d}_{\mathrm{CFO}}^{(l)}$ with $d^{(l)}[k] \in 1/\sqrt{2}\{-1-j, -1+j, 1-j, 1+j\}$ for $\breve{\mathbf{G}} = \mathbf{E} = \mathbf{I}$ (i.e., conventional OFDM), $\epsilon = 0.1$, $l = 0$ and $\sigma_v^2 = 0$.

data symbols are rotated by an angle φ_l. Isolating this effect would deliver symbols corresponding to the circles. Contrary to single-carrier based systems, all data symbols within one OFDM symbol share the very same phase rotation. Second, the data symbols experience an attenuation indicated by the diamond symbols. Third, each constellation point is additionally expanded to a cloud due to intercarrier interference, leading to the actually received symbols $\mathbf{d}_{\mathrm{CFO}}^{(l)}$ displayed as crosses. The latter two effects can be well explained from Fig. 6.3. In frequency domain, each subcarrier symbol is in fact represented by a sinc function (6.40) weighted

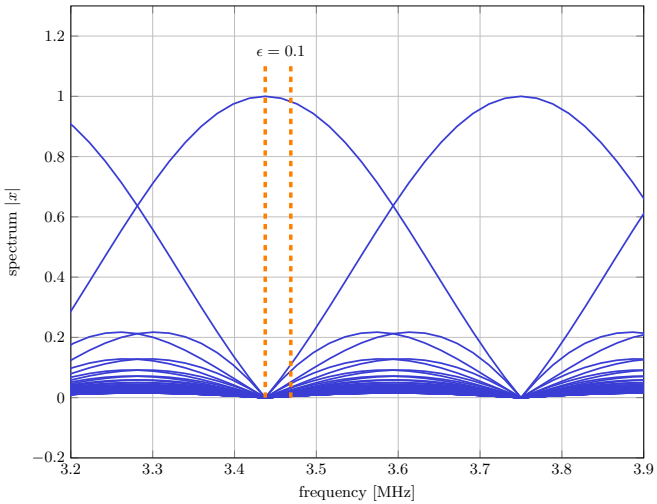


Figure 6.3.: Intercarrier interference and magnitude attenuation due to CFO shown at the decomposed magnitude spectrum of one data symbol vector $\mathbf{d}_{\mathrm{CFO}}^{(l)}$ from a conventional OFDM system.

with $d^{(l)}[k]$. This sinc function is one at the subcarrier position it belongs to and zero at all others. For $\epsilon = 0$ and thus perfect alignment, the effect of the sinc function is thus not noticeable at the receiver side. In case of $\epsilon \neq 0$, a data symbol experiences an attenuation and an additive interference term caused by intercarrier interference. For $\epsilon = 0.1$, the absolute value of the main diagonal entries of $\tilde{\mathbf{\Lambda}}'_{\mathrm{stat}}$ and thus the attenuation corresponds to 0.98. The entries off the main diagonal determine the interference and decrease with the distance to the main diagonal; for this particular ϵ value from 0.1 for the first to already 0.03 for the third minor diagonal. Hence, the performance degradation of these two effects is rather limited. The most prominent effect arises from the phase rotation by φ_l and is typically the only one which needs to be estimated and taken care of to achieve sufficient CFO compensation in current communications systems. This kind of compensation is also known as carrier phase synchronization in the literature.

Having elaborated on both border concepts, the next part investigates UW-OFDM. One interesting candidate is UW based SC/FDE [22, 23, 30] with a generator matrix defined as

$$\breve{\mathbf{G}} = \mathbf{G}_{\text{uw-sc}} = \mathbf{F}_N \begin{bmatrix} \mathbf{I} \\ \mathbf{0} \end{bmatrix} \tag{6.41}$$

and an estimator

$$\mathbf{E} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \end{bmatrix} \mathbf{F}_N^{-1}. \tag{6.42}$$

Inserting into (6.35) immediately yields

$$\mathbf{d}_{\mathrm{CFO}}^{(l)} = \mathrm{e}^{j\varphi_l} \mathrm{e}^{-j\psi} \mathbf{\Lambda}_{\mathrm{sub}} \mathbf{d}^{(l)} + \mathbf{n}, \tag{6.43}$$

where $\mathbf{\Lambda}_{\mathrm{sub}} \in \mathbb{C}^{(N-N_u)\times(N-N_u)}$ denotes a submatrix according to $\mathbf{\Lambda}' = \left[\begin{smallmatrix} \mathbf{\Lambda}_{\mathrm{sub}} & * \\ * & * \end{smallmatrix}\right]$. Obviously, data symbols in UW-SC/FDE experience the same linear incremental phase rotation as in classical SC based systems, only the range of the phase ramp differs due to the last N_u samples dedicated to the UW. Fig. 6.4 displays the resulting constellation diagram.

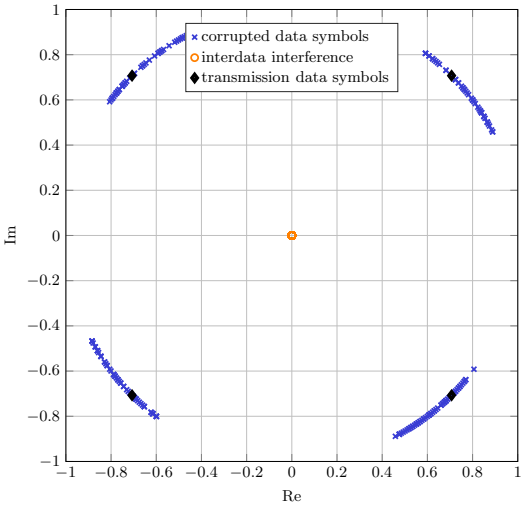


Figure 6.4.: Constellation diagram of data symbol vectors $\mathbf{d}_{\mathrm{CFO}}^{(l)}$ (several instances) with $d^{(l)}[k] \in 1/\sqrt{2}\{-1-j, -1+j, 1-j, 1+j\}$ for an UW based SC/FDE system $\breve{\mathbf{G}} = \mathbf{F}_N \begin{bmatrix} \mathbf{I} & \mathbf{0}^T \end{bmatrix}^T$, $\mathbf{E} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \end{bmatrix} \mathbf{F}_N^{-1}$, $\epsilon = 0.1$, $\sigma_v^2 = 0$ and $\varphi_l = 0$.

A little bit further away but still showing ties to SC systems is the generator matrix $\mathbf{G}''$ visualized in Fig. 4.5. With $\breve{\mathbf{G}} = \mathbf{B}\mathbf{G}''$ and $\mathbf{E} = (\mathbf{B}\mathbf{G}'')^{\dagger}$, the phase ramp is not linearly increasing in k anymore as for UW-SC/FDE (where k denotes the data symbol index in $d_{\mathrm{CFO}}^{(l)}[k]$), but there is still a ramp behavior when displaying the phase rotations in ascending order, cf. Fig. 6.5. As expected, $\breve{\mathbf{G}} = \mathbf{B}\mathbf{G}'$ (cf. Fig. 4.2) with $\mathbf{E} = (\mathbf{B}\mathbf{G}')^{\dagger}$ is closer to conventional OFDM systems. The data symbols of one OFDM symbol do not experience one common phase rotation. However, the

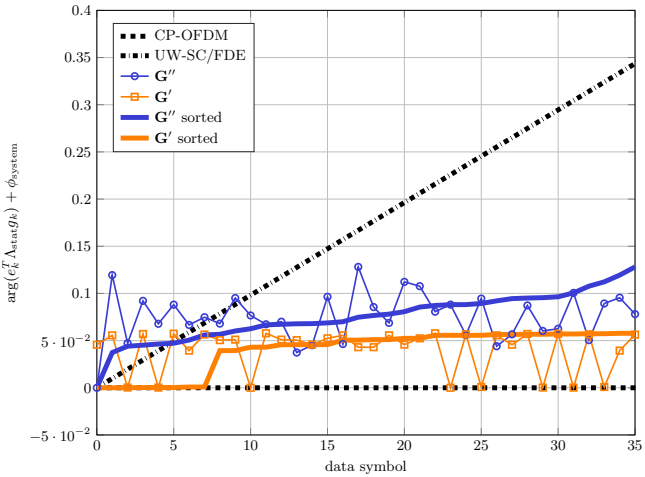


Figure 6.5.: Phase rotation experienced by the data symbols $d_{\mathrm{CFO}}^{(l)}[k]$ due to a CFO of $\epsilon = 0.1$. For easier comparison, an individual offset ϕ_{system} is added to the phase trajectory of each system to start at zero radiants.

number of different angles is still rather limited compared to SC systems. Fig. 6.6 and 6.7 visualize the constellation diagram of $\mathbf{d}_{\mathrm{CFO}}^{(l)}$ for $\mathbf{G}'$ and $\mathbf{G}''$, respectively, whereas $\mathbf{d}_{\mathrm{CFO}}^{(l)}$ is split up into two terms according to (6.36), with $\sigma_v^2 = 0$ and $\varphi_l = 0$ for easier comparison. The data symbols $d_{\mathrm{CFO}}^{(l)}[k]$ experience almost no attenuation in case of both matrices, but the phase rotations differ. The power of the interdata interference term (averaged over data and OFDM symbols) is very similar for both matrices, but lower than in a conventional OFDM system presented in Fig. 6.2 (about a factor 2 for the considered setups). This is one reason for a higher robustness of UW-OFDM against CFO performance degradation effects (Sec. 6.3) compared to conventional OFDM systems. The second reason is due to a better CFO estimation (Sec. 6.5) which again results from a difference in the interference power.

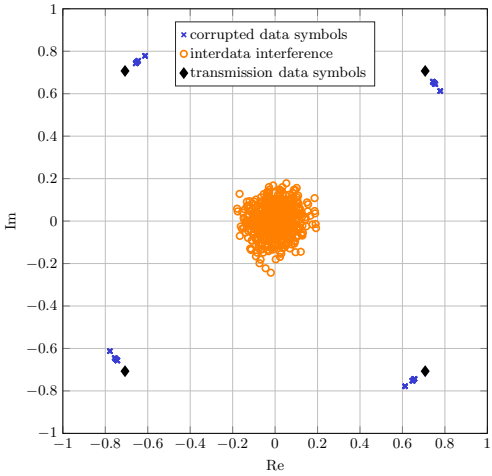


Figure 6.6.: Constellation diagram of data symbol vectors $\mathbf{d}_{\mathrm{CFO}}^{(l)}$ (several instances) with $d^{(l)}[k] \in 1/\sqrt{2}\{-1-j, -1+j, 1-j, 1+j\}$ for non-systematically encoded UW-OFDM $\breve{\mathbf{G}} = \mathbf{B}\mathbf{G}'$, $\mathbf{E} = (\mathbf{B}\mathbf{G}')^{\dagger}$, $\epsilon = 0.1$, $\sigma_v^2 = 0$ and $\varphi_l = 0$.

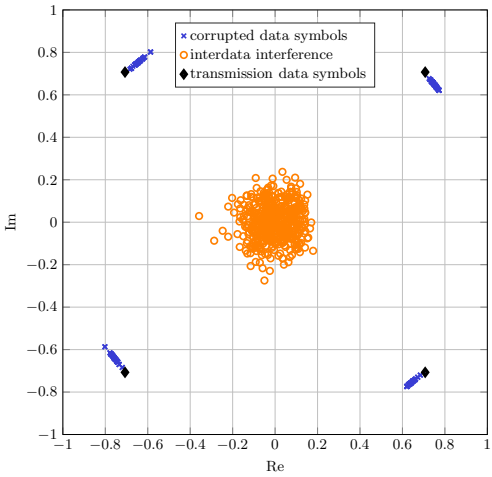


Figure 6.7.: Constellation diagram of data symbol vectors $\mathbf{d}_{\mathrm{CFO}}^{(l)}$ (several instances) with $d^{(l)}[k] \in 1/\sqrt{2}\{-1-j, -1+j, 1-j, 1+j\}$ for non-systematically encoded UW-OFDM $\breve{\mathbf{G}} = \mathbf{B}\mathbf{G}''$, $\mathbf{E} = (\mathbf{B}\mathbf{G}'')^{\dagger}$, $\epsilon = 0.1$, $\sigma_v^2 = 0$ and $\varphi_l = 0$.

Fig. 6.8 summarizes the relationship between CFO effects and the domain data symbols lie. Depending on the specific properties, UW-OFDM experiences CFO effects either related to single-carrier or multi-carrier systems. In a certain way, choosing a generator matrix allows also choosing the impact of the CFO effects on an UW-OFDM system.

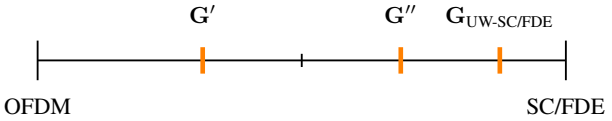


Figure 6.8.: Categorization of the carrier frequency offset effects on the different UW-OFDM generator matrices.

6.2. Receiver Model

After modelling the CFO effects properly, the next part evaluates the impact on the performance of an UW-OFDM system. For performance analysis it suffices to consider the non-zero subcarriers only and thus zero subcarriers are discarded, yielding the downsized vector

$$\tilde{\mathbf{y}}_d^{(l)} = \mathbf{B}^T \tilde{\mathbf{y}}_r^{(l)}. \tag{6.44}$$

Excluding zero subcarriers is straightforward in (5.7) due to the diagonal structure of the channel matrix $\tilde{\mathbf{H}}$. However, $\tilde{\mathbf{\Lambda}}'^{(l)}$ in (6.12) is a dense matrix, requiring thus more in-depth investigations. For an easier understanding, introducing a permutation matrix $\mathbf{P}_s$ yields a sorted vector $\tilde{\mathbf{x}}_s \in \mathbb{C}^{N\times 1}$ with

$$\tilde{\mathbf{x}}_s = \mathbf{P}_s \tilde{\mathbf{x}}''^{(l)} = \begin{bmatrix} \tilde{\mathbf{x}}_z \\ \tilde{\mathbf{x}}_n \end{bmatrix}, \tag{6.45}$$

where $\tilde{\mathbf{x}}_z \in \mathbb{C}^{N_z\times 1}$ denotes the load on the zero and $\tilde{\mathbf{x}}_n \in \mathbb{C}^{(N-N_z)\times 1}$ on the non-zero subcarriers. Note that for all vectors and matrices which are used only locally to derive the receiver model, the subindex l is dropped for the sake of simplicity. $\mathbf{P}_s^T\mathbf{P}_s = \mathbf{I}$, $\mathbf{B}^T = \begin{bmatrix}\mathbf{0} & \mathbf{I}\end{bmatrix}\mathbf{P}_s$, (6.12),(6.14) and (6.45) deliver

$$\begin{aligned}
\tilde{\mathbf{y}}_d^{(l)} &= \mathbf{B}^T \tilde{\mathbf{\Lambda}}'^{(l)} \tilde{\mathbf{H}}' \tilde{\mathbf{x}}''^{(l)} + \mathbf{B}^T \mathbf{v}' && (6.46)\\
&= \mathbf{B}^T \tilde{\mathbf{\Lambda}}'^{(l)} \mathbf{P}_s^T \mathbf{P}_s \tilde{\mathbf{H}}' \mathbf{P}_s^T \mathbf{P}_s \tilde{\mathbf{x}}''^{(l)} + \mathbf{v} && (6.47)\\
&= \begin{bmatrix}\mathbf{0} & \mathbf{I}\end{bmatrix} \mathbf{P}_s \tilde{\mathbf{\Lambda}}'^{(l)} \mathbf{P}_s^T \mathbf{P}_s \tilde{\mathbf{H}}' \mathbf{P}_s^T \mathbf{P}_s \tilde{\mathbf{x}}''^{(l)} + \mathbf{v} && (6.48)\\
&= \begin{bmatrix}\mathbf{0} & \mathbf{I}\end{bmatrix} \tilde{\mathbf{\Lambda}}'_s \tilde{\mathbf{H}}'_s \tilde{\mathbf{x}}_s + \mathbf{v}. && (6.49)
\end{aligned}$$

$\tilde{\mathbf{H}}'_s \in \mathbb{C}^{N\times N}$ and $\tilde{\mathbf{\Lambda}}'_s \in \mathbb{C}^{N\times N}$ represent sorted versions defined as

$$\tilde{\mathbf{H}}'_s = \mathbf{P}_s \tilde{\mathbf{H}}' \mathbf{P}_s^T = \begin{bmatrix} \tilde{\mathbf{H}}_z & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{H}}_n \end{bmatrix} \tag{6.50}$$

$$\tilde{\mathbf{\Lambda}}'_s = \mathbf{P}_s \tilde{\mathbf{\Lambda}}'^{(l)} \mathbf{P}_s^T = \begin{bmatrix} \tilde{\mathbf{\Lambda}}_{zz} & \tilde{\mathbf{\Lambda}}_{nz} \\ \tilde{\mathbf{\Lambda}}_{zn} & \tilde{\mathbf{\Lambda}}_{nn} \end{bmatrix}, \tag{6.51}$$

with the diagonal submatrices $\tilde{\mathbf{H}}_z \in \mathbb{C}^{N_z\times N_z}$ and $\tilde{\mathbf{H}}_n \in \mathbb{C}^{(N-N_z)\times(N-N_z)}$, and the dense matrices $\tilde{\mathbf{\Lambda}}_{zz} \in \mathbb{C}^{N_z\times N_z}$, $\tilde{\mathbf{\Lambda}}_{nz} \in \mathbb{C}^{N_z\times(N-N_z)}$, $\tilde{\mathbf{\Lambda}}_{zn} \in \mathbb{C}^{(N-N_z)\times N_z}$, and $\tilde{\mathbf{\Lambda}}_{nn} \in \mathbb{C}^{(N-N_z)\times(N-N_z)}$. The nomenclature of the submatrices of $\tilde{\mathbf{\Lambda}}'_s$ determines the origin and the sink, e.g., a subindex *nz* denotes the influence of the non-zero subcarrier symbols *on* the zero subcarriers. These definitions lead to

$$\tilde{\mathbf{y}}_d^{(l)} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{\Lambda}}_{zz} & \tilde{\mathbf{\Lambda}}_{nz} \\ \tilde{\mathbf{\Lambda}}_{zn} & \tilde{\mathbf{\Lambda}}_{nn} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{H}}_z & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{H}}_n \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{x}}_z \\ \tilde{\mathbf{x}}_n \end{bmatrix} + \mathbf{v} \tag{6.52}$$

$$= \tilde{\mathbf{\Lambda}}_{zn} \tilde{\mathbf{H}}_z \tilde{\mathbf{x}}_z + \tilde{\mathbf{\Lambda}}_{nn} \tilde{\mathbf{H}}_n \tilde{\mathbf{x}}_n + \mathbf{v}. \tag{6.53}$$

The matrices $\tilde{\mathbf{\Lambda}}_{nn}$ and $\tilde{\mathbf{H}}_n$ originate from excluding all elements with indices corresponding to zero subcarriers, hence

$$\tilde{\mathbf{\Lambda}}_{nn} = \mathbf{B}^T \tilde{\mathbf{\Lambda}}'^{(l)} \mathbf{B} = \tilde{\mathbf{\Lambda}}^{(l)} \in \mathbb{C}^{(N-N_z)\times(N-N_z)} \tag{6.54}$$

$$\tilde{\mathbf{H}}_n = \mathbf{B}^T \tilde{\mathbf{H}}' \mathbf{B} = \tilde{\mathbf{H}} \in \mathbb{C}^{(N-N_z)\times(N-N_z)}, \tag{6.55}$$

the latter already known from (2.63). Consequently, (6.53) follows to

$$\tilde{\mathbf{y}}_d^{(l)} = \tilde{\mathbf{\Lambda}}^{(l)} \tilde{\mathbf{H}} \tilde{\mathbf{x}}_n + \tilde{\mathbf{\Lambda}}_{zn} \tilde{\mathbf{H}}_z \tilde{\mathbf{x}}_z + \mathbf{v}. \tag{6.56}$$

Taking into account

$$\tilde{\mathbf{x}}_n = \begin{bmatrix} \mathbf{0} & \mathbf{I} \end{bmatrix} \tilde{\mathbf{x}}_s = \begin{bmatrix} \mathbf{0} & \mathbf{I} \end{bmatrix} \mathbf{P}_s \tilde{\mathbf{x}}''^{(l)} = \mathbf{B}^T \tilde{\mathbf{x}}''^{(l)}, \tag{6.57}$$

the first term in (6.56) translates to

$$\tilde{\mathbf{\Lambda}}^{(l)} \tilde{\mathbf{H}} \tilde{\mathbf{x}}_n = \tilde{\mathbf{\Lambda}}^{(l)} \tilde{\mathbf{H}} \mathbf{B}^T \tilde{\mathbf{x}}''^{(l)} \tag{6.58}$$

$$= \tilde{\mathbf{\Lambda}}^{(l)} \tilde{\mathbf{H}} \mathbf{B}^T \left(\mathbf{B} \mathbf{G}_d \mathbf{d}^{(l)} + \mathbf{B} \mathbf{G}_p \mathbf{p} + \tilde{\mathbf{x}}_u \right) \tag{6.59}$$

$$= \tilde{\mathbf{\Lambda}}^{(l)} \tilde{\mathbf{H}} \mathbf{G}_d \mathbf{d}^{(l)} + \tilde{\mathbf{\Lambda}}^{(l)} \tilde{\mathbf{H}} \mathbf{G}_p \mathbf{p} + \tilde{\mathbf{\Lambda}}^{(l)} \tilde{\mathbf{H}} \mathbf{B}^T \tilde{\mathbf{x}}_u. \tag{6.60}$$

Based on (6.8) and (6.45), the load on the zero subcarriers $\tilde{\mathbf{x}}_z$ can be decomposed into three components

$$\tilde{\mathbf{x}}_z = \tilde{\mathbf{x}}_{d,z} + \tilde{\mathbf{x}}_{p,z} + \tilde{\mathbf{x}}_{u,z}, \tag{6.61}$$

whereas the part of the data $\tilde{\mathbf{x}}_{d,z}$ and the pilots $\tilde{\mathbf{x}}_{p,z}$ is zero by definition, leaving $\tilde{\mathbf{x}}_z = \tilde{\mathbf{x}}_{u,z}$. This insight yields

$$\tilde{\mathbf{y}}_d^{(l)} = \tilde{\mathbf{\Lambda}}^{(l)} \tilde{\mathbf{H}} \mathbf{G}_d \mathbf{d}^{(l)} + \tilde{\mathbf{\Lambda}}^{(l)} \tilde{\mathbf{H}} \mathbf{G}_p \mathbf{p} + \tilde{\mathbf{\Lambda}}^{(l)} \tilde{\mathbf{H}} \mathbf{B}^T \tilde{\mathbf{x}}_u + \tilde{\mathbf{\Lambda}}_{zn} \tilde{\mathbf{H}}_z \tilde{\mathbf{x}}_{u,z} + \mathbf{v}. \tag{6.62}$$

In contrast to (5.7), the zero subcarriers cannot simply be excluded anymore. There is now a fourth term fed by the load of an UW on the actual zero subcarriers which impacts the non-zero subcarriers and thus the performance. Considering though that zero subcarriers follow a dedicated purpose like shaping the spectral mask, a UW shall be chosen to maintain these properties. Hence, in a real-world system $\tilde{\mathbf{x}}_{u,z}$ almost vanishes by design such that $\tilde{\mathbf{\Lambda}}_{zn}\tilde{\mathbf{H}}_z\tilde{\mathbf{x}}_{u,z} \to \mathbf{0}$, leading again to the affine receiver model

$$\tilde{\mathbf{y}}_d^{(l)} \approx \tilde{\mathbf{\Lambda}}^{(l)}\tilde{\mathbf{H}}\mathbf{G}_d\mathbf{d}^{(l)} + \tilde{\mathbf{\Lambda}}^{(l)}\tilde{\mathbf{H}}\mathbf{G}_p\mathbf{p} + \tilde{\mathbf{\Lambda}}^{(l)}\tilde{\mathbf{H}}\mathbf{B}^T\tilde{\mathbf{x}}_u + \mathbf{v} \tag{6.63}$$

with negligible approximation error.

In order to perform data detection according to Sec. 2.3, the CFO effects have to be compensated, followed by a subtraction of the known signal parts to transform the affine into a linear model. As already stated before, derotating by φ_l (also known as carrier phase synchronization) normally suffices for an adequate CFO compensation in current communication systems. Since φ_l is usually not known, an estimate $\hat{\varphi}_l$ with a certain estimation error is utilized instead. The estimation process is detailed in Sec. 6.5, at the moment $\hat{\varphi}_l$ is simply treated as given. The corrected received signal reads

$$\tilde{\mathbf{y}}_c^{(l)} = \mathrm{e}^{-j\hat{\varphi}_l}\tilde{\mathbf{\Lambda}}^{(l)}\tilde{\mathbf{H}}\mathbf{G}_d\mathbf{d}^{(l)} + \mathrm{e}^{-j\hat{\varphi}_l}\tilde{\mathbf{\Lambda}}^{(l)}\tilde{\mathbf{H}}\mathbf{G}_p\mathbf{p} + \mathrm{e}^{-j\hat{\varphi}_l}\tilde{\mathbf{\Lambda}}^{(l)}\tilde{\mathbf{H}}\mathbf{B}^T\tilde{\mathbf{x}}_u + \mathrm{e}^{-j\hat{\varphi}_l}\mathbf{v}. \tag{6.64}$$

As a next step, the known portions in the signal have to be subtracted to obtain a linear model as in (2.68). Assuming derotation to be already an adequate CFO compensation means, subtracting the known offset

$$\mathbf{x}_{\mathrm{off}} = \tilde{\mathbf{H}}\mathbf{G}_p\mathbf{p} + \tilde{\mathbf{H}}\mathbf{B}^T\tilde{\mathbf{x}}_u \tag{6.65}$$

caused by the UW and the pilots yields

$$\begin{aligned}
\tilde{\mathbf{y}}^{(l)} &= \tilde{\mathbf{y}}_c^{(l)} - \tilde{\mathbf{H}}\mathbf{G}_p\mathbf{p} - \tilde{\mathbf{H}}\mathbf{B}^T\tilde{\mathbf{x}}_u && (6.66)\\
&= \mathrm{e}^{-j\hat{\varphi}_l}\tilde{\mathbf{\Lambda}}^{(l)}\tilde{\mathbf{H}}\mathbf{G}_d\mathbf{d}^{(l)} + \left(\mathrm{e}^{-j\hat{\varphi}_l}\tilde{\mathbf{\Lambda}}^{(l)} - \mathbf{I}\right)\tilde{\mathbf{H}}\mathbf{G}_p\mathbf{p} && (6.67)\\
&\quad + \left(\mathrm{e}^{-j\hat{\varphi}_l}\tilde{\mathbf{\Lambda}}^{(l)} - \mathbf{I}\right)\tilde{\mathbf{H}}\mathbf{B}^T\tilde{\mathbf{x}}_u + \mathrm{e}^{-j\hat{\varphi}_l}\mathbf{B}^T\mathbf{v} && (6.68)\\
&= \mathrm{e}^{-j\hat{\varphi}_l}\tilde{\mathbf{\Lambda}}^{(l)}\tilde{\mathbf{H}}\mathbf{G}_d\mathbf{d}^{(l)} + \mathbf{e}_{x_u}^{(l)} + \mathbf{e}_{x_p}^{(l)} + \mathrm{e}^{-j\hat{\varphi}_l}\mathbf{B}^T\mathbf{v}, && (6.69)\\
&= \mathrm{e}^{-j\hat{\varphi}_l}\tilde{\mathbf{\Lambda}}^{(l)}\tilde{\mathbf{H}}\mathbf{G}_d\mathbf{d}^{(l)} + \mathbf{e}^{(l)} + \mathrm{e}^{-j\hat{\varphi}_l}\mathbf{B}^T\mathbf{v}, && (6.70)
\end{aligned}$$

with $\mathbf{e}_{x_u}^{(l)}$ and $\mathbf{e}_{x_p}^{(l)}$ determining the subtraction error

$$\mathbf{e}^{(l)} = \mathbf{e}_{x_u}^{(l)} + \mathbf{e}_{x_p}^{(l)} \tag{6.71}$$

originating from the UW and the pilots, respectively. A data estimate follows as

$$\hat{\mathbf{d}}^{(l)} = \mathbf{E}_{\mathrm{LMMSE}}\tilde{\mathbf{y}}^{(l)}, \tag{6.72}$$

with $\mathbf{E}_{\mathrm{LMMSE}}$ as defined in (2.81). Fig. 6.9 comprises the constellation diagram for several instances of $\hat{\mathbf{d}}^{(l)}$ in case of $\mathbf{H} = \mathbf{I}$, $\epsilon = 0.1$, $\sigma_v^2 = 0$ and w.l.o.g. $\tilde{\mathbf{x}}_u = \mathbf{0}$. Additionally, perfect pilot subtraction with $\mathbf{e}_{x_p}^{(l)} = \mathbf{0}$ is assumed. As expected, the estimated data symbols form clouds centered around the elements of the transmission symbol alphabet, implying therefore that $\hat{\varphi}_l$ obtained according to (6.149) in Sec. 6.5 constitutes a pretty accurate estimate of φ_l. Completely unexpected though, assuming perfect estimation and derotating with $\hat{\varphi}_l = \varphi_l$ results in decentered symbol clouds, leading to a worse performance as with imperfect estimation (e.g., in terms of the mean squared error w.r.t. the transmitted symbols $\mathbf{d}^{(l)}$). This surprising behavior motivates more in-depth investigations. As part

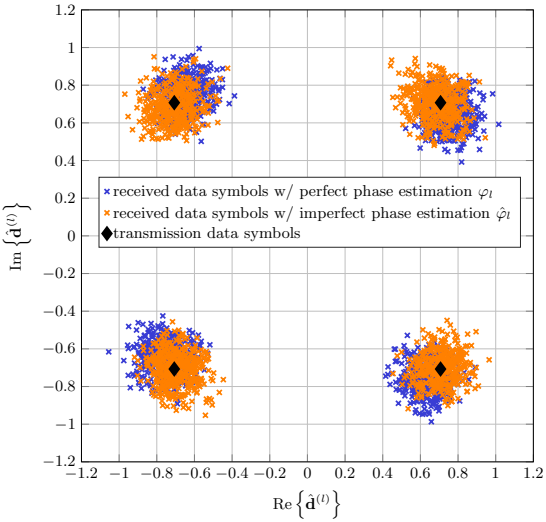


Figure 6.9.: Comparison of imperfect (pilot tone based) and perfect phase estimation $\hat{\varphi}_l$ based on the constellation diagram of the data symbol vector estimates $\hat{\mathbf{d}}^{(l)}$ for $d^{(l)}[k] \in 1/\sqrt{2}\{-1-j, -1+j, 1-j, 1+j\}$ $\tilde{\mathbf{H}} = \mathbf{I}$, $\tilde{\mathbf{x}}_u = \mathbf{0}$, $\mathbf{e}_{x_p}^{(l)} = \mathbf{0}$, $\epsilon = 0.1$, $l = 0 \ldots L-1$ and $\sigma_v^2 = 0$.

of a thought experiment let $\tilde{\mathbf{x}}_u = \mathbf{0}$, $\mathbf{e}_{x_p}^{(l)} = \mathbf{0}$ and $\sigma_v^2 = 0$, then (6.72) reduces to

$$\hat{\mathbf{d}}_{(\mathrm{exp})}^{(l)} = \mathrm{e}^{-j\hat{\varphi}_l} \mathbf{E}_{\mathrm{LMMSE}} \tilde{\mathbf{\Lambda}}^{(l)} \tilde{\mathbf{H}} \mathbf{G}_d \mathbf{d}^{(l)} \tag{6.73}$$

$$= \mathrm{e}^{-j\hat{\varphi}_l} \mathbf{E}_{\mathrm{LMMSE}} \mathrm{e}^{j\varphi_l} \tilde{\mathbf{\Lambda}}_{\mathrm{stat}} \tilde{\mathbf{H}} \mathbf{G}_d \mathbf{d}^{(l)}, \tag{6.74}$$

with

$$\tilde{\mathbf{\Lambda}}^{(l)} = \mathbf{B}^T \tilde{\mathbf{\Lambda}}'^{(l)} \mathbf{B} = \mathbf{B}^T \mathrm{e}^{j\varphi_l} \tilde{\mathbf{\Lambda}}'_{\mathrm{stat}} \mathbf{B} = \mathrm{e}^{j\varphi_l} \tilde{\mathbf{\Lambda}}_{\mathrm{stat}}. \tag{6.75}$$

A single symbol follows as

$$\hat{d}^{(l)}_{(\mathrm{exp})}[k] = \mathrm{e}^{-j\hat{\varphi}_l}\,[\mathbf{E}_{\mathrm{LMMSE}}]_{k,*}\,\mathrm{e}^{j\varphi_l}\,\tilde{\mathbf{\Lambda}}_{\mathrm{stat}}\tilde{\mathbf{H}}\left([\mathbf{G}_d]_{*,k}\,d^{(l)}[k] + \sum_{m=0,m\neq k}^{N_d-1}[\mathbf{G}_d]_{*,m}\,d^{(l)}[m]\right) \tag{6.76}$$

$$= \mathrm{e}^{-j\hat{\varphi}_l}\mathrm{e}^{j\varphi_l}\underbrace{\mathbf{e}_k^T\tilde{\mathbf{\Lambda}}_{\mathrm{stat}}\tilde{\mathbf{H}}\mathbf{g}_k}_{a_{\mathrm{SI},k}\mathrm{e}^{j\varphi_{\mathrm{SI},k}}}d^{(l)}[k] + \underbrace{\mathrm{e}^{-j\hat{\varphi}_l}\mathrm{e}^{j\varphi_l}\mathbf{e}_k^T\tilde{\mathbf{\Lambda}}_{\mathrm{stat}}\tilde{\mathbf{H}}\sum_{m=0,m\neq k}^{N_d-1}\mathbf{g}_m d^{(l)}[m]}_{\Delta_k} \tag{6.77}$$

$$= \mathrm{e}^{-j\hat{\varphi}_l}\mathrm{e}^{j\varphi_l}a_{\mathrm{SI},k}\mathrm{e}^{j\varphi_{\mathrm{SI},k}}d^{(l)}[k] + \Delta_k, \tag{6.78}$$

with $[\mathbf{E}_{\mathrm{LMMSE}}]_{k,*} = \mathbf{e}_k^T$ and $[\mathbf{G}_d]_{*,k} = \mathbf{g}_k$, whereas $[\cdot]_{k,*}$ and $[\cdot]_{*,k}$ denote the kth row and column of a matrix, respectively. There are now two effects degrading the performance of an estimate $\hat{d}^{(l)}_{(\mathrm{exp})}[k]$. First, Δ_k originates from the other data symbols as a consequence of intercarrier interference (compare with (6.36) and interdata interference) and has a similar impact as additive noise (with zero mean). Hence, the data symbol clouds in Fig. 6.9 may be enlarged in diameter, but the center of the cloud is not affected by Δ_k. Second, each transmitted data symbol $d^{(l)}[k]$ itself is corrupted, which can be modelled by a multiplication with a scalar

$$a_{\mathrm{SI},k}\mathrm{e}^{j\varphi_{\mathrm{SI},k}} = \mathbf{e}_k^T\tilde{\mathbf{\Lambda}}_{\mathrm{stat}}\tilde{\mathbf{H}}\mathbf{g}_k. \tag{6.79}$$

The actual value of this scalar depends on the CFO ϵ due to $\tilde{\mathbf{\Lambda}}_{\mathrm{stat}}$, the estimator, the channel instance and the generator matrix. The next paragraph will reveal that the corruption of $d^{(l)}[k]$ is in fact a result of self interference.

The amplitude $a_{\mathrm{SI},k}$ can be safely assumed to be always close to one and therefore negligible. This seems reasonable even without providing a detailed proof, considering that $\mathbf{e}_k^T$ normally compensates for $\tilde{\mathbf{H}}\mathbf{g}_k$ and $\tilde{\mathbf{\Lambda}}_{\mathrm{stat}}$ from (6.33) is similar to a diagonal matrix with entries close to one on the main diagonal. The phase rotation by $\varphi_{\mathrm{SI},k}$ causes the surprising effect visible in Fig. 6.9. In detail, a UW-OFDM symbol generator (modelled by $\mathbf{g}_k$) spreads a data symbol over several subcarriers. As well known, a CFO will introduce ICI (modelled by $\tilde{\mathbf{\Lambda}}_{\mathrm{stat}}$) into any OFDM system. In the context of UW-OFDM this means that the portions of a data symbol $d^{(l)}[k]$ spread over several subcarriers is leaked back due to ICI. In that sense, ICI in an UW-OFDM system causes self interference to a certain extent. Assuming now a perfect estimate $\hat{\varphi}_l = \varphi_l$, the phase offset $\varphi_{\mathrm{SI},k}$ in (6.78) remains and causes the rotation of the constellation diagram illustrated in Fig. 6.9. Since this rotation is not observable in case of $\hat{\varphi}_l \neq \varphi_l$, the deviation of the estimate $\hat{\varphi}_l$ from the actual phase φ_l seems to account for this effect. It is important to know at this point that in this work, $\hat{\varphi}_l$ is estimated on the basis of frequency pilot tones $\mathbf{p}$. The insertion of these pilot tones has already been discussed in chapter 5 and the

specific estimation algorithm is detailed later in Sec. 6.5, whereas knowledge of the latter is not required to understand the following explanation. Just studying the incorporation of pilot tones into the frequency domain by $\mathbf{G}_p\mathbf{p}$ in (5.6) shows that the pilot symbols – analogously to the data symbols – are spread over several subcarriers. Consequently, similar self interference effects due to intercarrier interference as in (6.78) can be expected, which will then be part of the estimate $\hat{\varphi}_l$. Without requiring any further information about the estimation algorithm presented later in Sec. 6.5, let us just utilize the model from (6.176) on page 163 that an estimate $\hat{\varphi}_l$ is given as

$$\hat{\varphi}_l = \varphi_l + \varphi_{\mathrm{pil}} + \Delta_l. \tag{6.80}$$

Here, φ_{pil} denotes the phase offset due to self interference of the pilot symbols[3] and Δ_l is a deviation corresponding to additive noise. With (6.80), the estimate of the kth data symbol in (6.78) can therefore be approximated with a remaining phase offset

$$\varphi_{\mathrm{off},k} = \varphi_{\mathrm{SI},k} - \varphi_{\mathrm{pil}} \tag{6.81}$$

as

$$\begin{aligned}
\hat{d}^{(l)}_{(\mathrm{exp})}[k] &= \mathrm{e}^{-j\hat{\varphi}_l}\mathrm{e}^{j\varphi_l}a_{\mathrm{SI},k}\mathrm{e}^{j\varphi_{\mathrm{SI},k}}d^{(l)}[k] + \Delta_k && (6.82)\\
&= \mathrm{e}^{-j\Delta_l}\mathrm{e}^{-j\varphi_l}\mathrm{e}^{j\varphi_l}\mathrm{e}^{j(\varphi_{\mathrm{SI},k}-\varphi_{\mathrm{pil}})}d^{(l)}[k] + \Delta_k && (6.83)\\
&\approx \mathrm{e}^{-j\varphi_l}\mathrm{e}^{j\varphi_l}\mathrm{e}^{j(\varphi_{\mathrm{SI},k}-\varphi_{\mathrm{pil}})}d^{(l)}[k] && (6.84)\\
&= \mathrm{e}^{j\varphi_{\mathrm{off},k}}d^{(l)}[k]. && (6.85)
\end{aligned}$$

According to (6.85), each data symbol experiences a different phase offset $\varphi_{\mathrm{off},k}$, which therefore requires a data symbol individual compensation. Experiments not detailed here proved a simplification of the phase offset model

$$\varphi_{\mathrm{off}} = \frac{1}{N_d}\sum_{k=0}^{N_d-1}\varphi_{\mathrm{off},k} \tag{6.86}$$

and a corresponding simplified compensation to have almost no degrading impact on the performance, leading to

$$\begin{aligned}
\hat{d}^{(l)}_{(\mathrm{exp})}[k] &= \mathrm{e}^{j\varphi_{\mathrm{off},k}}d^{(l)}[k] && (6.87)\\
&\approx \mathrm{e}^{j\varphi_{\mathrm{off}}}d^{(l)}[k]. && (6.88)
\end{aligned}$$

[3] Note that self interference is a source for φ_{pil} in any UW-OFDM system. In case of a non-zero UW and depending on the estimation model used for φ_l, the UW can be a second source for the resulting phase rotation φ_{pil}. To account for this optional source, the very general subcript $(\cdot)_{\mathrm{pil}}$ for 'pilots' instead of $(\cdot)_{\mathrm{SI}}$ for 'self interference' has been chosen. Details can be found in Sec. 6.5.

Based on the insights gained from this thought experiment, the estimate of a data vector is given as

$$\hat{\mathbf{d}}^{(l)} = \mathrm{e}^{-j\hat{\varphi}_{\mathrm{off}}}\mathbf{E}_{\mathrm{LMMSE}}\tilde{\mathbf{y}}^{(l)} \tag{6.89}$$

$$= \mathrm{e}^{-j\hat{\varphi}_{\mathrm{off}}}\mathrm{e}^{-j\hat{\varphi}_l}\mathbf{E}_{\mathrm{LMMSE}}\tilde{\mathbf{\Lambda}}^{(l)}\tilde{\mathbf{H}}\mathbf{G}_d\mathbf{d}^{(l)} + \mathrm{e}^{-j\hat{\varphi}_{\mathrm{off}}}\mathbf{E}_{\mathrm{LMMSE}}\mathbf{e}^{(l)} + \mathrm{e}^{-j\hat{\varphi}_{\mathrm{off}}}\mathrm{e}^{-j\hat{\varphi}_l}\mathbf{E}_{\mathrm{LMMSE}}\mathbf{B}^T\mathbf{v}. \tag{6.90}$$

The estimate $\hat{\varphi}_{\mathrm{off}}$ required for offset compensation is assumed to be given at this point, details are provided in (6.190) within Sec. 6.5.

Fig. 6.10 visualizes the Bayesian MSE per data symbol

$$\theta_d = \frac{1}{L}\sum_{l=0}^{L-1}\frac{1}{N_d}\mathrm{E}\left\{\left\|\hat{\mathbf{d}}^{(l)} - \mathbf{d}^{(l)}\right\|_2^2\right\} \tag{6.91}$$

as a function of the CFO ϵ. The MSE has been evaluated for 10^4 multipath chan-

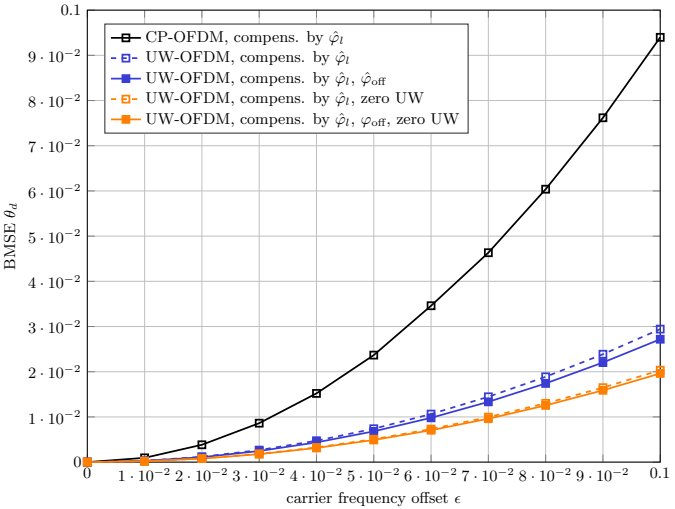


Figure 6.10.: Carrier frequency offset induced Bayesian MSE θ_d in UW-OFDM and CP-OFDM after derotating by $\hat{\varphi}_l$ as CFO compensation means. The estimate $\hat{\mathbf{d}}^{(l)}$ has been obtained according to (6.89). Investigated scenario: $\mathbf{G}'_d$, CAZAC sequence as UW, and multipath environment with $\tau_{\mathrm{RMS}} = 100\,\mathrm{ns}$.

nel instances according to Sec. 2.4 and averaged over $L = 200$ OFDM symbols. The presented results are obtained for $\mathbf{G}'_d$ as generator matrix and a CAZAC sequence as UW $\mathbf{x}_u$ [65]. Note that the results for different generator matrices like

$\mathbf{G}_d''$ and other UWs are very similar and therefore not explicitly shown. In order to better assess the performance, two main reference curves are illustrated. The upper bound determines the performance of the reference CP-OFDM system. At this point it is already fair to say that UW-OFDM outperforms CP-OFDM in terms of CFO robustness in any way. One main reason is a stronger inter-carrier/interdata interference in CP-OFDM, a second originates from the missing redundancy.

The lower bound represents the case of a zero UW, perfect pilot subtraction with $\mathbf{e}_{x_p}^{(l)} = \mathbf{0}$ and perfect offset compensation with $\hat{\varphi}_{\text{off}} = \varphi_{\text{off}}$. The dashed line with square symbols depicts the same but without offset compensation. In case of a zero UW (and only then), the gain due to additional derotation (even if φ_{off} is perfectly known) is thus rather limited. For all setups investigated in this work, the phase offset φ_{off} in case of a zero UW is rather small and correcting it delivers only a minor performance improvement. Only for some very special channel realizations, the effort would pay off. However, this statement cannot be generalized and has to be verified for each new setup. In order to avoid confusions regarding the UW-OFDM setup used for the bound, please note that the case of a zero UW and $\mathbf{e}_{x_p}^{(l)} = \mathbf{0}$ slightly differs from a non-zero UW-OFDM system with $\mathbf{e}_{x_u}^{(l)} = \mathbf{0}$ and $\mathbf{e}_{x_p}^{(l)} = \mathbf{0}$. Both share the ideal assumption of $\mathbf{e}_{x_u}^{(l)} = \mathbf{0}$, but in the non-zero case, the UW will influence the estimation of φ_l. Although there is only a minor impact on the estimation performance (as shown later in Sec. 6.5) and therefore on the resulting MSE in Fig. 6.10, the dedicated purpose of this bound is to show the zero UW case.

The performance of an UW-OFDM system with a non-zero UW-OFDM lies between the bounds, whereas Fig. 6.10 incorporates one curve with and one without offset compensation. Note that here the applied angle is only an estimate $\hat{\varphi}_{\text{off}}$, but it almost yields the same performance as the utilization of the true angle φ_{off}. The latter case is not in the figure to keep it as simple as possible. Clearly, the gain of offset compensation is higher in the non-zero UW than in the zero UW case. Furthermore, it increases with increasing CFO. Nonetheless, there is still room for improvement, even with offset compensation:

- The first performance gap between the non-zero UW case and the lower bound originates from subtracting the wrong UW and pilot portion leading to the error $\mathbf{e}^{(l)} = \mathbf{e}_{x_u}^{(l)} + \mathbf{e}_{x_p}^{(l)}$.

- The second performance gap between the lower bound and the abscissa is due to the incomplete CFO cancelation by only derotating with $\hat{\varphi}_l$.

Both gaps will be analyzed in the following and alternatives are pointed out.

6.3. Impact of Imperfect UW and Pilot Offset Subtraction

In order to minimize $\mathbf{e}^{(l)}$ in (6.71) further, let us additionally investigate the subtraction of an offset

- $\mathbf{x}_{\mathrm{off},i} = \mathrm{e}^{-j\hat{\varphi}_l}\hat{\tilde{\mathbf{\Lambda}}}^{(l)}\tilde{\mathbf{H}}\left(\mathbf{B}^T\tilde{\mathbf{x}}_u + \mathbf{G}_p\mathbf{p}\right)$ and
- $\mathbf{x}_{\mathrm{off},c} = \mathrm{e}^{-j\hat{\varphi}_{\mathrm{pil}}}\hat{\tilde{\mathbf{\Lambda}}}_{\mathrm{stat}}\tilde{\mathbf{H}}\left(\mathbf{B}^T\tilde{\mathbf{x}}_u + \mathbf{G}_p\mathbf{p}\right)$

instead of $\mathbf{x}_{\mathrm{off}} = \tilde{\mathbf{H}}\left(\mathbf{B}^T\tilde{\mathbf{x}}_u + \mathbf{G}_p\mathbf{p}\right)$ from (6.65). In the first approach with $\mathbf{x}_{\mathrm{off},i}$, the term to be subtracted is specifically tailored to each individual OFDM symbol l. The dependence on l requires an update for every symbol and thus denotes a certain computational overhead. Note that the estimate $\hat{\tilde{\mathbf{\Lambda}}}^{(l)}$ is calculated based on a estimated CFO

$$\hat{\epsilon} = \epsilon + \Delta_{\epsilon_l} \tag{6.92}$$

with estimation error Δ_{ϵ_l}, whereas $\hat{\epsilon}$ is derived from $\hat{\varphi}_l$ as presented later in Sec. 6.5. It suffices to treat $\hat{\epsilon}$ as given, the actual estimation algorithm is not of relevance in the following. This approach takes the applied derotation angle $\hat{\varphi}_l$ of each individual OFDM symbol into account, in this sense utilizing all the knowledge available and thus expected to deliver the best results.

In the second approach with $\mathbf{x}_{\mathrm{off},c}$ the subtracted term is common for all OFDM symbols. It is built on the hypothetical assumption of having a noise free estimate $\hat{\varphi}_l$. Noise free in this context means that φ_l is estimated exactly up to the constant offset φ_{pil}, cf. (6.80). It follows that $\hat{\varphi}_l = \varphi_l + \varphi_{\mathrm{pil}}$ yielding $\mathrm{e}^{-j(\varphi_l+\varphi_{\mathrm{pil}})}\tilde{\mathbf{\Lambda}}^{(l)} = \mathrm{e}^{-j\varphi_{\mathrm{pil}}}\tilde{\mathbf{\Lambda}}_{\mathrm{stat}}$, therefore dissolving the dependence on l and thus reducing complexity. In other words, the UW and pilot portion do not have to be updated for each OFDM symbol, when assuming constant CFO and channel conditions. In reality, a noise free estimate is of course not possible and thus only an estimate $\hat{\varphi}_{\mathrm{pil}}$ of φ_{pil} is available for calculating $\mathbf{x}_{\mathrm{off},c}$. Further, $\hat{\tilde{\mathbf{\Lambda}}}_{\mathrm{stat}}$ is based on $\hat{\epsilon}$ and approximates $\mathbf{B}^T\tilde{\mathbf{\Lambda}}'_{\mathrm{stat}}\mathbf{B}$, cf. (6.34).

Contrary to any intuitive assumption, the second and less complex method with $\mathbf{x}_{\mathrm{off},c}$ reaches almost the performance bound of an UW-OFDM system with zero word and perfect pilot subtraction, whereas the first one with constantly updated subtraction term $\mathbf{x}_{\mathrm{off},i}$ shows devastating performance (Fig. 6.11). Since the latter uses all the information available, it seemed nearby that it performs better than the less complex method and closes the remaining little gap to the bound. This motivates a more in-depth analysis of the resulting error $\mathbf{e}^{(l)}$ for both methods.

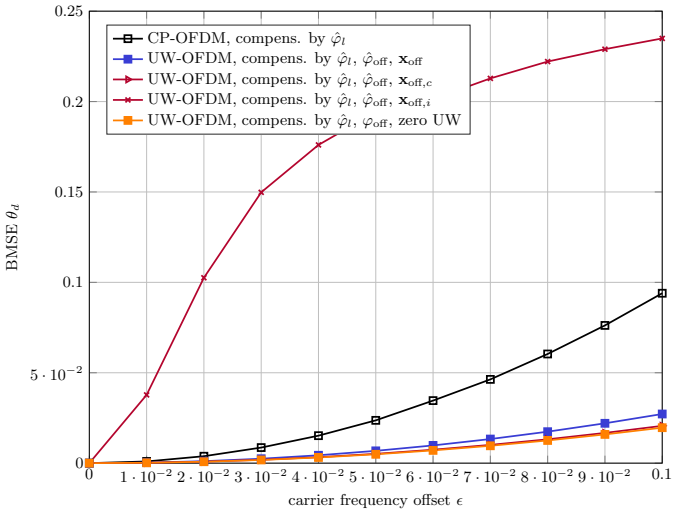


Figure 6.11.: Carrier frequency offset induced Bayesian MSE θ_d in UW-OFDM and CP-OFDM after derotating by $\hat{\varphi}_l$ as CFO compensation means. Investigation of different UW and pilot subtraction techniques that incorporate different levels of CFO information. The estimate $\hat{\mathbf{d}}^{(l)}$ has been obtained according to (6.89), with $\tilde{\mathbf{y}}^{(l)} = \tilde{\mathbf{y}}_c^{(l)} - \mathbf{x}_{\text{off}}$, $\tilde{\mathbf{y}}^{(l)} = \tilde{\mathbf{y}}_c^{(l)} - \mathbf{x}_{\text{off},i}$ and $\tilde{\mathbf{y}}^{(l)} = \tilde{\mathbf{y}}_c^{(l)} - \mathbf{x}_{\text{off},c}$. Investigated scenario: $\mathbf{G}'_d$, CAZAC sequence as UW, and multipath environment with $\tau_{\text{RMS}} = 100\,\text{ns}$.

6.3.1. Individual UW and Pilot Offset Compensation

In the first approach with a subtraction of $\mathbf{x}_{\text{off},i}$, the error vector in (6.71) reads

$$\mathbf{e}^{(l)} = \mathrm{e}^{-j\hat{\varphi}_l}\tilde{\mathbf{\Lambda}}^{(l)}\tilde{\mathbf{H}}\left(\mathbf{B}^T\tilde{\mathbf{x}}_u + \mathbf{G}_p\mathbf{p}\right) - \mathrm{e}^{-j\hat{\varphi}_l}\hat{\tilde{\mathbf{\Lambda}}}^{(l)}\tilde{\mathbf{H}}\left(\mathbf{B}^T\tilde{\mathbf{x}}_u + \mathbf{G}_p\mathbf{p}\right) \tag{6.93}$$

$$= \mathrm{e}^{-j\hat{\varphi}_l}\tilde{\mathbf{\Lambda}}^{(l)}\mathbf{x}_{\text{off}} - \mathrm{e}^{-j\hat{\varphi}_l}\hat{\tilde{\mathbf{\Lambda}}}^{(l)}\mathbf{x}_{\text{off}}, \tag{6.94}$$

with $\mathbf{x}_{\text{off}}$ from (6.65). For the sake of simplicity and w.l.o.g., the following derivations focus only on the kth element $e^{(l)}[k]$ instead of the whole error vector $\mathbf{e}^{(l)}$. With (6.5) and (6.54), $e^{(l)}[k]$ is written as

$$e^{(l)}[k] = \mathrm{e}^{-j\hat{\varphi}_l}\left[\tilde{\mathbf{\Lambda}}^{(l)}\right]_{k,*}\mathbf{x}_{\text{off}} - \mathrm{e}^{-j\hat{\varphi}_l}\left[\hat{\tilde{\mathbf{\Lambda}}}^{(l)}\right]_{k,*}\mathbf{x}_{\text{off}} \tag{6.95}$$

$$= \mathrm{e}^{-j\hat{\varphi}_l}\mathrm{e}^{j\psi_l}\left[\tilde{\mathbf{\Lambda}}\right]_{k,*}\mathbf{x}_{\text{off}} - \mathrm{e}^{-j\hat{\varphi}_l}\mathrm{e}^{j\frac{2\pi\hat{\epsilon}(Nl+N_u)}{N}}\left[\hat{\tilde{\mathbf{\Lambda}}}\right]_{k,*}\mathbf{x}_{\text{off}}. \tag{6.96}$$

For further analysis, let us investigate $\tilde{\mathbf{\Lambda}}$ and $\hat{\tilde{\mathbf{\Lambda}}}$ more in detail. $\tilde{\mathbf{\Lambda}}$ originates from $\tilde{\mathbf{\Lambda}}' \in \mathbb{C}^{N\times N}$ defined as

$$\left[\tilde{\mathbf{\Lambda}}'\right]_{k,m} = \frac{1}{N}\sum_{u=0}^{N-1} \mathrm{e}^{j\frac{2\pi}{N}(m+\epsilon-k)u} \qquad k=0\dots N-1;\, m=0\dots N-1 \tag{6.97}$$

$$= \frac{1}{N}\sum_{u=0}^{N-1} \mathrm{e}^{j\frac{2\pi\epsilon u}{N}} \mathrm{e}^{j\frac{2\pi(m-k)u}{N}} \tag{6.98}$$

by simply deleting rows and columns corresponding to zero subcarriers. Alternatively, this can be expressed as $\tilde{\mathbf{\Lambda}} = \mathbf{B}^T\tilde{\mathbf{\Lambda}}'\mathbf{B}$ as already demonstrated in (6.54), the matrix $\hat{\tilde{\mathbf{\Lambda}}}$ is analogously derived from $\hat{\tilde{\mathbf{\Lambda}}}'$. The subsequent derivations are based on $\tilde{\mathbf{\Lambda}}'$ and $\hat{\tilde{\mathbf{\Lambda}}}'$, the link to $\tilde{\mathbf{\Lambda}}$ and $\hat{\tilde{\mathbf{\Lambda}}}$ is provided at the end. A notation like in (6.98) enables a compact description of $\tilde{\mathbf{\Lambda}}$ and $\hat{\tilde{\mathbf{\Lambda}}}$, but further simplifications of the error $e^{(l)}[k]$ are rather limited. Introducing the vectors $\mathbf{a}^T \in \mathbb{C}^{1\times N}$ and $\mathbf{m}_m^{(k)} \in \mathbb{C}^{N\times 1}$ with

$$\mathbf{a}^T = \mathrm{e}^{j\frac{2\pi\epsilon\mathbf{u}^T}{N}} \qquad \mathbf{u}^T = [0,1,\dots N-1] \tag{6.99}$$

$$\mathbf{m}_m^{(k)} = \frac{1}{N}\mathrm{e}^{j\frac{2\pi(m-k)\mathbf{u}}{N}} \tag{6.100}$$

delivers $\left[\tilde{\mathbf{\Lambda}}'\right]_{k,m} = \mathbf{a}^T\mathbf{m}_m^{(k)}$. Consequently, the kth row of $\tilde{\mathbf{\Lambda}}'$ is obtained as

$$\left[\tilde{\mathbf{\Lambda}}'\right]_{k,*} = \mathbf{a}^T\mathbf{M}'^{(k)}, \tag{6.101}$$

where

$$\mathbf{M}'^{(k)} = \left[\mathbf{m}_0^{(k)} \quad \mathbf{m}_1^{(k)} \dots \mathbf{m}_{N-1}^{(k)}\right] \in \mathbb{C}^{N\times N}. \tag{6.102}$$

In a very similar fashion, it follows that

$$\left[\hat{\tilde{\mathbf{\Lambda}}}'\right]_{k,*} = \hat{\mathbf{a}}^T\mathbf{M}'^{(k)}, \tag{6.103}$$

with $\hat{\epsilon} = \epsilon + \Delta_{\epsilon_l}$ and

$$\hat{\mathbf{a}}^T = \mathrm{e}^{j\frac{2\pi\hat{\epsilon}\mathbf{u}^T}{N}} = \mathrm{e}^{j\frac{2\pi(\epsilon+\Delta_{\epsilon_l})\mathbf{u}^T}{N}} \qquad \mathbf{u}^T = [0,1,\dots N-1]. \tag{6.104}$$

$\tilde{\mathbf{\Lambda}}'$ and $\hat{\tilde{\mathbf{\Lambda}}}'$ in (6.101) and (6.103) have now been disaggregated in a constant term $\mathbf{M}'^{(k)}$, which both matrices share together, and a variable and different term $\mathbf{a}^T$ and $\hat{\mathbf{a}}^T$, respectively. Before utilizing this notation for the error $e^{(l)}[k]$ described in (6.96), there is still the relationship $\tilde{\mathbf{\Lambda}} = \mathbf{B}^T\tilde{\mathbf{\Lambda}}'\mathbf{B}$ to resolve. Based on (6.101), $\tilde{\mathbf{\Lambda}}'$ translates to

$$\tilde{\mathbf{\Lambda}}' = \begin{bmatrix} \mathbf{a}^T\mathbf{M}'^{(0)} \\ \mathbf{a}^T\mathbf{M}'^{(1)} \\ \vdots \\ \mathbf{a}^T\mathbf{M}'^{(N-1)} \end{bmatrix}. \tag{6.105}$$

Multiplying $\tilde{\mathbf{\Lambda}}'\mathbf{B}$ yields $\mathbf{a}^T\mathbf{M}'^{(k)}\mathbf{B} = \mathbf{a}^T\mathbf{M}^{(k)}$ for all $k = 0, 1, \dots N-1$, whereas $\mathbf{M}^{(k)} \in \mathbb{C}^{N\times(N-N_z)}$ is a matrix consisting only of columns with indices corresponding to non-zero subcarriers

$$\mathbf{M}^{(k)} = \begin{bmatrix} \mathbf{m}_{i_0}^{(k)} & \mathbf{m}_{i_1}^{(k)} \dots \mathbf{m}_{i_{N-N_z-1}}^{(k)} \end{bmatrix}. \tag{6.106}$$

The subscript i_k addresses the kth element of the ordered index set of all non-zero subcarriers $\mathcal{I}_{nz,o}$ defined in (3.41). Finally, $\tilde{\mathbf{\Lambda}}$ is expressed as

$$\tilde{\mathbf{\Lambda}} = \mathbf{B}^T\tilde{\mathbf{\Lambda}}'\mathbf{B} = \mathbf{B}^T \begin{bmatrix} \mathbf{a}^T\mathbf{M}^{(0)} \\ \mathbf{a}^T\mathbf{M}^{(1)} \\ \vdots \\ \mathbf{a}^T\mathbf{M}^{(N-1)} \end{bmatrix} = \begin{bmatrix} \mathbf{a}^T\mathbf{M}^{(i_0)} \\ \mathbf{a}^T\mathbf{M}^{(i_1)} \\ \vdots \\ \mathbf{a}^T\mathbf{M}^{(i_{N-N_z-1})} \end{bmatrix}, \tag{6.107}$$

and the kth row of $\tilde{\mathbf{\Lambda}}$ and $\hat{\tilde{\mathbf{\Lambda}}}$ reads

$$\left[\hat{\tilde{\mathbf{\Lambda}}}\right]_{k,*} = \hat{\mathbf{a}}^T\mathbf{M}^{(i_k)} \tag{6.108}$$

$$\left[\tilde{\mathbf{\Lambda}}\right]_{k,*} = \mathbf{a}^T\mathbf{M}^{(i_k)}. \tag{6.109}$$

Inserting (6.108), (6.109) and the estimate $\hat{\varphi}_l$ from (6.80) into (6.96), the error due to subtracting the wrong UW and pilot term follows as

$$e^{(l)}[k] = \mathrm{e}^{-j\hat{\varphi}_l}\left(\mathrm{e}^{j\psi_l}\mathbf{a}^T\mathbf{M}^{(i_k)}\mathbf{x}_{\mathrm{off}} - \mathrm{e}^{j\frac{2\pi\hat{\epsilon}(Nl+N_u)}{N}}\hat{\mathbf{a}}^T\mathbf{M}^{(i_k)}\mathbf{x}_{\mathrm{off}}\right) \tag{6.110}$$

$$= \mathrm{e}^{-j\left(\varphi_l+\varphi_{\mathrm{pil}}+\Delta_l\right)}\left(\mathrm{e}^{j\psi_l}\mathbf{a}^T - \mathrm{e}^{j\frac{2\pi(\epsilon+\Delta_{\epsilon_l})(Nl+N_u)}{N}}\hat{\mathbf{a}}^T\right)\mathbf{M}_k\mathbf{x}_{\mathrm{off}} \tag{6.111}$$

$$= \mathrm{e}^{-j\left(\psi+\varphi_{\mathrm{pil}}+\Delta_l\right)}\left(\mathbf{a}^T - \mathrm{e}^{j\frac{2\pi\Delta_{\epsilon_l}(Nl+N_u)}{N}}\hat{\mathbf{a}}^T\right)\mathbf{M}_k\mathbf{x}_{\mathrm{off}}, \tag{6.112}$$

the last step as a consequence of $\psi_l - \varphi_l = -\psi$, cf. (6.31). For reasons of compact notation, the nomenclature $\mathbf{M}_k := \mathbf{M}^{(i_k)}$ has been introduced. With the definition of $\mathbf{a}^T$ and $\hat{\mathbf{a}}^T$ according to (6.99) and (6.104), the error corresponds to

$$e^{(l)}[k] = \left(\mathrm{e}^{j\frac{2\pi\epsilon\mathbf{u}^T}{N}} - \mathrm{e}^{j\frac{2\pi\Delta_{\epsilon_l}(Nl+N_u)}{N}}\mathrm{e}^{j\frac{2\pi\hat{\epsilon}\mathbf{u}^T}{N}}\right)\mathrm{e}^{-j\left(\psi+\varphi_{\mathrm{pil}}+\Delta_l\right)}\mathbf{M}_k\mathbf{x}_{\mathrm{off}} \tag{6.113}$$

$$= \left(\mathrm{e}^{j\frac{2\pi\epsilon\mathbf{u}^T}{N}} - \mathrm{e}^{j\frac{2\pi\Delta_{\epsilon_l}(Nl+N_u)}{N}}\mathrm{e}^{j\frac{2\pi(\epsilon+\Delta_{\epsilon_l})\mathbf{u}^T}{N}}\right)\mathbf{x}'_{\mathrm{off}} \tag{6.114}$$

$$= \sum_{u=0}^{N-1}\left(\mathrm{e}^{j\frac{2\pi\epsilon u}{N}} - \mathrm{e}^{j\frac{2\pi\Delta_{\epsilon_l}(Nl+N_u)}{N}}\mathrm{e}^{j\frac{2\pi(\epsilon+\Delta_{\epsilon_l})u}{N}}\right)x'_{\mathrm{off}}[u] \tag{6.115}$$

$$= \sum_{u=0}^{N-1}\left(1 - \mathrm{e}^{j\frac{2\pi\Delta_{\epsilon_l}(Nl+N_u+u)}{N}}\right)\mathrm{e}^{j\frac{2\pi\epsilon u}{N}}x'_{\mathrm{off}}[u]. \tag{6.116}$$

There are now three terms influencing $e^{(l)}[k]$. The uth element $x'_{\text{off}}[u]$ of

$$\mathbf{x}'_{\text{off}} = \mathrm{e}^{-j\left(\psi+\varphi_{\text{pil}}+\Delta_l\right)}\mathbf{M}_k\mathbf{x}_{\text{off}} \tag{6.117}$$

cannot be deterministically determined due to the random components Δ_l (6.80) and $\tilde{\mathbf{H}}$. Each element $x'_{\text{off}}[u]$ experiences a different phase rotation $\mathrm{e}^{j\frac{2\pi\epsilon u}{N}}$ dependent on the index u. Nonetheless, the rotations do not change the statistical characteristics of the elements $x'_{\text{off}}[u]$ and can be neglected for this kind of analysis. Independent of the concrete realization of $x'_{\text{off}}[u]$ though, the expression in parentheses in (6.116) works as a scaling factor and the error vanishes, as soon as the expression in the parentheses approaches zero. Following this idea and taking into account that $1 = \mathrm{e}^{j2\pi m}$ for $m \in \mathbb{Z}$ yields

$$\left(1 - \mathrm{e}^{j\frac{2\pi\Delta_{\epsilon_l}(Nl+N_u+u)}{N}}\right) \stackrel{!}{=} 0 \tag{6.118}$$

$$\frac{2\pi}{N}\Delta_{\epsilon_l}\left(Nl + N_u + nu\right) = 2\pi m \tag{6.119}$$

$$l = \left[\frac{1}{\Delta_{\epsilon_l}}m - \left(\frac{N_u}{N} + \frac{u}{N}\right)\right]_R, \tag{6.120}$$

whereas $[\cdot]_R$ denotes rounding to the nearest integer. Obviously, the scaling factor slightly varies within u and the optimum l is thus ambiguous. Inserting the borders of n delivers

$$l_{u=0} = \left[\frac{1}{\Delta_{\epsilon_l}}m - \left(\frac{N_u}{N} + \frac{0}{N}\right)\right]_R = \frac{1}{\Delta_{\epsilon_l}}m \tag{6.121}$$

$$l_{u=N-1} = \left[\frac{1}{\Delta_{\epsilon_l}}m - \left(\frac{N_u}{N} + \frac{N-1}{N}\right)\right]_R \tag{6.122}$$

$$= \left[\frac{1}{\Delta_{\epsilon_l}}m - 1 - \frac{N_u+1}{N}\right]_R = \frac{1}{\Delta_{\epsilon_l}}m - 1. \tag{6.123}$$

The results suggest that the minimum error is achieved for either $l = \frac{1}{\epsilon}m$ or $l = \frac{1}{\epsilon}m - 1$, simulations prove the first case to deliver the minimum. Despite of this, a big problem is the dependence of the scaling factor and thus of the error on the OFDM symbol index l. With a period of $\frac{1}{\Delta_{\epsilon_l}}$ the scaling factor varies between 0 and 2, the latter causing the devastating performance illustrated in Fig. 6.11.

6.3.2. Common UW and Pilot Offset Compensation

A second alternative approach for CFO compensation is subtracting $\mathbf{x}_{\text{off},c}$, which is based on the assumption of a noise free estimate $\hat{\varphi}$ with $\Delta_l = 0$ in (6.80),

i.e.,

$$\hat{\varphi}_l \stackrel{!}{=} \varphi_l + \varphi_{\text{pil}}. \tag{6.124}$$

The remaining additive term of the UW and pilots (assuming a noise free estimate $\hat{\varphi}_l$) together with (6.75) would then be

$$\mathrm{e}^{-j\hat{\varphi}_l}\tilde{\mathbf{\Lambda}}^{(l)}\mathbf{x}_{\text{off}} = \mathrm{e}^{-j\left(\varphi_l+\varphi_{\text{pil}}\right)}\mathrm{e}^{j\psi_l}\tilde{\mathbf{\Lambda}}\mathbf{x}_{\text{off}} \tag{6.125}$$

$$= \mathrm{e}^{-j\varphi_{\text{pil}}}\mathrm{e}^{-j\psi}\tilde{\mathbf{\Lambda}}\mathbf{x}_{\text{off}} = \mathrm{e}^{-j\varphi_{\text{pil}}}\tilde{\mathbf{\Lambda}}_{\text{stat}}\mathbf{x}_{\text{off}}. \tag{6.126}$$

Following this approach, the portion to be subtracted is now static, meaning that it does not change with the OFDM symbol index l anymore. In reality though, a noise free estimate of ϵ and φ_l is not available, requiring an estimate

$$\mathrm{e}^{-j\hat{\varphi}_{\text{pil}}}\hat{\tilde{\mathbf{\Lambda}}}_{\text{stat}}\mathbf{x}_{\text{off}} = \mathrm{e}^{-j\hat{\varphi}_{\text{pil}}}\mathrm{e}^{-j\frac{\pi}{N}\hat{\epsilon}(N-1)}\hat{\tilde{\mathbf{\Lambda}}}\mathbf{x}_{\text{off}}. \tag{6.127}$$

Consequently, the error after subtraction can be described as

$$\mathbf{e}^{(l)} = \mathrm{e}^{-j\hat{\varphi}_l}\tilde{\mathbf{\Lambda}}^{(l)}\mathbf{x}_{\text{off}} - \mathrm{e}^{-j\hat{\varphi}_{\text{pil}}}\hat{\tilde{\mathbf{\Lambda}}}_{\text{stat}}\mathbf{x}_{\text{off}}. \tag{6.128}$$

The kth element reads

$$e^{(l)}[k] = \mathrm{e}^{-j\hat{\varphi}_l}\left[\tilde{\mathbf{\Lambda}}^{(l)}\right]_{k,*}\mathbf{x}_{\text{off}} - \mathrm{e}^{-j\hat{\varphi}_{\text{pil}}}\left[\hat{\tilde{\mathbf{\Lambda}}}_{\text{stat}}\right]_{k,*}\mathbf{x}_{\text{off}} \tag{6.129}$$

$$= \mathrm{e}^{-j\hat{\varphi}_l}\mathrm{e}^{j\psi_l}\left[\tilde{\mathbf{\Lambda}}\right]_{k,*}\mathbf{x}_{\text{off}} - \mathrm{e}^{-j\hat{\varphi}_{\text{pil}}}\mathrm{e}^{-j\frac{2\pi\hat{\epsilon}(N-1)}{2N}}\left[\hat{\tilde{\mathbf{\Lambda}}}\right]_{k,*}\mathbf{x}_{\text{off}}. \tag{6.130}$$

Please bear in mind that the definition of $\hat{\varphi}_l$ in (6.124) has just been part of a thought experiment to obtain $\tilde{\mathbf{\Lambda}}_{\text{stat}}$ in (6.126), for the following derivation of the actual error, the definition according to (6.80) with $\Delta_l \neq 0$ applies. Carrying out similar operations as in (6.110)-(6.116), the error follows with $\mathbf{x}'_{\text{off}}$ from (6.117) as

$$e^{(l)}[k] = \mathrm{e}^{-j\left(\varphi_l+\varphi_{\text{pil}}+\Delta_l\right)}\mathrm{e}^{j\psi_l}\mathbf{a}^T\mathbf{M}_k\mathbf{x}_{\text{off}}$$

$$- \mathrm{e}^{-j\hat{\varphi}_{\text{pil}}}\mathrm{e}^{-j\frac{2\pi(\epsilon+\Delta_{\epsilon l})(N-1)}{2N}}\hat{\mathbf{a}}^T\mathbf{M}_k\mathbf{x}_{\text{off}} \tag{6.131}$$

$$= \left(\mathrm{e}^{-j\left(\psi+\varphi_{\text{pil}}+\Delta_l\right)}\mathbf{a}^T - \mathrm{e}^{-j\hat{\varphi}_{\text{pil}}}\mathrm{e}^{-j\frac{2\pi(\epsilon+\Delta_{\epsilon l})(N-1)}{2N}}\hat{\mathbf{a}}^T\right)\mathbf{M}_k\mathbf{x}_{\text{off}} \tag{6.132}$$

$$= \left(\mathbf{a}^T - \mathrm{e}^{-j\hat{\varphi}_{\text{pil}}}\mathrm{e}^{j\left(\varphi_{\text{pil}}+\Delta_l\right)}\mathrm{e}^{-j\frac{2\pi\Delta_{\epsilon l}(N-1)}{2N}}\hat{\mathbf{a}}^T\right)\mathrm{e}^{-j\left(\varphi_{\text{pil}}+\Delta_l\right)}$$

$$\cdot\, \mathrm{e}^{-j\psi}\mathbf{M}_k\mathbf{x}_{\text{off}} \tag{6.133}$$

$$= \left(\mathrm{e}^{j\frac{2\pi}{N}\epsilon\mathbf{u}^T} - \mathrm{e}^{-j\left(\hat{\varphi}_{\text{pil}}-\varphi_{\text{pil}}\right)}\mathrm{e}^{j\Delta_l}\mathrm{e}^{-j\frac{2\pi\Delta_{\epsilon l}(N-1)}{2N}}\mathrm{e}^{j\frac{2\pi}{N}(\epsilon+\Delta_{\epsilon l})\mathbf{u}^T}\right)\mathbf{x}'_{\text{off}} \tag{6.134}$$

$$= \sum_{u=0}^{N-1}\left(\mathrm{e}^{j\frac{2\pi}{N}\epsilon u} - \mathrm{e}^{-j\left(\hat{\varphi}_{\text{pil}}-\varphi_{\text{pil}}\right)}\mathrm{e}^{j\Delta_l}\mathrm{e}^{-j\frac{2\pi\Delta_{\epsilon l}(N-1)}{2N}}\mathrm{e}^{j\frac{2\pi}{N}(\epsilon+\Delta_{\epsilon l})u}\right)x'_{\text{off}}[u] \tag{6.135}$$

$$= \sum_{u=0}^{N-1} \left(1 - \mathrm{e}^{-j\left(\hat{\varphi}_{\mathrm{pil}} - \varphi_{\mathrm{pil}}\right)} \mathrm{e}^{j\Delta_l} \mathrm{e}^{j\frac{2\pi}{N}\Delta_{\epsilon_l}\left(u - \frac{N-1}{2}\right)}\right) \mathrm{e}^{j\frac{2\pi}{N}\epsilon u} x'_{\mathrm{off}}[u]. \tag{6.136}$$

Different to the first alternative CFO compensation approach in Sec. 6.3.1, the scaling factor within the parentheses in (6.136) does not vary with the OFDM symbol index l anymore, it only depends on the estimation errors Δ_l and Δ_{ϵ_l}. Assuming moderate estimation errors, the scaling factor always remains close to 0 and never reaches the upper bound of 2, as it does in the first approach.

In a very surprising way, taking into account all available information does not deliver the best results in this case, the MSE of this simplified method is significantly lower than of the more costly approach from Sec. 6.3.1. Moreover, the simplified approach almost reaches the performance bound of perfect subtraction (Fig. 6.11). In summary, calculating $\hat{\tilde{\mathbf{\Lambda}}}_{\mathrm{stat}} \mathbf{x}_{\mathrm{off}}$ to obtain $\mathbf{x}_{\mathrm{off},c}$ results in a little bit more effort for UW-OFDM to handle CFO properly compared to a CP-OFDM system. However, the additive term requires an update only when ϵ or $\tilde{\mathbf{H}}$ change to a certain extent, keeping thus the effort manageable.

6.4. Impact of Imperfect CFO Compensation

The CFO induced error caused by subtracting the wrong UW and pilot portion can be kept very small with manageable computational effort (see previous section). Fig. 6.11 confirms that the performance loss due to incomplete CFO compensation by derotating with $\hat{\varphi}_l$ is much more dominant, motivating more in-depth investigations. This section aims at reducing the remaining MSE between the lower bound and the abscissa in Fig. 6.11. A zero UW and perfect pilot subtraction is assumed for the subsequent considerations to separately carry out the effects of imperfect CFO compensation, yielding from (6.64) the system model

$$\tilde{\mathbf{y}}_c^{(l)} = \mathrm{e}^{-j\hat{\varphi}_l} \tilde{\mathbf{\Lambda}}^{(l)} \tilde{\mathbf{H}} \mathbf{G}_d \mathbf{d}^{(l)} + \mathrm{e}^{-j\hat{\varphi}_l} \mathbf{v}. \tag{6.137}$$

The received signal has already been corrected by derotating with $-\hat{\varphi}_l$. Consequently, perfect CFO compensation would correspond to multiplying with $\mathrm{e}^{j\hat{\varphi}_l} \tilde{\mathbf{\Lambda}}^{(l)-1}$ and result in a zero MSE (given zero noise). This is of course not a realistic scenario. Starting with the hypothetical assumption of a noise free estimate $\hat{\varphi}_l$ ($\Delta_l = 0$) has delivered very good results in the previous section and seems a promising entry point for this task as well. With $\hat{\varphi}_l = \varphi_l + \varphi_{\mathrm{pil}}$, the system model translates to

$$\tilde{\mathbf{y}}_{c,(\mathrm{exp})}^{(l)} = \mathrm{e}^{-j(\varphi_l + \varphi_{\mathrm{pil}})} \tilde{\mathbf{\Lambda}}^{(l)} \tilde{\mathbf{H}} \mathbf{G}_d \mathbf{d}^{(l)} + \mathrm{e}^{-j(\varphi_l + \varphi_{\mathrm{pil}})} \mathbf{v} \tag{6.138}$$

$$= \mathrm{e}^{-j\varphi_{\mathrm{pil}}} \tilde{\mathbf{\Lambda}}_{\mathrm{stat}} \tilde{\mathbf{H}} \mathbf{G}_d \mathbf{d}^{(l)} + \mathrm{e}^{-j(\varphi_l + \varphi_{\mathrm{pil}})} \mathbf{v}. \tag{6.139}$$

The main idea of this thought experiment is that the CFO effects are the same for all OFDM symbols, thus simplifying further compensation techniques significantly.

Multiplying with $\mathrm{e}^{j\varphi_{\mathrm{pil}}}\tilde{\mathbf{\Lambda}}_{\mathrm{stat}}^{-1}$ would cancel out all CFO effects, but both, the constant phase offset and the matrix depend on ϵ, hence only estimates of them are available. Furthermore, a matrix inversion comes along with an additional computational overhead. According to (6.34), $\tilde{\mathbf{\Lambda}}_{\mathrm{stat}}$ shows high similarities to a unitary matrix, therefore the inversion may be approximated as

$$\tilde{\mathbf{\Lambda}}_{\mathrm{stat}}^{-1} \approx \tilde{\mathbf{\Lambda}}_{\mathrm{stat}}^{H}. \tag{6.140}$$

Based on the results of this thought experiment, the data estimate follows as

$$\begin{aligned}
\hat{\mathbf{d}}^{(l)} &= \mathbf{E}\mathrm{e}^{j\hat{\varphi}_{\mathrm{pil}}}\hat{\tilde{\mathbf{\Lambda}}}_{\mathrm{stat}}\tilde{\mathbf{y}}_c^{(l)} && (6.141)\\
&= \mathbf{E}\mathrm{e}^{j\hat{\varphi}_{\mathrm{pil}}}\hat{\tilde{\mathbf{\Lambda}}}_{\mathrm{stat}}^{H}\mathrm{e}^{-j\hat{\varphi}_l}\tilde{\mathbf{\Lambda}}^{(l)}\tilde{\mathbf{H}}\mathbf{G}_d\mathbf{d}^{(l)} + \mathbf{E}\mathrm{e}^{j\hat{\varphi}_{\mathrm{pil}}}\hat{\tilde{\mathbf{\Lambda}}}_{\mathrm{stat}}^{H}\mathrm{e}^{-j\hat{\varphi}_l}\mathbf{v} && (6.142)\\
&= \mathbf{E}\mathrm{e}^{j\hat{\varphi}_{\mathrm{pil}}}\hat{\tilde{\mathbf{\Lambda}}}_{\mathrm{stat}}^{H}\mathrm{e}^{-j(\varphi_l+\varphi_{\mathrm{pil}}+\Delta_l)}\tilde{\mathbf{\Lambda}}^{(l)}\tilde{\mathbf{H}}\mathbf{G}_d\mathbf{d}^{(l)}\\
&\quad + \mathbf{E}\mathrm{e}^{j\hat{\varphi}_{\mathrm{pil}}}\hat{\tilde{\mathbf{\Lambda}}}_{\mathrm{stat}}^{H}\mathrm{e}^{-j(\varphi_l+\varphi_{\mathrm{pil}}+\Delta_l)}\mathbf{v} && (6.143)\\
&= \mathrm{e}^{j(\hat{\varphi}_{\mathrm{pil}}-\varphi_{\mathrm{pil}}-\Delta_l)}\mathbf{E}\hat{\tilde{\mathbf{\Lambda}}}_{\mathrm{stat}}^{H}\tilde{\mathbf{\Lambda}}_{\mathrm{stat}}\tilde{\mathbf{H}}\mathbf{G}_d\mathbf{d}^{(l)} + \mathrm{e}^{j(\hat{\varphi}_{\mathrm{pil}}-\varphi_{\mathrm{pil}}-\Delta_l-\varphi_l)}\mathbf{E}\hat{\tilde{\mathbf{\Lambda}}}_{\mathrm{stat}}^{H}\mathbf{v}. && (6.144)
\end{aligned}$$

In contrast to (6.90) with $\hat{\varphi}_{\mathrm{off}} = \hat{\varphi}_{\mathrm{SI}} - \hat{\varphi}_{\mathrm{pil}}$, the phase offset compensation comprises only the impact of φ_{pil}, φ_{SI} is expected to vanish based on $\hat{\tilde{\mathbf{\Lambda}}}_{\mathrm{stat}}^{H}\tilde{\mathbf{\Lambda}}_{\mathrm{stat}} \approx \mathbf{I}$. Fig. 6.12 unveils a significant reduction of the MSE due to the improved CFO compensation method. Additionally, a reference curve with perfect conditions, i.e., $\tilde{\mathbf{\Lambda}}_{\mathrm{stat}}^{-1}$ and $\hat{\varphi}_{\mathrm{pil}} = \varphi_{\mathrm{pil}}$, shows only a minor performance loss due to the applied approximations. Furthermore, the benefit of offset compensation becomes more dominant, since φ_{pil} and φ_{SI} do not cancel each other as it happens in (6.90) to a certain extent. Even in absolute terms the gain of offset compensation approximately doubles compared to the case of an UW-OFDM system which only applies simple derotation by $\hat{\varphi}_l$ as a CFO compensation means (Fig. 6.10, UW-OFDM with a CAZAC sequence as UW). Also for CP-OFDM, the performance improves significantly due to the advanced CFO compensation technique. The difference to UW-OFDM is reduced compared to the phase derotation case in Fig. 6.10, but it still performs worse than UW-OFDM.

6.5. Estimation of Phase Rotation and Carrier Frequency Offset

Usually, a sufficient countermeasure against CFO is derotating with $\mathrm{e}^{-j\varphi_l}$ at the receiver, which is also known as carrier phase synchronization. According to (6.29), the lth OFDM symbol experiences a phase rotation of

$$\varphi_l = \frac{2\pi}{N}\epsilon\left(Nl + N_u + \frac{N-1}{2}\right). \tag{6.145}$$

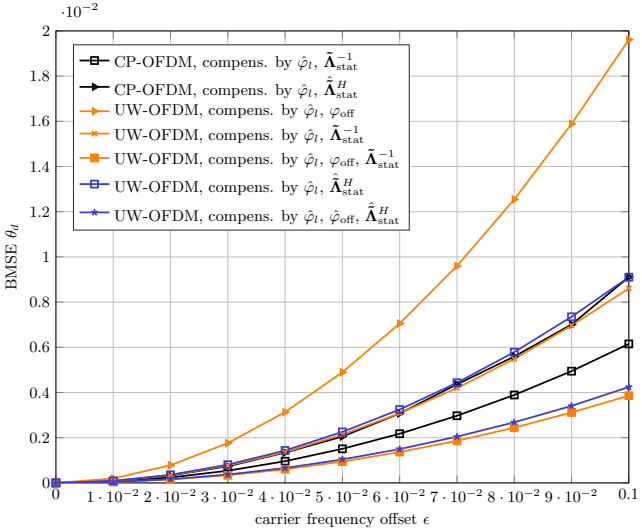


Figure 6.12.: Carrier frequency offset induced Bayesian MSE θ_d in UW-OFDM and CP-OFDM. Investigation of advanced CFO compensation techniques on top of a phase derotation by $\hat{\varphi}_l$. Zero UW and perfect pilot subtraction with $\mathbf{e}_{x_p}^{(l)} = \mathbf{0}$ is assumed for the UW-OFDM case. The estimate $\hat{\mathbf{d}}^{(l)}$ has been obtained according to (6.144). Investigated scenario: $\mathbf{G}_d'$ and multipath environment with $\tau_{\text{RMS}} = 100\,\text{ns}$.

The phase is therefore linearly increasing from OFDM symbol to OFDM symbol with

$$\varphi_\Delta = \varphi_l - \varphi_{l-1} \tag{6.146}$$

$$= \frac{2\pi}{N}\epsilon\left(Nl + N_u + \frac{N-1}{2}\right) - \frac{2\pi}{N}\epsilon\left(N(l-1) + N_u + \frac{N-1}{2}\right) \tag{6.147}$$

$$= \frac{2\pi}{N}\epsilon\left(Nl - N(l-1)\right) = \frac{2\pi}{N}\epsilon N = 2\pi\epsilon. \tag{6.148}$$

There are various approaches for estimating the phase rotation φ_l in OFDM, e.g., based on data symbols [84, 85] known as decision directed schemes or cyclic prefix based methods [86, 87]. Another common way is the utilization of pilot symbols $\mathbf{p}$ in the frequency domain [82]

$$\hat{\varphi}_l = \arg\left(\mathbf{p}^H \mathbf{W}_p \hat{\mathbf{p}}^{(l)}\right). \tag{6.149}$$

Here, $\hat{\mathbf{p}}^{(l)}$ denotes the extracted pilot symbols from the lth received OFDM symbol and $\mathbf{W}_p = \text{diag}\left(\mathbf{w}_p\right)$ is a diagonal weighting matrix with $\mathbf{w}_p \in \mathbb{R}_+^{0\,(N_p \times 1)}$ to rate the

pilots regarding their estimation quality (e.g., based on the inverse main diagonal of an error covariance matrix). In order to evaluate the applicability of this concept to UW-OFDM as well, let us investigate the effects of a CFO on the pilot symbols more in detail. The general definition of a received OFDM symbol introduced in (6.63) serves as starting point

$$\tilde{\mathbf{y}}_d^{(l)} \approx \tilde{\mathbf{\Lambda}}^{(l)}\tilde{\mathbf{H}}\mathbf{G}_p\mathbf{p} + \tilde{\mathbf{\Lambda}}^{(l)}\tilde{\mathbf{H}}\mathbf{G}_d\mathbf{d}^{(l)} + \tilde{\mathbf{\Lambda}}^{(l)}\tilde{\mathbf{H}}\mathbf{B}^T\tilde{\mathbf{x}}_u + \mathbf{v}. \tag{6.150}$$

To estimate φ_l (which is part of $\tilde{\mathbf{\Lambda}}^{(l)}$ and will be extracted a little bit later for reasons of compact notation) based on the pilot symbols, estimation theory suggests to transform the affine model into a linear one, cf. Sec. 2.3. Since the data vector $\mathbf{d}^{(l)}$ is unknown, subtracting the corresponding term would be quite difficult. However, data and pilot symbols are arranged orthogonal in the frequency domain, hence considering only the pilot subcarriers would solve this issue. The only degrading effect of the data symbols on the estimate of the pilot symbols arises then from intercarrier interference, which is normally in an acceptable range. This effect is experienced by all OFDM systems. Specific to UW-OFDM though is the second additive term caused by the UW. Contrary to the data symbols, the UW itself is known to the receiver, but this term is also problematic due to $\tilde{\mathbf{\Lambda}}^{(l)}$. In other words, we aim to estimate φ_l and therefore $\tilde{\mathbf{\Lambda}}^{(l)}$ on the one hand, but also need to know them for the estimation process itself on the other hand. Investigations not detailed here verify that subtracting an approximation $\tilde{\mathbf{H}}\mathbf{B}^T\tilde{\mathbf{x}}_u$ instead leads to poor estimates of φ_l, hence this is also not an adequate solution. Of course, iterative algorithms could solve this issue, but would lead to a high computational complexity. Therefore, an affine model serves as starting point anyway and an estimator $\mathbf{E}_p' = \mathbf{E}_p\tilde{\mathbf{H}}^{-1}$ consisting of two stages is introduced. The first stage inverts the multipath channel, and the second stage $\mathbf{E}_p = \begin{bmatrix}\mathbf{I} & \mathbf{0}\end{bmatrix}\mathbf{P}_p^T$ extracts the pilots from the frequency domain vector, with a permutation matrix $\mathbf{P}_p$ introduced in (5.11). $\mathbf{E}_p$ is only a suboptimum estimator from a performance point of view, as it is not capable of exploiting the portion of the pilots spread on the non-pilot subcarriers due to $\mathbf{G}_p\mathbf{p}$. This for instance would be possible with $\mathbf{E}_p = \mathbf{G}_p^\dagger$, which denotes the Moore-Penrose Pseudo-Inverse of the pilot generator matrix. Nevertheless, the suboptimum approach does not require the subtraction of the data and UW dependent offset term in (6.150) and thus simplifies the estimation task greatly. Applying $\mathbf{E}_p'$ leads to

$$\begin{aligned}
\hat{\mathbf{p}}^{(l)} &= \mathbf{E}_p\tilde{\mathbf{H}}^{-1}\tilde{\mathbf{y}}^{(l)} && (6.151)\\
&= \mathbf{E}_p\tilde{\mathbf{H}}^{-1}\tilde{\mathbf{\Lambda}}^{(l)}\tilde{\mathbf{H}}\mathbf{G}_p\mathbf{p} + \mathbf{E}_p\tilde{\mathbf{H}}^{-1}\tilde{\mathbf{\Lambda}}^{(l)}\tilde{\mathbf{H}}^{-1}\mathbf{G}_d\mathbf{d}^{(l)} + \mathbf{E}_p\tilde{\mathbf{H}}^{-1}\tilde{\mathbf{\Lambda}}^{(l)}\tilde{\mathbf{H}}\mathbf{B}^T\tilde{\mathbf{x}}_u\\
&\quad + \mathbf{E}_p\tilde{\mathbf{H}}^{-1}\mathbf{B}^T\mathbf{v} && (6.152)\\
&= \mathbf{E}_p\tilde{\mathbf{H}}^{-1}\tilde{\mathbf{\Lambda}}^{(l)}\tilde{\mathbf{H}}\left(\mathbf{G}_p\mathbf{p} + \mathbf{B}^T\tilde{\mathbf{x}}_u\right) + \mathbf{d}_{\mathrm{ICI}}^{(l)} + \mathbf{v}'' && (6.153)\\
&= \mathbf{E}_p\tilde{\mathbf{\Lambda}}_H^{(l)}\left(\mathbf{G}_p\mathbf{p} + \mathbf{B}^T\tilde{\mathbf{x}}_u\right) + \mathbf{d}_{\mathrm{ICI}}^{(l)} + \mathbf{v}''. && (6.154)
\end{aligned}$$

As already mentioned, pilot and data symbols are actually orthogonal in frequency domain, but $\tilde{\mathbf{\Lambda}}^{(l)}$ introduces intercarrier interference resulting in

$$\mathbf{d}_{\mathrm{ICI}}^{(l)} = \mathbf{E}_p \tilde{\mathbf{H}}^{-1} \tilde{\mathbf{\Lambda}}^{(l)} \tilde{\mathbf{H}} \mathbf{G}_d \mathbf{d}^{(l)}, \tag{6.155}$$

with $\mathbf{d}_{\mathrm{ICI}}^{(l)} \in \mathbb{C}^{(N_p \times 1)}$. The vector $\mathbf{v}'' \in \mathbb{C}^{(N_p \times 1)}$ represents additive noise according to

$$\mathbf{v}'' = \mathbf{E}_p \tilde{\mathbf{H}}^{-1} \mathbf{B}^T \mathbf{v}. \tag{6.156}$$

Since $\tilde{\mathbf{\Lambda}}^{(l)}$ has entries off the main diagonal and is only approximately a diagonal matrix, multiplying with $\tilde{\mathbf{H}}^{-1}$ cannot fully equalize the multipath channel $\tilde{\mathbf{H}}$, these effects are incorporated in

$$\tilde{\mathbf{\Lambda}}_H^{(l)} = \tilde{\mathbf{H}}^{-1} \tilde{\mathbf{\Lambda}}^{(l)} \tilde{\mathbf{H}}. \tag{6.157}$$

For an easier analysis, a *single* pilot symbol is considered in the following, which is given as

$$\hat{p}^{(l)}[k] = [\mathbf{E}_p]_{k,*} \, \tilde{\mathbf{\Lambda}}_H^{(l)} \left(\mathbf{G}_p \mathbf{p} + \mathbf{B}^T \tilde{\mathbf{x}}_u \right) + d_{\mathrm{ICI}}^{(l)}[k] + v''[k] \tag{6.158}$$

$$= [\mathbf{E}_p]_{k,*} \, \tilde{\mathbf{\Lambda}}_H^{(l)} \left([\mathbf{G}_p]_{*,k} \, p[k] + \mathbf{B}^T \tilde{\mathbf{x}}_u \right) + [\mathbf{E}_p]_{k,*} \, \tilde{\mathbf{\Lambda}}_H^{(l)} \sum_{m=0, m \neq k}^{N_p - 1} [\mathbf{G}_p]_{*,m} \, p[m] + d_{\mathrm{ICI}}^{(l)}[k] + v''[k] \tag{6.159}$$

$$= \mathbf{e}_k^T \tilde{\mathbf{\Lambda}}_H^{(l)} \left(\mathbf{g}_k p[k] + \mathbf{B}^T \tilde{\mathbf{x}}_u \right) + \mathbf{e}_k^T \tilde{\mathbf{\Lambda}}_H^{(l)} \sum_{m=0, m \neq k}^{N_p - 1} \mathbf{g}_m p[m] + d_{\mathrm{ICI}}^{(l)}[k] + v''[k] \tag{6.160}$$

$$= \mathbf{e}_k^T \tilde{\mathbf{\Lambda}}_H^{(l)} \left(\mathbf{g}_k p[k] + \mathbf{B}^T \tilde{\mathbf{x}}_u \right) + p_{\mathrm{ICI}}^{(l)}[k] + d_{\mathrm{ICI}}^{(l)}[k] + v''[k]. \tag{6.161}$$

Consequently, there will be four terms influencing the measured angle. The additive noise term $v''[k]$ should not require further explanation. The third term $d_{\mathrm{ICI}}^{(l)}[k]$ models the intercarrier interference induced by data symbols. This random term has similar influence as additive noise and degrades the estimation quality to a certain extent. The second term

$$p_{\mathrm{ICI}}^{(l)}[k] = \mathbf{e}_k^T \tilde{\mathbf{\Lambda}}_H^{(l)} \sum_{m=0, m \neq k}^{N_p - 1} \mathbf{g}_m p[m] \tag{6.162}$$

comprises the intercarrier interference induced by the other pilot symbols. Since the pilot symbols are constant, this term is also constant (given a fixed CFO ϵ), which leads to a constant estimation error. The pilot symbols are usually uniformly distributed over the available spectrum, which results in a distance of several subcarriers among them. Hence, the influence of this term is limited and can be neglected. The first term in (6.161) comprises the actual pilot symbol,

an additive offset caused by the UW and a part causing the phase rotation to be estimated.

Merging the last three terms in (6.161) in a new additive noise term $v'''[k] = p_{\mathrm{ICI}}^{(l)}[k] + d_{\mathrm{ICI}}^{(l)}[k] + v''[k]$, the estimation of φ_l reads

$$\hat{\varphi}_l = \arg\left(\mathbf{p}^H \mathbf{W}_p \hat{\mathbf{p}}^{(l)}\right) \tag{6.163}$$

$$= \arg\left(\sum_{k=0}^{N_p-1} p[k]^H w_p[k] \hat{p}^{(l)}[k]\right) \tag{6.164}$$

$$= \arg\left(\sum_{k=0}^{N_p-1} \mathbf{e}_k^T \tilde{\mathbf{\Lambda}}_H^{(l)} \left(\mathbf{g}_k w_p[k]|p[k]|^2 + \mathbf{B}^T \tilde{\mathbf{x}}_u w_p[k] p[k]^H\right) + v'''[k] w_p[k] p[k]^H\right) \tag{6.165}$$

$$= \arg\Bigg(\underbrace{\sum_{k=0}^{N_p-1} \mathrm{e}^{j\varphi_l} \mathbf{e}_k^T \tilde{\mathbf{\Lambda}}_{H,\mathrm{stat}} \left(\mathbf{g}_k w_p[k]|p[k]|^2 + \mathbf{B}^T \tilde{\mathbf{x}}_u w_p[k] p[k]^H\right)}_{\text{pilot term}} + \underbrace{\sum_{k=0}^{N_p-1} v'''[k] w_p[k] p[k]^H}_{\text{noise term}}\Bigg), \tag{6.166}$$

with

$$\tilde{\mathbf{\Lambda}}_{H,\mathrm{stat}} = \tilde{\mathbf{H}}^{-1} \tilde{\mathbf{\Lambda}}_{\mathrm{stat}} \tilde{\mathbf{H}}. \tag{6.167}$$

The angle $\hat{\varphi}_l$ depends now on two sources, whereas the first one is denoted as *pilot term* and the second one as *noise term*. As a next step, the effects of both sources on $\hat{\varphi}_l$ shall be separated. Taking into account the average power of the pilot term to be much larger than of the noise term, the argument in (6.166) will mainly be determined by the pilot term, i.e., it almost equals an argument only taken from the pilot term. This translates to the model

$$\hat{\varphi}_l = \arg\left(\sum_{k=0}^{N_p-1} \mathrm{e}^{j\varphi_l} \mathbf{e}_k^T \tilde{\mathbf{\Lambda}}_{H,\mathrm{stat}} \left(\mathbf{g}_k w_p[k]|p[k]|^2 + \mathbf{B}^T \tilde{\mathbf{x}}_u w_p[k] p[k]^H\right)\right) + \Delta_l \tag{6.168}$$

$$= \arg\left(\mathrm{e}^{j\varphi_l} \sum_{k=0}^{N_p-1} \mathbf{e}_k^T \tilde{\mathbf{\Lambda}}_{H,\mathrm{stat}} \mathbf{g}_k w_p[k]|p[k]|^2 + \mathbf{e}_k^T \tilde{\mathbf{\Lambda}}_{H,\mathrm{stat}} \mathbf{B}^T \tilde{\mathbf{x}}_u w_p[k] p[k]^H\right) + \Delta_l, \tag{6.169}$$

whereas Δ_l denotes a minor additive deviation approximated as a function of

$$\Delta_l \approx f\left(\sum_{k=0}^{N_p-1} v'''[k] w_p[k] p[k]^H\right). \tag{6.170}$$

In reality, Δ_l would of course depend on both, the noise as well as the pilot term in (6.166). Nonetheless, due to the inequality in power distribution, modelling $\hat{\varphi}_l$ as a linear combination of two angles and reducing the dependence of Δ_l to only the noise term seems to be appropriate. A further specification of the function $f(\cdot)$ is not of relevance for this analysis and is therefore disregarded. The expression in (6.169) is still quite complex and requires additional simplification to allow an intuitive interpretation of the impact factors on $\hat{\varphi}_l$. Let us therefore introduce the definitions

$$a_{\mathrm{pil},k} \mathrm{e}^{j\varphi_{\mathrm{pil},k}} = \mathbf{e}_k^T \tilde{\mathbf{\Lambda}}_{H,\mathrm{stat}} \mathbf{g}_k w_p[k] |p[k]|^2 + \mathbf{e}_k^T \tilde{\mathbf{\Lambda}}_{H,\mathrm{stat}} \mathbf{B}^T \tilde{\mathbf{x}}_u w_p[k] p[k]^H \tag{6.171}$$

$$a_{\mathrm{pil}} \mathrm{e}^{j\varphi_{\mathrm{pil}}} = \sum_{k=0}^{N_p-1} a_{\mathrm{pil},k} \mathrm{e}^{j\varphi_{\mathrm{pil},k}}, \tag{6.172}$$

then the estimated phase rotation $\hat{\varphi}_l$ can finally be written as

$$\hat{\varphi}_l = \arg\left(\mathrm{e}^{j\varphi_l} \sum_{k=0}^{N_p-1} \mathbf{e}_k^T \tilde{\mathbf{\Lambda}}_{H,\mathrm{stat}} \mathbf{g}_k w_p[k] |p[k]|^2 + \mathbf{e}_k^T \tilde{\mathbf{\Lambda}}_{H,\mathrm{stat}} \mathbf{B}^T \tilde{\mathbf{x}}_u w_p[k] p[k]^H\right) + \Delta_l \tag{6.173}$$

$$= \arg\left(\mathrm{e}^{j\varphi_l} \sum_{k=0}^{N_p-1} a_{\mathrm{pil},k} \mathrm{e}^{j\varphi_{\mathrm{pil},k}}\right) + \Delta_l \tag{6.174}$$

$$= \arg\left(\mathrm{e}^{j\varphi_l} a_{\mathrm{pil}} \mathrm{e}^{j\varphi_{\mathrm{pil}}}\right) + \Delta_l \tag{6.175}$$

$$= \varphi_l + \varphi_{\mathrm{pil}} + \Delta_l. \tag{6.176}$$

Contrary to conventional OFDM, an estimate based on frequency pilot tones incorporates an additional phase offset φ_{pil}. This phase offset is influenced by the CFO ϵ due to $\tilde{\mathbf{\Lambda}}_{\mathrm{stat}}$, the multipath channel $\tilde{\mathbf{H}}$, the UW $\tilde{\mathbf{x}}_u$, the utilized estimator $\mathbf{E}_p$ as well as the generator matrix $\mathbf{G}_p$, and impacts the detection performance of the data symbols in (6.66)-(6.72). In detail, assuming first that derotating by φ_l sufficiently cancels the CFO effects and thus $\tilde{\mathbf{\Lambda}}^{(l)}$, and second that the estimator neutralizes the channel and the generator matrix, a phase offset φ_{pil} remains to be compensated. This angle denotes an offset as the result of averaging over several single estimates. To avoid the averaging operation for the moment due to reasons of easier explanation, the offset from a single estimate $\varphi_{\mathrm{pil},k}$ is considered in the following. At the end, a link back to φ_{pil} will complete the analysis. Going back to (6.173), there are two sources causing the phase offset. The UW is the first

source and a pretty obvious one, as it has not been subtracted at the beginning of the estimation process in (6.150). The second source is an additional scaling and rotating of the pilot symbol, which in turn originates from the fact that $\mathbf{G}_p$ places a pilot symbol at the dedicated subcarrier position, but also spreads portions of it over the remaining subcarriers to fulfill the zero word constraint of UW-OFDM, cf. (5.2). In combination with $\tilde{\mathbf{\Lambda}}^{(l)}$, the portions of a pilot symbol spread over several subcarriers are leaked back. This leads in a certain way to a self interference of the pilot symbols, which has to be accounted for. The same problem has already been described for the data symbols in Sec. 6.2. Note that this effect is specific for UW-OFDM and does not occur in conventional OFDM.

6.5.1. Phase Offset Compensation

The first step towards offset compensation in (6.176) is now to obtain an estimate of φ_{pil}. There are two problems emerging in this context as a result of the dependence of φ_{pil} on ϵ due to $\tilde{\mathbf{\Lambda}}_{H,\text{stat}}$. First, a straightforward relationship between φ_{pil} and ϵ is not immediately apparent from (6.173). Second, an estimate $\hat{\varphi}_{\text{pil}}$ requires also an estimate $\hat{\epsilon}$, which in turn has somehow be obtained from $\hat{\varphi}_l$ based on the relationship between φ_l and ϵ in (6.145). The following part will establish a simple model between φ_{pil} and ϵ in a first step, and resolve the issues resulting from the mutual dependencies between φ_l, φ_{pil} and ϵ in a second step. Fig. 6.13 plots the phase offset $\varphi_{\text{pil},k}$ with $k = 0$ for the pilot generator matrices $\mathbf{G}'_p$ and $\mathbf{G}''_p$ from chapter 5 in case of $\tilde{\mathbf{H}} = \mathbf{I}$, $\mathbf{W}_p = |\tilde{\mathbf{H}}_p|^2 = \mathbf{I}$ and $\tilde{\mathbf{x}}_u \neq \mathbf{0}$. Here, $\tilde{\mathbf{H}}_p \in \mathbb{C}^{N_p \times N_p}$ denotes a diagonal matrix with the channel coefficients corresponding to the pilot subcarriers on the main diagonal. It turns out that the offset is approximately an affine function of ϵ for the relevant CFO range for a given setup. As a first step to account for the offset, the affine model

$$\varphi_{\text{pil},k} = m_{\text{pil},k}\epsilon + q_k \tag{6.177}$$

is introduced. Considering the case $\epsilon = 0$ for any realization of $\tilde{\mathbf{H}}$, the matrix $\tilde{\mathbf{\Lambda}}_{H,\text{stat}}$ in (6.173) collapses to an identity matrix, resulting in $\mathbf{e}_k^T \tilde{\mathbf{\Lambda}}_{H,\text{stat}} \mathbf{g}_k = 1$ and $\mathbf{e}_k^T \tilde{\mathbf{\Lambda}}_{H,\text{stat}} \mathbf{B}^T \tilde{\mathbf{x}}_u = \tilde{x}_u[i_{p,k}]$, with $i_{p,k}$ addressing the kth element of the ordered pilot subcarrier index set $\mathcal{I}_{p,o} = (\mathcal{I}_p, <)$ defined as

$$\mathcal{I}_{p,o} = \left\{ i_{p,0}, i_{p,1}, \dots i_{p,N_p-1} \right\}. \tag{6.178}$$

In this context, a single estimate follows then as

$$\begin{aligned}
\hat{\varphi}_{\text{pil},k|\epsilon=0} &= q_k && (6.179)\\
&= \arg\left(w_p[k]|p[k]|^2 + \tilde{x}_u[i_{p,k}] w_p[k] p[k]^H \right) + \Delta_l && (6.180)\\
&= w_p[k] \arg\left(|p[k]|^2 + \tilde{x}_u[i_{p,k}] p[k]^H \right) + \Delta_l && (6.181)\\
&\approx w_p[k] \arg\left(|p[k]|^2 + \tilde{x}_u[i_{p,k}] p[k]^H \right). && (6.182)
\end{aligned}$$

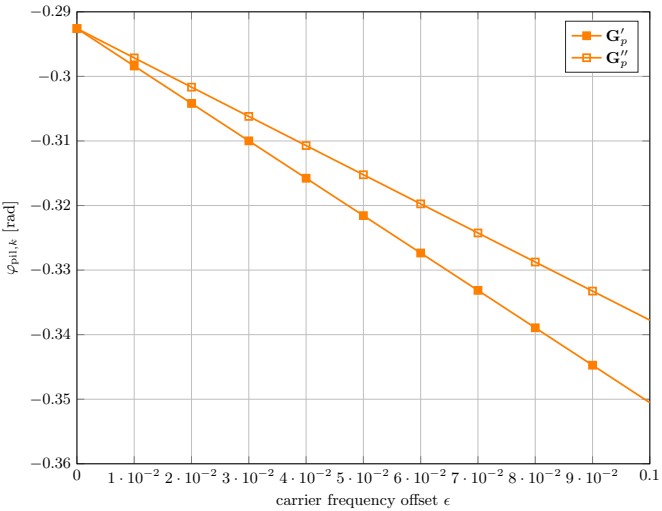


Figure 6.13.: UW-OFDM specific phase offset $\varphi_{\mathrm{pil},k}$ (with $k = 0$) when estimating the phase rotation φ_l caused by a CFO for $\tilde{\mathbf{H}} = \mathbf{I}$, $\tilde{\mathbf{x}}_u \neq \mathbf{0}$, $\sigma_v^2 = 0$ and $\mathbf{W}_p = \mathbf{I}$.

Clearly, the constant phase offset q_k originates from the presence of a non-zero UW and vanishes in case of $\tilde{\mathbf{x}}_u = \mathbf{0}$ and thus $\tilde{x}_u[i_{p,k}] = 0$. The actual value of q_k depends on the UW offset at the specific subcarrier and the pilot symbol itself. The weighting factor $w_p[k]$ only plays a role when determining the averaged estimate $\hat{\varphi}_{\mathrm{pil}}$ in (6.172). A simple way of avoiding the constant offset q_k in the estimate would be the definition of new pilots $\mathbf{p}' = \mathbf{p} + \tilde{\mathbf{x}}_{u,\mathcal{I}_p}$, where $\tilde{\mathbf{x}}_{u,\mathcal{I}_p}$ denotes a vector with elements out of $\tilde{\mathbf{x}}_u$ at the corresponding pilot subcarrier positions. The estimation of φ_l follows then according to $\hat{\varphi}_l = \arg\left(\mathbf{p}'^H \mathbf{W}_p \hat{\mathbf{p}}^{(l)}\right)$. The only difference is the transform of the affine into a linear model in (6.177), everything else in the estimation process stays the same.

The variable part $m_{\mathrm{pil},k}\epsilon$ of the phase offset exists independent of the presence of a zero or non-zero UW. The slope $m_{\mathrm{pil},k}$ is determined by both sources, the UW and the rotation and scaling of the pilot symbol

$$\begin{aligned} m_{\mathrm{pil},k}\epsilon = \arg\Big(&\mathbf{e}_k^T \tilde{\mathbf{\Lambda}}_{H,\mathrm{stat}} \mathbf{g}_k w_p[k] |p[k]|^2 \\ &+ \mathbf{e}_k^T \tilde{\mathbf{\Lambda}}_{H,\mathrm{stat}} \mathbf{B}^T \tilde{\mathbf{x}}_u w_p[k] p[k]^H \Big) - q_k, \end{aligned} \tag{6.183}$$

which in turn depend on the channel realization $\tilde{\mathbf{H}}$, the generator matrix $\mathbf{G}_p$, the estimator $\mathbf{E}_p$, and the pilot symbols $\mathbf{p}$. Regardless of the specific values of these

parameters, the affine model for $\varphi_{\mathrm{pil},k}$ as a function of ϵ holds in the relevant range.

The very same model also holds for the averaged estimate

$$\varphi_{\mathrm{pil}} = m_{\mathrm{pil}}\epsilon + q. \tag{6.184}$$

Note that in general, m_{pil} and q are not obtained from simply averaging over $m_{\mathrm{pil},k}$ and q_k, respectively. However, the relationship between the single and the averaged estimate is not required for compensating the offset. As already established, the model of a single estimate has only been introduced to simplify explanations.

In a real communication system, the parameters of the affine model for φ_{pil} are easily determined. Given a certain setup and assuming knowledge of the channel $\tilde{\mathbf{H}}$ or an estimate of it (which does not induce an additional effort as it is required for other purposes anyway), the parameters can be derived by numerically evaluating φ_{pil} at two different points, e.g., for $\epsilon = 0$ and $\epsilon = 0.1$. Inserting then (6.184) into (6.176) yields

$$\hat{\varphi}_l = \varphi_l + \varphi_{\mathrm{pil}} + \Delta_l \tag{6.185}$$

$$= \varphi_l + m_{\mathrm{pil}}\epsilon + q + \Delta_l. \tag{6.186}$$

Compensating the offset delivers a new estimate

$$\hat{\tilde{\varphi}}_l = \hat{\varphi}_l - \hat{\varphi}_{\mathrm{pil}} = \hat{\varphi}_l - m_{\mathrm{pil}}\hat{\epsilon} - q. \tag{6.187}$$

Offset Compensation for Data Estimation As already discussed in Sec. 6.2, the data symbols $d^{(l)}[k]$ in (6.78) experience also phase rotation effects labeled as φ_{SI} to stress the source self interference. Conducting the very same experiments as for φ_{pil} shows that except for a zero offset $q = 0$, the same linearization model

$$\varphi_{\mathrm{SI}} = m_{\mathrm{SI}}\epsilon \tag{6.188}$$

applies. In this context, the remaining phase offset φ_{off} in (6.88) can be modelled and estimated as

$$\varphi_{\mathrm{off}} = \varphi_{\mathrm{SI}} - \varphi_{\mathrm{pil}} = (m_{\mathrm{SI}} - m_{\mathrm{pil}})\,\epsilon \tag{6.189}$$

$$\hat{\varphi}_{\mathrm{off}} = (m_{\mathrm{SI}} - m_{\mathrm{pil}})\,\hat{\epsilon}. \tag{6.190}$$

Offset Compensation for CFO Estimation The presented linearization highlights the simple relationship of the discussed phase offsets and the CFO ϵ. For compensation though, there is still the estimate $\hat{\epsilon}$ missing, which will be provided next. Note that the CFO ϵ is a linear function of φ_l as shown in (6.145) and could therefore easily be derived. Unfortunately, the available estimate $\hat{\varphi}_l$ in (6.176)

additionally incorporates an unknown offset φ_{pil}, which is not accounted for in (6.145). Fortunately, the results presented in this section support an approximation of φ_{pil} as a linear function of ϵ. Therefore, introducing the prefactor $\frac{2\pi}{N}\epsilon$ from (6.145) to obtain $m_{\text{pil}} = N_{\text{pil}}\frac{2\pi}{N}\epsilon$ with a constant N_{pil}, $\hat{\epsilon}$ can elegantly be calculated as

$$\hat{\epsilon}_l = \epsilon + \Delta_{\epsilon_l} \tag{6.191}$$

$$= \hat{\varphi}_l \frac{N}{2\pi\left(\frac{N-1}{2} + (Nl + N_u) + N_{\text{pil}}\right)}. \tag{6.192}$$

The error Δ_{ϵ_l} in $\hat{\epsilon}_l$ decreases with increasing $\hat{\varphi}_l$ due to Nl in the denominator. However, the proposed estimation algorithm $\hat{\varphi}_l = \arg\left(\mathbf{p}^H \mathbf{W}_p \hat{\mathbf{p}}^{(l)}\right)$ implicitly applies a modulo operation on $\hat{\varphi}_l$, as it can only deliver values within the range $[0, 2\pi)$, whereas the estimator in (6.192) requires the total angle accumulated from the beginning of the burst up to and including OFDM symbol l. This restriction limits the applicability to angles φ_l not exceeding 2π and thus also the estimation accuracy. One way to circumvent this restriction is averaging over several estimates $\hat{\epsilon}_l$, gained from the knowledge that the angle increases linearly with $\Delta_l = \varphi_l - \varphi_{l-1} = 2\pi\epsilon$ between two OFDM symbols, cf. (6.148). Of course, this requires the CFO to be constant for the period to be averaged over.

6.5.2. Performance Evaluation

The preceding analysis has unveiled a few differences in pilot based CFO estimation between UW-OFDM and CP-OFDM. The question remains, whether these differences have an impact on the estimation quality, which is answered next. So far, the utilized non-systematically encoded generator matrices have been normalized to be energy invariant, yielding $\mathbf{G}^H\mathbf{G} = \alpha\mathbf{I}$ with $\alpha = 1$. For the BER simulations presented in chapter 4, the value of α is actually irrelevant for the performance, as it is considered in the E_b/N_0 calculation. In other words, given a fixed SNR, the noise increases proportionally with α. For the CFO estimation part, the value of α actually does matter. In case of zero additive noise as considered in all simulation results throughout this chapter, $\mathbf{d}_{\text{ICI}}^{(l)}$ defined in (6.155) is the only relevant source of disturbance. Hence, decreasing α automatically enhances the estimation quality by decreasing $\mathbf{d}_{\text{ICI}}^{(l)}$, but this is simply due to normalization issues. In order to enable a fair comparison with CP-OFDM, let $\alpha = \sqrt{N_d'/N_d}$, with N_d' denoting the number of data subcarriers of the reference CP-OFDM system, to ensure that the data induced mean power per non-pilot subcarrier is the same for both systems. This is even slightly advantageous for CP-OFDM, as in UW-OFDM the total mean power per non-pilot subcarrier is even higher due to the spread of the pilots according to $\mathbf{G}_p\mathbf{p}$. Despite this advantage, the estimation quality in CP-OFDM is significantly beyond that of UW-OFDM, see Fig. 6.14. The results for the latter are obtained for a CAZAC sequence as UW, however, they only slightly vary for a zero or any

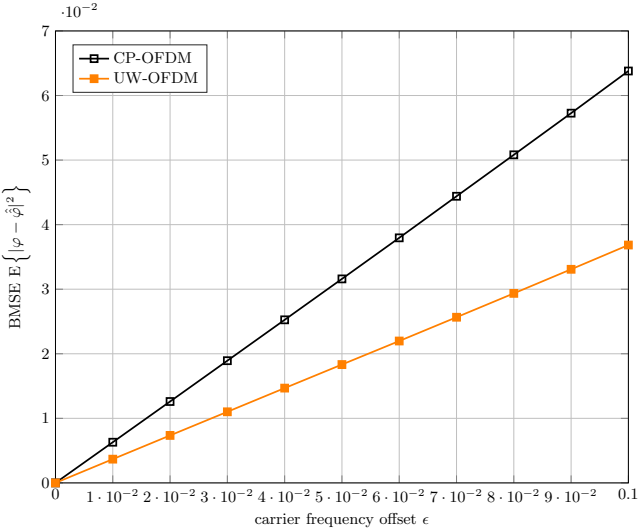


Figure 6.14.: Comparison of estimation error of the CFO induced symbol rotation φ_l between UW-OFDM and CP-OFDM in terms of Bayesian MSE. Investigated scenario: $\mathbf{G}'_d$, CAZAC sequence as UW, $\mathbf{W}_p = |\tilde{\mathbf{H}}_p|^2$, and multipath environment with $\tau_{\mathrm{RMS}} = 100\,\mathrm{ns}$.

other non-zero UW. Hence, the UW does not cause the difference in estimation performance. The reason for the difference gets apparent, when evaluating the mean power of the data induced intercarrier interference

$$\sigma^2_{d_{\mathrm{ICI}}} = \frac{1}{N_p}\sum_{k=0}^{N_p-1} \mathrm{E}\left\{ d^{(l)}_{\mathrm{ICI}}[k] d^{(l)}_{\mathrm{ICI}}[k]^H \right\} \tag{6.193}$$

$$= \frac{1}{N_p}\sum_{k=0}^{N_p-1} E\mathbf{e}_k^T \tilde{\mathbf{\Lambda}}_H^{(l)} \mathbf{G}_d \mathbf{d}^{(l)} \mathbf{d}^{(l)H} \mathbf{G}_d^H \tilde{\mathbf{\Lambda}}_H^{(l)H} (\mathbf{e}_k^T)^H \tag{6.194}$$

$$= \sigma_d^2 \frac{1}{N_p}\sum_{k=0}^{N_p-1} \mathbf{e}_k^T \tilde{\mathbf{\Lambda}}_{H,\mathrm{stat}} \mathbf{G}_d \mathbf{G}_d^H \tilde{\mathbf{\Lambda}}_{H,\mathrm{stat}}^H (\mathbf{e}_k^T)^H, \tag{6.195}$$

where

$$\tilde{\mathbf{\Lambda}}_H^{(l)} = \mathrm{e}^{j\varphi_l} \tilde{\mathbf{\Lambda}}_{H,\mathrm{stat}} \tag{6.196}$$

and $d^{(l)}_{\mathrm{ICI}}[k]$ denotes the kth element of the vector $\mathbf{d}^{(l)}_{\mathrm{ICI}}$ in (6.155). In order to meet the notation in (6.195), CP-OFDM is modelled by a generator matrix $\mathbf{G}_d = \mathbf{P}\begin{bmatrix}\mathbf{I} & \mathbf{0}\end{bmatrix}^T$, with a permutation matrix determining the correct position of the

data symbols. The correlations introduced by the UW-OFDM generator matrix reduce the resulting interferences significantly, yielding $\sigma^2_{d_{\mathrm{ICI}},\mathrm{cp}} > \sigma^2_{d_{\mathrm{ICI}},\mathrm{uw}}$ for an increasing CFO as shown in Fig. 6.15. The mean power of the pilot induced

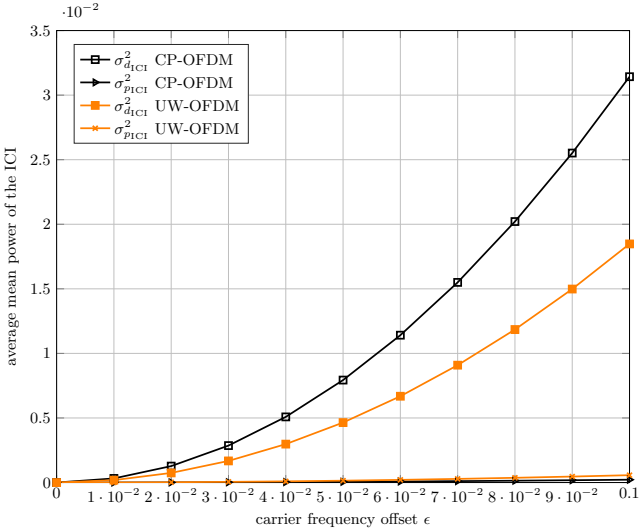


Figure 6.15.: Comparison between UW-OFDM and CP-OFDM of the average mean power of the ICI experienced by a pilot symbol. The ICI is separated into the data part and the part induced by the other pilots. Investigated scenario: $\mathbf{G}'_d$, CAZAC sequence as UW, and multipath environment with $\tau_{\mathrm{RMS}} = 100\,\mathrm{ns}$.

intercarrier interference

$$\sigma^2_{p_{\mathrm{ICI}}} = \frac{1}{N_p} \sum_{k=0}^{N_p-1} \mathrm{E}\left\{ p^{(l)}_{\mathrm{ICI}}[k] p^{(l)}_{\mathrm{ICI}}[k]^H \right\} \tag{6.197}$$

with $p^{(l)}_{\mathrm{ICI}}[k] = \mathbf{e}_k^T \tilde{\mathbf{\Lambda}}^{(l)}_H \sum_{m=0,m\neq k}^{N_p-1} \mathbf{g}_m p[m]$ from (6.162) is a little bit higher for UW-OFDM, but as expected of no relevance. In the end, UW-OFDM clearly outperforms CP-OFDM in terms of pilot based estimation of the CFO induced phase rotation φ_l.

6.5.3. Estimation Setup

For the results presented in Sec. 6.3 and 6.4, ϵ has been estimated once per burst based on the estimated phase rotation of the first OFDM symbol $\hat{\varphi}_0$

$$\hat{\epsilon}_l = \hat{\epsilon} = \epsilon + \Delta_{\epsilon_0} = \hat{\varphi}_0 \frac{N}{2\pi \left(\frac{N-1}{2} + N_u + N_{\text{pil}}\right)}. \tag{6.198}$$

As mentioned, averaging over several estimates would significantly improve the estimation accuracy, the presented results utilizing $\hat{\epsilon}$ from (6.198) represent thus a lower performance bound. The phase rotations $\hat{\varphi}_l$ are estimated according to (6.149) with $\mathbf{W}_p = |\tilde{\mathbf{H}}_p|^2$. The equalizer is the same as for CP-OFDM and has been chosen to be $\mathbf{E}_p = \begin{bmatrix}\mathbf{0} & \mathbf{I}\end{bmatrix} \mathbf{P}_p^T$. The results are obtained for the noise free case to distinctly elaborate the degrading effects of a CFO. Note that the main purpose of this chapter is not to find the best estimation method for φ_l and ϵ, the aim is to show the sensibility of UW-OFDM regarding CFO and to compare it against CP-OFDM. Therefore, a simple pilot based estimation is not only sufficient, the same estimation method can also be used for both UW-OFDM and CP-OFDM. This e.g., would not be possible, if the UW would be used in the estimation process). Consequently, this decision allows to focus all considerations on the CFO resilience of both systems.

6.6. Bit Error Ratio Simulations

The previous sections have isolated the CFO effect on UW-OFDM to determine its impact without being superimposed by other effects. This section aims at investigating the CFO impairments on UW-OFDM in the context of a whole communication system. BER simulations for coded as well as uncoded transmission in a frequency selective environment serve as a performance means. The results are obtained in the same way as in Sec. 3.5.3 and 4.3.2 by averaging over a fixed set of 10^4 different channel instances (see Sec. 2.4.3) drawn from the same channel model (see Sec. 2.4.2). Simulations have been conducted for various CFO values within the range $0.0 \leq \epsilon \leq 0.1$. Since all lead to a performance within the corridor spanned by the boarder values 0.0 and 0.1, the results corresponding to intermediate CFO values are omitted to enhance clarity in the figures. Further, a reference curve determining the BER results without CFO is included in the following figures to provide principle performance bounds. Note that the UW-OFDM system utilized in this chapter originates from chapter 5 and is not completely identical to the one used in chapter 4. Hence, the curves without CFO presented in the following are also not identical with the ones in chapter 4. Since the pilots take over the required estimation task of the CFO induced phase rotation φ_l, a zero-word as UW is applied. Except for the beforementioned reference curves, all presented figures always

incorporate a CFO compensation by multiplying the lth OFDM symbol with $\mathrm{e}^{-j\hat{\varphi}_l}$, which is therefore not explicitly labeled in the legend.

Fig. 6.16 compares the BER performance of UW-OFDM and CP-OFDM for $\epsilon = 0.0$ and $\epsilon = 0.1$ for uncoded transmission. As expectable and similar to the results without CFO in Fig. 4.9, UW-OFDM significantly outperforms CP-OFDM in the high E_b/N_0 region. Both systems though experience a saturating BER characteristic in case of $\epsilon = 0.1$. Note that the curves without CFO slightly differ from $\epsilon = 0.0$, as in the latter case, the CFO estimation algorithm might still detect $\hat{\varphi}_l \neq 0$ (or eqivalently $\hat{\epsilon} \neq 0$) and therefore introduce a certain error. However, the simulations reveal differences only in the low E_b/N_0 region, which is better visible in Fig. 6.17, a zoomed version of Fig. 6.16.

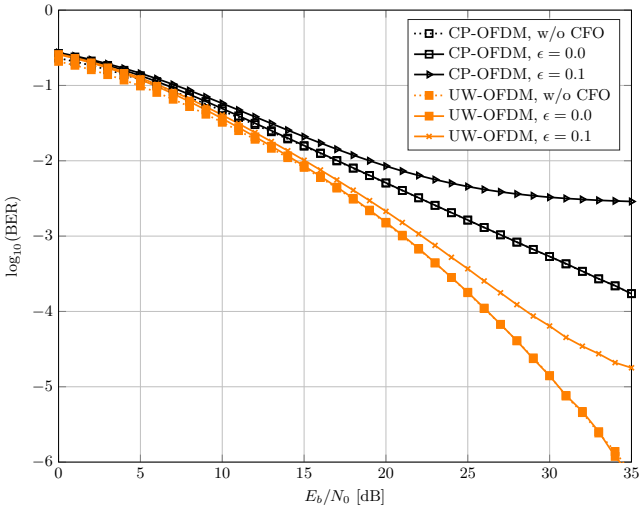


Figure 6.16.: BER comparison of UW-OFDM and CP-OFDM for uncoded transmission in the presence of a CFO in a multipath environment with $\tau_{\mathrm{RMS}} = 100\,\mathrm{ns}$.

BER results for coded transmission with a coding rate of $r = 1/2$ are illustrated in Fig. 6.18. The two most right curves correspond to $\epsilon = 0.1$ and testify a performance difference of 0.8 dB in favor of UW-OFDM. The two most left curves denote the case without CFO with a difference of 0.9 dB, again advantageous for UW-OFDM. The tendency of both results were expectable this way. Hence, this makes the outcome for $\epsilon = 0.0$ represented by the two middle curves even more surprising. Of course, both UW-OFDM and CP-OFDM would experience a performance loss compared to the case without CFO, a conclusion already indicated by

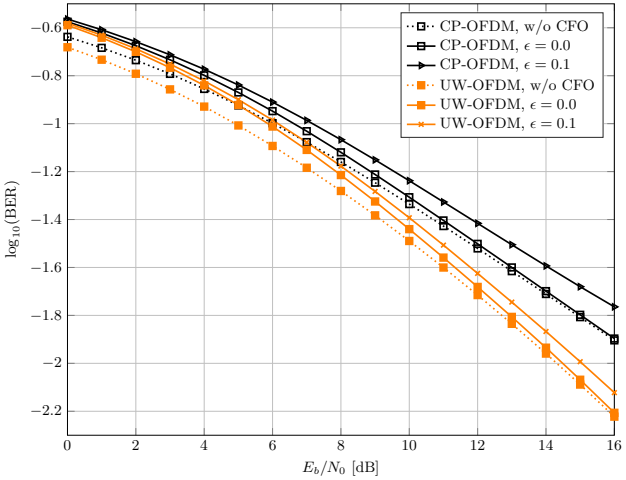


Figure 6.17.: BER comparison of UW-OFDM and CP-OFDM for uncoded transmission in the presence of a CFO in a multipath environment with $\tau_{\mathrm{RMS}} = 100\,\mathrm{ns}$ (zoom).

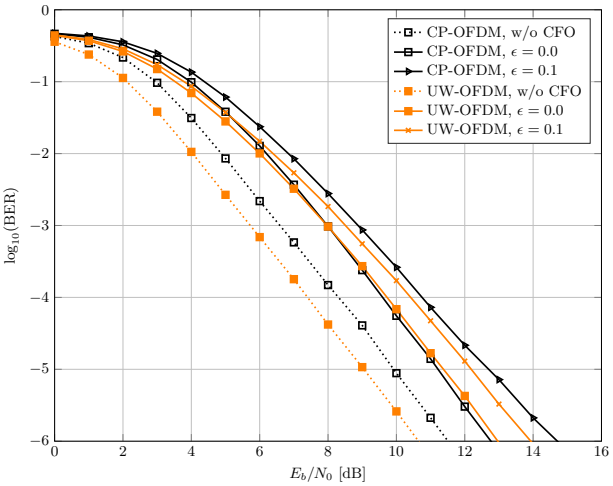


Figure 6.18.: BER comparison of UW-OFDM and CP-OFDM for coded transmission with $r = 1/2$ in the presence of a CFO in a multipath environment with $\tau_{\mathrm{RMS}} = 100\,\mathrm{ns}$.

the uncoded BER curves in Fig. 6.17. However, the results in Fig. 6.17 also identify UW-OFDM as the better system, it is thus not immediately apparent why this has changed in the coded case. In order to determine the cause, Fig. 6.19 displays the results for coded transmission with hard decision decoding. This means that the

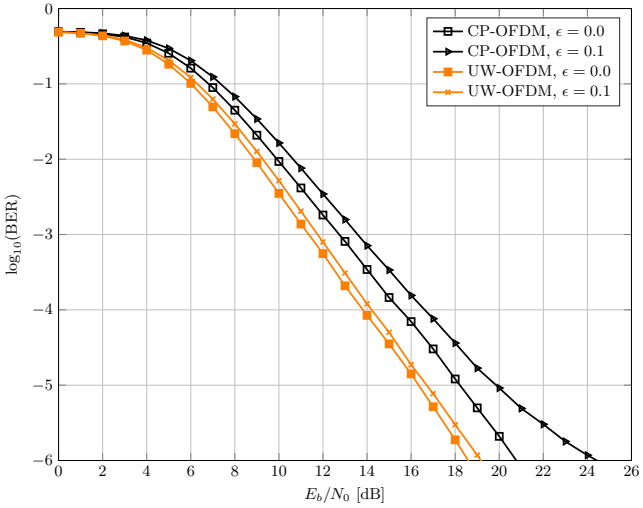


Figure 6.19.: BER comparison of UW-OFDM and CP-OFDM for coded transmission with $r = 1/2$ and hard decision decoding in the presence of a CFO in a multipath environment with $\tau_{\mathrm{RMS}} = 100\,\mathrm{ns}$.

reliability information available in form of LLRs is not used (details to the LLRs are given in appendix A). Instead, the Viterbi decoder is simply fed with a sequence of binary symbols associated with the transmit data symbols that are closest to the data estimates (e.g., obtained after LMMSE estimation) in terms of the Euclidean distance. Here, UW-OFDM beats CP-OFDM again in any case, namely by 2.2 dB and 5.8 dB for $\epsilon = 0.0$ and $\epsilon = 0.1$, respectively.

To sum up all findings so far, UW-OFDM beats CP-OFDM in any case for uncoded transmission as well as coded transmission with hard decision decoding. Thus, the reason for the unexpected performance loss of UW-OFDM over CP-OFDM in case of soft decision decoding has to lie in an inaccurate reliability information that is handed over to the Viterbi decoder. This inaccuracy results from not incorporating the CFO effects into the LLR values, but nevertheless, both UW-OFDM and CP-OFDM suffer from the same lack of information. Hence, it seems that UW-OFDM is more sensitive on the CFO induced inaccuracy than CP-OFDM. According to the analysis in Sec. 6.1, a CFO induces three effects on a UW-OFDM subcarrier symbol. First, the symbol will be attenuated, but this effect can be neglected. Second,

each symbol experiences intercarrier interference, which has mathematically been modelled by $\tilde{\mathbf{\Lambda}}_{\text{stat}}$ in this chapter. Third, each UW-OFDM frequency domain symbol experiences a phase rotation by φ_l. In order to determine the critical source of error within the LLRs, Fig. 6.20 visualizes the BER performance for soft decision decoding in case of $\hat{\varphi}_l = \varphi_l$, leaving the intercarrier interference and the negligible attenuation as only sources not reflected in the reliability information. Now, UW-OFDM outperforms CP-OFDM again in any case, even if soft decision decoding is applied. In conclusion, this leaves the non-incorporation of the error induced by derotating with $\hat{\varphi}_l \neq \varphi_l$ in the LLRs as responsible source of the performance degradation of UW-OFDM. In other words, only with an adequate error model, UW-OFDM maybe able to surpass CP-OFDM in every scenario.

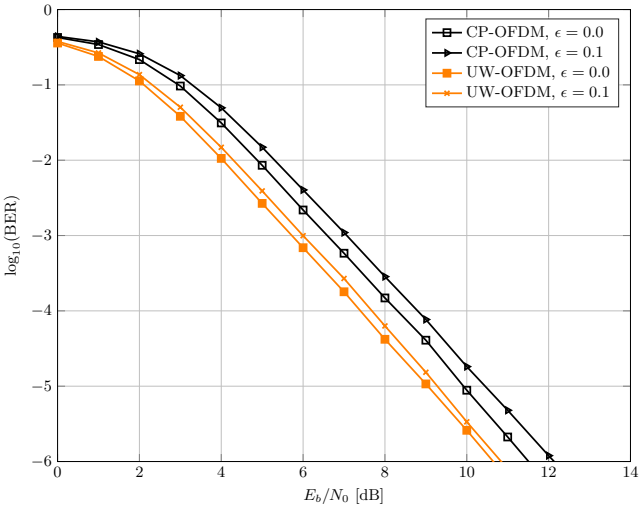


Figure 6.20.: BER comparison of UW-OFDM and CP-OFDM for coded transmission with $r = 1/2$ and perfect derotation with $\hat{\varphi}_l = \varphi_l$ in the presence of a CFO in a multipath environment with $\tau_{\text{RMS}} = 100\,\text{ns}$.

Fig. 6.21 provides BER values for coded transmission with a coding rate of $r = 3/4$. Similar to $r = 1/2$, the results for the case without CFO defeat the case with $\epsilon = 0.0$. However, a substantial difference is only visible in the lower E_b/N_0 region till approximately 12 dB. After that the error in $\hat{\varphi}_l$ is small enough such that it is not reflected in the BER performance anymore. Contrary to $r = 1/2$, UW-OFDM is now superior to CP-OFDM in any case with soft decision decoding and therefore regardless of the error in the LLRs introduced by inaccurate derotations with $\hat{\varphi}_l$. The gain of UW-OFDM over CP-OFDM is 1.5 dB for $\epsilon = 0.0$ and increases to

3.4 dB for $\epsilon = 0.1$. For the high E_b/N_0 range, this behavior is easily explained by the accurate estimates $\hat{\varphi}_l$, leading to LLR values with negligible approximation error. For the low E_b/N_0 range, errors in the LLRs in the same scale as for $r = 1/2$ occur. Nonetheless, the principle gain of UW-OFDM over CP-OFDM, i.e., without a CFO, is significantly larger for $r = 3/4$ compared to $r = 1/2$. Further, this punctured code seems to be less sensitive to errors in the reliability information. In contrast to the case for $r = 1/2$, these errors do not translate to a BER performance degradation of UW-OFDM in a scale such that in certain cases CP-OFDM shows better results than UW-OFDM.

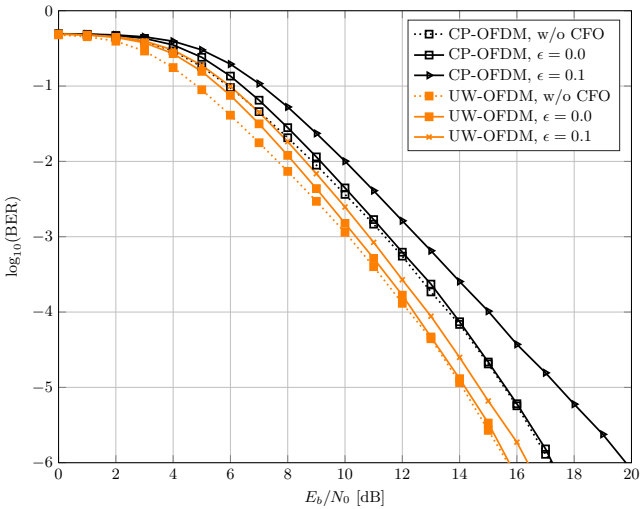


Figure 6.21.: BER comparison of UW-OFDM and CP-OFDM for coded transmission with $r = 3/4$ in the presence of a CFO in a multipath environment with $\tau_{\mathrm{RMS}} = 100\,\mathrm{ns}$.

So far, only the simple method of derotating by $\hat{\varphi}_l$ has been considered in the BER simulations. In a next step, the more sophisticated methods from Sec. 6.4 shall be investigated. Since a considerable gain over the derotating approach can only be expected in the high CFO range, cf. Fig. 6.12, only the case of $\epsilon = 0.1$ is treated. Furthermore, simulations demonstrated that the biggest gain is achieved for $r = 3/4$, hence elaborations are restricted to this case. Two advanced methods are considered with the following compensation parameters on top of derotating by $\hat{\varphi}_l$:

- φ_{off} and $\tilde{\mathbf{\Lambda}}_{\text{stat}}^{-1}$, or
- $\hat{\varphi}_{\text{off}}$ and $\hat{\tilde{\mathbf{\Lambda}}}_{\text{stat}}^{H}$.

Both concepts have already been analyzed in terms of their MSE in Fig. 6.12, whereas the first one denotes a principle performance bound, and the second one a low cost implementation of the first concept. For the latter, $\hat{\epsilon}$ is obtained by averaging over 200 estimates from 200 OFDM symbols. Fig. 6.22 shows for UW-OFDM and CP-OFDM in each case 4 curves, whereas two serve as reference. One denotes the performance, if the CFO is compensated only by derotating with $\hat{\varphi}_l$, and one curve represents the performance in case of $\epsilon = 0.0$. Both have already been part in Fig. 6.21. Starting with the best method, i.e., additionally compensating with $\tilde{\mathbf{\Lambda}}_{\text{stat}}^{-1}$ and φ_{off} improves the BER performance of UW-OFDM by 0.5 dB, reducing the distance to the optimum without CFO to 0.2 dB. The most impressive gain, however, is obtained for the CP-OFDM system. Here, the performance improves by 2.3 dB at a BER of 10^{-6}, leaving a residual gap to the optimum case of 0.3 dB. These results have already been indicated by the MSE analysis, which showed that the performance difference between UW-OFDM and CP-OFDM reduces for the advanced CFO compensation methods.

Unfortunately and not supported by the MSE analysis, the computationally less complex method with $\hat{\tilde{\mathbf{\Lambda}}}_{\text{stat}}^{H}$ and $\hat{\varphi}_{\text{off}}$ cannot keep up with the other advanced compensation method. The gain over simply derotating with $\hat{\varphi}_l$ reduces to 0.2 dB for UW-OFDM and 0.8 dB for CP-OFDM. These results come by some surprise, since the MSE analysis in Fig. 6.12 identified both approaches as almost equally good, therefore a similar BER performance might be expected. More in-depth investigations by evaluating the results for $\tilde{\mathbf{\Lambda}}_{\text{stat}}^{H}$ and $\hat{\varphi}_{\text{off}}$ revealed that the delimiter is not the simplification of the matrix inversion by the Hermitian operator rather than the estimation accuracy of $\hat{\epsilon}$ to form $\hat{\tilde{\mathbf{\Lambda}}}_{\text{stat}}$. More specifically, in the MSE analysis of Sec. 6.4, the only source of disturbance was given by the data subcarrier interference, but now an additional additive noise term is present. In this case, it does not even help that $\hat{\epsilon}$ is the result of averaging over 200 single estimates.

In conclusion, compensating the CFO effects by advanced methods on top of a derotation by $\hat{\varphi}_l$ only pays off, if a very accurate estimate of ϵ is available. Oth-

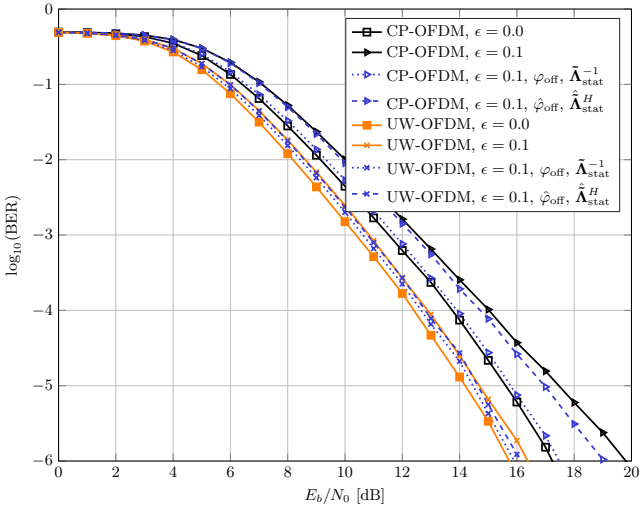


Figure 6.22.: BER comparison of UW-OFDM and CP-OFDM for coded transmission with $r = 3/4$ in the presence of a CFO in a multipath environment with $\tau_{\text{RMS}} = 100\,\text{ns}$.

erwise, the expectable gain is rather limited and the computational effort might prevent an application.

7. Conclusion

In this work, Unique Word OFDM as an attractive alternative to current OFDM schemes was considered, promising to surmount the well-known inefficiency inherent to guard intervals by implementing the interval in a fashion that induces a variety of beneficial properties. Unique Word OFDM denotes a signaling framework rather than a single scheme, offering a vast amount of different realizations. Main focus of this work has been laid on the different signal designs and their impact on the performance, with special focus on BER results in AWGN and multipath channels for coded as well as uncoded transmission. CP-OFDM based on the IEEE 802.11a standard served as reference system in most cases.

Systematically encoded UW-OFDM was presented as the first principle approach to generate UW-OFDM signals, implemented by applying appropriate generator matrices. Characteristic to this concept is the introduction of dedicated data and redundant subcarriers, the latter loaded with a defined linear mapping of the data to fulfill the specific UW time domain properties. The applicability of the systematic approach highly depends on the energy spent for the redundant subcarriers. Investigations shed light on the influence of setup parameters such as amount of total subcarriers, number and position of zero subcarriers as well as the proportion to the data energy. The most crucial factor determining the redundant energy turned out to be the distribution of the redundant subcarriers among the available bandwidth. Optimization algorithms for finding good distributions were investigated and suitable ones highlighted. Subsequent elaborations proved that designing this class of UW-OFDM signals based on minimizing the redundant energy is indeed also optimum w.r.t. the whole transceiver performance. Investigated cost functions in this context were the minimization of the sum of the error variances after a BLUE or an LMMSE estimator in AWGN at a given SNR. Analytical as well as numerical investigations confirmed independent of the cost function that a two-step generation of the UW signal is always superior to a direct generation approach. It is thus most beneficial to generate a UW-OFDM signal with a zero-word in a first, and add the desired UW in a second step. In order to reduce energy further, additional redundant subcarriers and the concept of systematic noise were introduced. The resulting BER improvement entailed a decreased bandwidth efficiency in the first and an inevitable BER error floor most of the time in the second case.

Non-systematically encoded UW-OFDM solves the problem of the excess redundant energy by lifting the role dedication from the subcarriers and introducing

all-purpose subcarriers instead. A thorough signal analysis established the properties of optimum generator matrices that overall minimize the sum of the error variances after a BLUE or an LMMSE estimator in an AWGN channel at a given SNR. A presented steepest descent algorithm is able to deliver different optimum generator matrices, with the resulting design as a function of the initialization of the algorithm. Two exemplary UW-OFDM systems, one rather sharing ties with conventional OFDM and one with single carrier systems, were studied in detail. The generator matrix delivering the best BER performance varied depending on the utilized channel encoder. Further experiments confirmed BER performance advantages of UW-OFDM over CP-OFDM regardless of the specific setup. Simulations were also conducted for higher order modulations such as 16-QAM and multipath channels with a channel delay spread exceeding the guard interval, again in favor of UW-OFDM. Both systems showed a performance degradation in the same scale for imperfect channel knowledge.

The insertion of pilot tones into UW-OFDM frequency domain symbols was discussed and meaningful optimization criteria as well as parameters identified. Splitting the frequency domain symbol into independent data and pilot terms enabled an optimization of the OFDM symbol towards data and system parameter estimation at the same time.

Comprehensive elaborations identified the CFO effects experienced by UW-OFDM subcarrier symbols to be a combination of those for conventional OFDM and single carrier systems. The specific appearance of the three CFO induced effects attenuation, phase rotation and intercarrier interference depends on the specific UW-OFDM realization. The CFO induced error due to subtracting the wrong UW and pilot offset was shown to be corrected almost completely with moderate effort. Advanced CFO compensation methods on top of a simple OFDM symbol phase derotation only pay off, if very accurate CFO estimates are available. Since UW-OFDM is already quite robust against CFO, the bigger potential of these methods is given for CP-OFDM. As for the case without CFO and supported by MSE and BER simulations, UW-OFDM generally outperforms CP-OFDM, whereas the gain increases with increasing CFO. In selected scenarios, UW-OFDM requires a more accurate reliability information fed to the channel decoder than CP-OFDM to preserve its strengths.

A. LLR calculation

Appendix A comprises the calculation of the so-called log-likelihood ratios (LLRs). The LLRs serve as reliability information of the received and disturbed binary symbols which shall be recovered at the receiver and are fed as soft information into the decoder. Note that major parts of the derivations are adaptations from [88, 89, 90]. Based on the linear model in (2.68), let us start from

$$\tilde{\mathbf{y}} = \underbrace{\tilde{\mathbf{H}}\mathbf{G}}_{\mathbf{H}}\mathbf{d} + \mathbf{v}, \tag{A.1}$$

where $\mathbf{H} \in \mathbb{C}^{(N_d+N_r)\times N_d}$ denotes a generalized channel matrix and $\mathbf{v} \in \mathbb{C}^{(N_d+N_r)\times 1}$ an additive white Gaussian noise vector with $\mathbf{v} \sim \mathcal{CN}(\mathbf{0}, \sigma_v^2\mathbf{I})$. Further, $\mathbf{d} \in \mathcal{A}^{N_d\times 1}$ consists of N_d data symbols, where each symbol is drawn from a symbol alphabet $\mathcal{A}$, usually quadrature amplitude modulation (QAM), phase shift keying (PSK) or amplitude shift keying (ASK) constellations. Depending on the specific constellation, each symbol is formed by a certain number of binary symbols b_i represented by $b_i \in \{0, 1\}$. A linear estimator $\mathbf{E}$ is assumed throughout this work, yielding an estimate

$$\hat{\mathbf{d}} = \mathbf{E}\tilde{\mathbf{y}}. \tag{A.2}$$

For reasons of notational simplicity, let us focus on the kth estimated element

$$\hat{d}_k = \mathbf{e}_k^H\tilde{\mathbf{y}}, \tag{A.3}$$

with $\mathbf{e}_k^H$ denoting the kth row of $\mathbf{E}$. For this general estimator, the LLRs of the bits mapped onto any symbol constellation can be written as

$$\mathrm{L}\left(b_{ik}\middle|\hat{d}_k\right) = \ln\frac{\Pr\left(b_{ik}=1\middle|\hat{d}_k\right)}{\Pr\left(b_{ik}=0\middle|\hat{d}_k\right)}, \tag{A.4}$$

where b_{ik} is the ith bit of the kth received symbol. Hence, $\mathrm{L}\left(b_{ik}\middle|\hat{d}_k\right)$ denotes the LLR of the ith bit within the kth estimated data symbol. Let $S\left(b_{ik}=1\right)$ and $S(b_{ik}=0)$ be the set of symbol indices corresponding to $b_{ik}=1$ and $b_{ik}=0$, respectively. Furthermore, let $s^{(q)}$ be the symbol corresponding to $q \in S$, then

$$\Pr\left(b_{ik}=1\middle|\hat{d}_k\right) = \sum_{q\in S(b_{ik}=1)}\Pr\left(d_k=s^{(q)}\middle|\hat{d}_k\right), \tag{A.5}$$

$$\Pr\left(b_{ik}=0\middle|\hat{d}_k\right)=\sum_{q\in S(b_{ik}=0)}\Pr\left(d_k=s^{(q)}\middle|\hat{d}_k\right). \tag{A.6}$$

Applying Bayes' theorem and taking into account equiprobable input symbols, the probabilities can be changed to probability densities, transforming (A.4) into

$$\mathrm{L}\left(b_{ik}\middle|\hat{d}_k\right)=\ln\frac{\sum\limits_{q\in S(b_{ik}=1)}p\left(\hat{d}_k\middle|s^{(q)}\right)}{\sum\limits_{q\in S(b_{ik}=0)}p\left(\hat{d}_k\middle|s^{(q)}\right)}, \tag{A.7}$$

where $p\left(\hat{d}_k\middle|s^{(q)}\right)$ denotes the conditional PDF of the estimate $\hat{d}_k$ given that the actually transmitted symbol was $s^{(q)}$. Assuming now that the data estimate $\hat{d}_k$ is the result of the data symbol $d_k=s^{(q)}$ disturbed by AWGN, $p\left(\hat{d}_k\middle|s^{(q)}\right)$ can be written in the general form

$$p\left(\hat{d}_k\middle|s^{(q)}\right)=\frac{1}{\pi\mathrm{var}\left(\hat{d}_k\middle|s^{(q)}\right)}\mathrm{e}^{-\frac{1}{\mathrm{var}\left(\hat{d}_k\middle|s^{(q)}\right)}\left|\hat{d}_k-\mathrm{E}\left\{\hat{d}_k\middle|s^{(q)}\right\}\right|^2}. \tag{A.8}$$

Hence, in order to determine $\mathrm{L}\left(b_{ik}\middle|\hat{d}_k\right)$, the conditional variance $\mathrm{var}\left(\hat{d}_k\middle|s^{(q)}\right)$ and the conditional mean $\mathrm{E}\left\{\hat{d}_k\middle|s^{(q)}\right\}$ are required, which are derived in the following.

To enable the formulation of the LLRs for different estimators, let us take a generalized approach from [89] and reformulate the linear model in (A.1) as

$$\tilde{\mathbf{y}}=\mathbf{H}\mathbf{d}+\mathbf{v} \tag{A.9}$$

$$=\mathbf{h}_k d_k+\bar{\mathbf{H}}_k\bar{\mathbf{d}}_k+\mathbf{v}, \tag{A.10}$$

where $\mathbf{h}_k$ is the kth column of $\mathbf{H}$, $\bar{\mathbf{H}}_k$ is $\mathbf{H}$ without $\mathbf{h}_k$ and $\bar{\mathbf{d}}_k$ is $\mathbf{d}$ without d_k. Then $\hat{d}_k$ defined in (A.3) is given as

$$\hat{d}_k=\mathbf{e}_k^H(\mathbf{h}_k d_k+\bar{\mathbf{H}}_k\bar{\mathbf{d}}_k+\mathbf{v}) \tag{A.11}$$

$$=\underbrace{\mathbf{e}_k^H\mathbf{h}_k}_{\text{Scaling}}d_k+\underbrace{\mathbf{e}_k^H\bar{\mathbf{H}}_k\bar{\mathbf{d}}_k}_{\text{IPI}}+\underbrace{\mathbf{e}_k^H\mathbf{v}}_{\text{Noise}}. \tag{A.12}$$

An estimate $\hat{d}_k$ incorporates now three effects, namely a scaling of the true parameter value, an inter-parameter interference (IPI) term and a classical additive noise term. The specific realization of each effect varies from estimator to estimator, dedicated values are given for a BLUE and an LMMSE estimator at the end of this appendix.

The additive noise term in (A.12) has already been specified to be Gaussian, the IPI term can due to central limit theorem arguments usually also be assumed to be

Gaussian, especially when considering the rather high values for N_d in this work. Hence, the primarily assumption of $p\left(\hat{d}_k \middle| s^{(q)}\right)$ following a Gaussian distribution is justified.

The conditional mean of $\hat{d}_k$ becomes

$$\mathrm{E}_{\tilde{\mathbf{y}}|d_k}\left\{\hat{d}_k \middle| d_k\right\} = \mathbf{e}_k^H \mathbf{h}_k d_k + \mathbf{e}_k^H \bar{\mathbf{H}}_k \mathrm{E}_{\bar{\mathbf{d}}_k|d_k}\left\{\bar{\mathbf{d}}_k \middle| d_k\right\}. \tag{A.13}$$

Considering the elements of $\mathbf{d}$ to be statistically independent and with zero mean, (A.13) simplifies to

$$\mathrm{E}_{\tilde{\mathbf{y}}|d_k}\left\{\hat{d}_k \middle| d_k\right\} = \underbrace{\mathbf{e}_k^H \mathbf{h}_k}_{\alpha_k} d_k, \tag{A.14}$$

with α_k corresponding to a scaling factor. The conditional variance of the general linear estimator follows to

$$\sigma_k^2 = \mathrm{var}\left(\hat{d}_k \middle| d_k\right) = \mathrm{E}_{\tilde{\mathbf{y}}|d_k}\left\{\left(\hat{d}_k - \mathrm{E}_{\tilde{\mathbf{y}}|d_k}\left\{\hat{d}_k \middle| d_k\right\}\right)\left(\hat{d}_k - \mathrm{E}_{\tilde{\mathbf{y}}|d_k}\left\{\hat{d}_k \middle| d_k\right\}\right)^H \middle| d_k\right\}. \tag{A.15}$$

Inserting (A.12) and (A.14) into (A.15) yields

$$\begin{aligned}
\sigma_k^2 =& \mathrm{E}_{\tilde{\mathbf{y}}|d_k}\left\{\left(\mathbf{e}_k^H(\bar{\mathbf{H}}_k\bar{\mathbf{d}}_k + \mathbf{v})\right)\left(\mathbf{e}_k^H(\bar{\mathbf{H}}_k\bar{\mathbf{d}}_k + \mathbf{v})\right)^H \middle| d_k\right\} && \text{(A.16)}\\
=& \mathbf{e}_k^H\left(\bar{\mathbf{H}}_k \mathbf{C}_{\bar{\mathbf{d}}_k\bar{\mathbf{d}}_k} \bar{\mathbf{H}}_k^H + \mathbf{C}_{\mathbf{vv}}\right)\mathbf{e}_k && \text{(A.17)}\\
=& \mathbf{e}_k^H \mathbf{C}_k \mathbf{e}_k, && \text{(A.18)}\\
& && \text{(A.19)}
\end{aligned}$$

where $\mathbf{C}_k = \left(\bar{\mathbf{H}}_k \mathbf{C}_{\bar{\mathbf{d}}_k\bar{\mathbf{d}}_k} \bar{\mathbf{H}}_k^H + \mathbf{C}_{\mathbf{vv}}\right)$. With (A.14) and σ_k^2, the conditional PDF $p\left(\hat{d}_k \middle| s^{(q)}\right)$ is given as

$$p\left(\hat{d}_k \middle| s^{(q)}\right) = \frac{1}{\pi\sigma_k^2} \mathrm{e}^{-\frac{1}{\sigma_k^2}\left|\hat{d}_k - \alpha_k s^{(q)}\right|^2}, \tag{A.20}$$

and the LLR for the ith bit of the kth symbol finally follows as

$$\mathrm{L}\left(b_{ik} \middle| \hat{d}_k\right) = \ln \frac{\sum\limits_{q \in S(b_{ik}=1)} \frac{1}{\pi\sigma_k^2} \mathrm{e}^{-\frac{1}{\sigma_k^2}\left|\hat{d}_k - \alpha_k s^{(q)}\right|^2}}{\sum\limits_{q \in S(b_{ik}=0)} \frac{1}{\pi\sigma_k^2} \mathrm{e}^{-\frac{1}{\sigma_k^2}\left|\hat{d}_k - \alpha_k s^{(q)}\right|^2}}. \tag{A.21}$$

Note that (A.21) is a very general definition that applies for various estimators and an arbitrary symbol alphabet.

The next part evaluates (A.21) for the specific case of QPSK as modulation alphabet, where one symbol is formed out of bit b_0 and b_1. Further, the impact of a BLUE

and an LMMSE estimator on the LLRs is shown. Splitting the squared absolute values $\left|\hat{d}_k - \alpha_k s^{(q)}\right|^2$ into their real and imaginary part yields

$$\left|\hat{d}_k - \alpha_k s^{(q)}\right|^2 = \mathrm{Re}\left\{\hat{d}_k - \alpha_k s^{(q)}\right\}^2 + \mathrm{Im}\left\{\hat{d}_k - \alpha_k s^{(q)}\right\}^2 \tag{A.22}$$

$$= \left(\hat{d}_{k,\mathbb{R}} - \alpha_k s_{\mathbb{R}}^{(q)}\right)^2 + \left(\hat{d}_{k,\mathbb{I}} - \alpha_k s_{\mathbb{I}}^{(q)}\right)^2, \tag{A.23}$$

and the LLR definition in (A.21) translates to

$$\mathrm{L}\left(b_{ik}\middle|\hat{d}_k\right) = \ln \frac{\sum\limits_{q\in S(b_{ik}=1)} \frac{1}{\pi\sigma_k^2} \mathrm{e}^{-\frac{1}{\sigma_k^2}\left(\left(\hat{d}_{k,\mathbb{R}} - \alpha_k s_{\mathbb{R}}^{(q)}\right)^2 + \left(\hat{d}_{k,\mathbb{I}} - \alpha_k s_{\mathbb{I}}^{(q)}\right)^2\right)}}{\sum\limits_{q\in S(b_{ik}=0)} \frac{1}{\pi\sigma_k^2} \mathrm{e}^{-\frac{1}{\sigma_k^2}\left(\left(\hat{d}_{k,\mathbb{R}} - \alpha_k s_{\mathbb{R}}^{(q)}\right)^2 + \left(\hat{d}_{k,\mathbb{I}} - \alpha_k s_{\mathbb{I}}^{(q)}\right)^2\right)}}. \tag{A.24}$$

Assuming a mapping of the bit combinations $b_0 b_1 \in \{11, 10, 01, 00\}$ on the QPSK constellation points $s^{(q)} \in \rho\{1+j, 1-j, -1+j, -1-j\}$ with arbitrary scaling factor ρ, the LLR for the bit b_{0k} is given as

$$\mathrm{L}\left(b_{0k}\middle|\hat{d}_k\right) = \ln \frac{\mathrm{e}^{-\frac{1}{\sigma_k^2}\left(\left(\hat{d}_{k,\mathbb{R}} - \alpha_k\rho\right)^2 + \left(\hat{d}_{k,\mathbb{I}} - \alpha_k\rho\right)^2\right)} + \mathrm{e}^{-\frac{1}{\sigma_k^2}\left(\left(\hat{d}_{k,\mathbb{R}} - \alpha_k\rho\right)^2 + \left(\hat{d}_{k,\mathbb{I}} + \alpha_k\rho\right)^2\right)}}{\mathrm{e}^{-\frac{1}{\sigma_k^2}\left(\left(\hat{d}_{k,\mathbb{R}} + \alpha_k\rho\right)^2 + \left(\hat{d}_{k,\mathbb{I}} - \alpha_k\rho\right)^2\right)} + \mathrm{e}^{-\frac{1}{\sigma_k^2}\left(\left(\hat{d}_{k,\mathbb{R}} + \alpha_k\rho\right)^2 + \left(\hat{d}_{k,\mathbb{I}} + \alpha_k\rho\right)^2\right)}} \tag{A.25}$$

$$= \ln \frac{\mathrm{e}^{-\frac{1}{\sigma_k^2}\left(\hat{d}_{k,\mathbb{R}} - \alpha_k\rho\right)^2} \left[\mathrm{e}^{-\frac{1}{\sigma_k^2}\left(\hat{d}_{k,\mathbb{I}} - \alpha_k\rho\right)^2} + \mathrm{e}^{-\frac{1}{\sigma_k^2}\left(\hat{d}_{k,\mathbb{I}} + \alpha_k\rho\right)^2}\right]}{\mathrm{e}^{-\frac{1}{\sigma_k^2}\left(\hat{d}_{k,\mathbb{R}} + \alpha_k\rho\right)^2} \left[\mathrm{e}^{-\frac{1}{2\sigma_k^2}\left(\hat{d}_{k,\mathbb{I}} - \alpha_k\rho\right)^2} + \mathrm{e}^{-\frac{1}{\sigma_k^2}\left(\hat{d}_{k,\mathbb{I}} + \alpha_k\rho\right)^2}\right]} \tag{A.26}$$

$$= \ln \frac{\mathrm{e}^{-\frac{1}{\sigma_k^2}\left(\hat{d}_{k,\mathbb{R}} - \alpha_k\rho\right)^2}}{\mathrm{e}^{-\frac{1}{\sigma_k^2}\left(\hat{d}_{k,\mathbb{R}} + \alpha_k\rho\right)^2}} \tag{A.27}$$

$$= -\frac{1}{\sigma_k^2}\left(\hat{d}_{k,\mathbb{R}} - \alpha_k\rho\right)^2 + \frac{1}{\sigma_k^2}\left(\hat{d}_{k,\mathbb{R}} + \alpha_k\rho\right)^2 \tag{A.28}$$

$$= \frac{4\hat{d}_{k,\mathbb{R}}\alpha_k\rho}{\sigma_k^2} \tag{A.29}$$

A very similar derivation can be conducted for the LLR of the first bit of the kth symbol b_{1k}. With the usual scaling factor of $\rho = \frac{1}{\sqrt{2}}$ for QPSK to obtain symbols with unit variance, the LLR in case of QPSK finally simplifies to

$$\mathrm{L}\left(b_{ik}\middle|\hat{d}_k\right) = \frac{4\hat{d}_{k,\mathbb{R}}\alpha_k}{\sqrt{2}\sigma_k^2} = \frac{2\sqrt{2}\hat{d}_{k,\mathbb{R}}\alpha_k}{\sigma_k^2}, \tag{A.30}$$

with

$$\alpha_k = \mathbf{e}_k^H \mathbf{h}_k, \tag{A.31}$$

$$\sigma_k^2 = \mathbf{e}_k^H (\bar{\mathbf{H}}_k \mathbf{C}_{\bar{\mathbf{d}}_k \bar{\mathbf{d}}_k} \bar{\mathbf{H}}_k^H + \mathbf{C}_{\mathbf{vv}}) \mathbf{e}_k = \mathbf{e}_k^H \mathbf{C}_k \mathbf{e}_k. \tag{A.32}$$

In context of the model in (A.12), α_k represents the scaling effect, whereas σ_k^2 incorporates the effects of the IPI and the noise term together. As shown in [89], both variables depend on the specific estimator:

- In case of the BLUE, $\alpha_k = 1$ and the IPI term vanishes, yielding $\sigma_k^2 = \mathbf{e}_k^H (\mathbf{C}_{\mathbf{vv}}) \mathbf{e}_k$.
- In case of an LMMSE estimator, however,

$$\alpha_k < 1, \tag{A.33}$$

  which reflects the circumstance that an LMMSE estimator is not unbiased anymore in the classical sense. More specifically, unbiasedness is only reached in the Bayesian sense when averaging over the data symbols as well. Fig. 4.8 shows the effect resulting from the biasedness for the example of a non-systematically encoded UW-OFDM system in AWGN. If several estimates would then be plotted in a constellation diagram, the estimate clouds would not be centered around the actually transmitted symbols anymore as it is the case for the BLUE, but the centers would lie a little bit off in the direction of the origin. Finally, not only the value for α_k, but also the IPI term in (A.12) differs compared to the BLUE. In case of the LMMSE estimator, it does not become zero and hence σ_k^2 does not simplify in the same manner. Nevertheless, it can be shown that $[\mathbf{C}_{ee}]_{ii} = \frac{\sigma_k^2}{\alpha_k}$ with $\mathbf{C}_{ee}$ denoting the error covariance matrix of the error $\mathbf{e} = \hat{\mathbf{d}} - \mathbf{d}$, therefore providing some simplification potential as well.

B. Derivation of the Global Minimum of J_{BLUE}

Appendix B will analytically prove that $\partial J_{\mathrm{BLUE}}/\partial \mathbf{s} = \mathbf{0}$, if and only if all singular values of $\mathbf{G}$ are non-zero and identical. Furthermore, it will be shown that $J_{\mathrm{BLUE,min}} = \sigma_d^2 N_d/\gamma$ is also the global minimum.

The cost function defined in (4.7) as

$$J_{\mathrm{BLUE}}(\mathbf{A}) = \frac{\sigma_d^2}{\gamma N_d} \mathrm{tr}\left(\mathbf{G}^H\mathbf{G}\right) \mathrm{tr}\left(\left(\mathbf{G}^H\mathbf{G}\right)^{-1}\right) \tag{B.1}$$

serves as starting point, with the relationship $\mathbf{G} = \mathbf{A}\begin{bmatrix}\mathbf{I} & \mathbf{T}^T\end{bmatrix}^T$ from (2.42). Using the SVD of the generator matrix $\mathbf{G} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$ from (4.16) and $\mathbf{D}^2 = \mathbf{\Sigma}^H\mathbf{\Sigma}$ delivers

$$\begin{aligned}\mathrm{tr}\left(\mathbf{G}^H\mathbf{G}\right) &= \mathrm{tr}\left(\mathbf{V}\mathbf{\Sigma}^H\mathbf{U}^H\mathbf{U}\mathbf{\Sigma}\mathbf{V}^H\right) = \mathrm{tr}\left(\mathbf{V}\mathbf{D}^2\mathbf{V}^H\right)\\ &= \mathrm{tr}\left(\mathbf{D}^2\right), \qquad \text{(B.2)}\\ (\mathbf{G}^H\mathbf{G})^{-1} &= (\mathbf{V}\mathbf{\Sigma}^H\mathbf{U}^H\mathbf{U}\mathbf{\Sigma}\mathbf{V}^H)^{-1} = (\mathbf{V}\mathbf{D}^2\mathbf{V}^H)^{-1}\\ &= \mathbf{V}(\mathbf{D}^2)^{-1}\mathbf{V}^H. \qquad \text{(B.3)}\end{aligned}$$

Inserting (B.2) and (B.3) into (B.1), the cost function is developed to

$$\begin{aligned}J_{\mathrm{BLUE}}(\mathbf{A}) &= \frac{\sigma_d^2}{\gamma N_d}\mathrm{tr}\left(\mathbf{D}^2\right)\mathrm{tr}\left(\mathbf{V}(\mathbf{D}^2)^{-1}\mathbf{V}^H\right)\\ &= \frac{\sigma_d^2}{\gamma N_d}\mathrm{tr}\left(\mathbf{D}^2\right)\mathrm{tr}\left((\mathbf{D}^2)^{-1}\right)\\ &= \frac{\sigma_d^2}{\gamma N_d}\left(s_1^2 + s_2^2 + \cdots s_{N_d}^2\right)\left(\frac{1}{s_1^2} + \frac{1}{s_2^2} + \cdots \frac{1}{s_{N_d}^2}\right), \qquad \text{(B.4)}\end{aligned}$$

which can now be regarded as a function of the singular values of $\mathbf{G}$. The gradient of J_{BLUE} with respect to $\mathbf{s}$ is given by

$$\frac{\partial J_{\mathrm{BLUE}}}{\partial \mathbf{s}} = \frac{2\sigma_d^2}{\gamma N_d} \cdot \begin{bmatrix} s_1\left(\frac{1}{s_1^2} + \cdots \frac{1}{s_{N_d}^2}\right) - s_1^{-3}\left(s_1^2 + \cdots s_{N_d}^2\right)\\ \vdots \\ s_{N_d}\left(\frac{1}{s_1^2} + \cdots \frac{1}{s_{N_d}^2}\right) - s_{N_d}^{-3}\left(s_1^2 + \cdots s_{N_d}^2\right)\end{bmatrix}. \tag{B.5}$$

Setting the gradient to zero leads to the system of equations

$$\begin{aligned} s_1^4 \left(\frac{1}{s_1^2} + \cdots \frac{1}{s_{N_d}^2} \right) &= s_1^2 + \cdots s_{N_d}^2 \\ &\vdots \\ s_{N_d}^4 \left(\frac{1}{s_1^2} + \cdots \frac{1}{s_{N_d}^2} \right) &= s_1^2 + \cdots s_{N_d}^2. \end{aligned} \tag{B.6}$$

Based on (B.6), a general expression for the kth equation can be derived as

$$s_k^4 \sum_{m=1}^{N_d} \frac{1}{s_m^2} = \sum_{m=1}^{N_d} s_m^2. \tag{B.7}$$

Subtracting the lth from the kth equation

$$s_k^4 \sum_{m=1}^{N_d} \frac{1}{s_m^2} - s_l^4 \sum_{m=1}^{N_d} \frac{1}{s_m^2} = \sum_{m=1}^{N_d} s_m^2 - \sum_{m=1}^{N_d} s_m^2 \tag{B.8}$$

$$\left(s_k^4 - s_l^4\right) \underbrace{\sum_{m=1}^{N_d} \frac{1}{s_m^2}}_{>0} = 0 \tag{B.9}$$

for all $k, l \in \{1, ..., N_d\}$ with $k \neq l$ immediately reveals

$$s_1 = s_2 = \cdots s_{N_d} \tag{B.10}$$

as the only possible solution to this system of equations.

Consequently, every possible candidate $\mathbf{G}$ for a local minimum satisfies $\mathbf{G}^H\mathbf{G} = s^2\mathbf{I}$ (cf. (4.17) and its implications). Inserting $\mathbf{G}^H\mathbf{G} = s^2\mathbf{I}$ into the cost function (4.7) leads to the same expression as in (4.13) that corresponds to the numerically found local minima. Hence, every $\mathbf{G}$ fulfilling $\mathbf{G}^H\mathbf{G} = s^2\mathbf{I}$ results in the same (and minimum) value $J_{\mathrm{BLUE,min}} = \sigma_d^2 N_d/\gamma$ which therefore constitutes the global minimum of the cost function.

C. Derivation of the Global Minimum of J_{LMMSE}

Appendix C proves that $\partial J_{\text{LMMSE}}/\partial \mathbf{s} = \mathbf{0}$, if and only if all singular values of $\mathbf{G}$ are non-zero and identical. Furthermore, it is also shown that every local minimum of the cost function J_{LMMSE} is also a global minimum with $J_{\text{LMMSE,min}} = \sigma_d^2 N_d/(\gamma + 1)$. Starting with the cost function defined in (4.10) as

$$J_{\text{LMMSE}}(\mathbf{A}) = \sigma_d^2 \text{tr}\left(\left(\frac{\gamma N_d}{\text{tr}\left(\mathbf{G}^H\mathbf{G}\right)}\mathbf{G}^H\mathbf{G} + \mathbf{I}\right)^{-1}\right), \tag{C.1}$$

whereas $\mathbf{G} = \mathbf{A}\begin{bmatrix}\mathbf{I} & \mathbf{T}^T\end{bmatrix}^T$ according to (2.42), and inserting (B.2) and (B.3) into (C.1) leads to

$$J_{\text{LMMSE}} = \sigma_d^2 \text{tr}\left(\left(\frac{\gamma N_d}{\text{tr}\left(\mathbf{D}^2\right)}\mathbf{V}\mathbf{D}^2\mathbf{V}^H + \mathbf{I}\right)^{-1}\right). \tag{C.2}$$

Applying the matrix inversion lemma immediately (Sherma-Morrison-Woodbury formula [91]) delivers

$$\begin{aligned} J_{\text{LMMSE}} &= \sigma_d^2 \text{tr}\left(\mathbf{I} - \mathbf{V}\left(\mathbf{V}^H\mathbf{V} + \frac{\text{tr}\left(\mathbf{D}^2\right)}{\gamma N_d}(\mathbf{D}^2)^{-1}\right)^{-1}\mathbf{V}^H\right) \\ &= \sigma_d^2\left(N_d - \text{tr}\left(\left(\mathbf{I} + \frac{\text{tr}\left(\mathbf{D}^2\right)}{\gamma N_d}(\mathbf{D}^2)^{-1}\right)^{-1}\right)\right). \end{aligned} \tag{C.3}$$

The advantage of this notation is that every matrix in (C.3), that needs to be inverted, shows a diagonal structure. Having in mind that $\mathbf{D}^2$ is a diagonal matrix with the squared singular values of $\mathbf{G}$ on its main diagonal such that $\text{tr}\left(\mathbf{D}^2\right) = \sum_{i=1}^{N_d} s_i^2$, and after some rearrangements, the cost function reads as a function of the singular values

$$J_{\text{LMMSE}} = \sigma_d^2 N_d - \sigma_d^2 \gamma N_d \sum_{i=1}^{N_d} \frac{s_i^2}{\gamma N_d s_i^2 + \text{tr}\left(\mathbf{D}^2\right)}. \tag{C.4}$$

The gradient of J_{LMMSE} with respect to $\mathbf{s}$ is obtained by calculating the partial derivatives $\frac{\partial J}{\partial s_k}$ for $k \in \{1, \ldots, N_d\}$. In the following, a derivation for $\frac{\partial J}{\partial s_1}$ is given and subsequently generalized to s_k.

$$\frac{\partial J}{\partial s_1} = -\sigma_d^2 \gamma N_d \left[\frac{2s_1 \left(\gamma N_d^2 s_1^2 + \mathrm{tr}\left(\mathbf{D}^2\right)\right) - s_1^2 \left(2\gamma N_d s_1 + 2s_1\right)}{\left(\gamma N_d s_1^2 + \mathrm{tr}\left(\mathbf{D}^2\right)\right)^2} + \frac{-2s_1 s_2^2}{\left(\gamma N_d s_2^2 + \mathrm{tr}\left(\mathbf{D}^2\right)\right)^2} + \frac{-2s_1 s_3^2}{\left(\gamma N_d s_3^2 + \mathrm{tr}\left(\mathbf{D}^2\right)\right)^2} + \cdots + \frac{-2s_1 {s_N}_d^2}{\left(\gamma N_d {s_N}_d^2 + \mathrm{tr}\left(\mathbf{D}^2\right)\right)^2} \right] \tag{C.5}$$

$$= -2\sigma_d^2 \gamma N_d s_1 \left[\frac{\gamma N_d^2 s_1^2 + \mathrm{tr}\left(\mathbf{D}^2\right) - s_1^2 \left(\gamma N_d + 1\right)}{\left(\gamma N_d s_1^2 + \mathrm{tr}\left(\mathbf{D}^2\right)\right)^2} - \frac{s_2^2}{\left(\gamma N_d s_2^2 + \mathrm{tr}\left(\mathbf{D}^2\right)\right)^2} - \frac{s_3^2}{\left(\gamma N_d s_3^2 + \mathrm{tr}\left(\mathbf{D}^2\right)\right)^2} - \cdots - \frac{{s_N}_d^2}{\left(\gamma N_d {s_N}_d^2 + \mathrm{tr}\left(\mathbf{D}^2\right)\right)^2} \right] \tag{C.6}$$

$$= -2\sigma_d^2 \gamma N_d s_1 \left[\frac{s_2^2 + s_3^2 + \cdots + {s_N}_d^2}{\left(\gamma N_d s_1^2 + \mathrm{tr}\left(\mathbf{D}^2\right)\right)^2} - \frac{s_2^2}{\left(\gamma N_d s_2^2 + \mathrm{tr}\left(\mathbf{D}^2\right)\right)^2} - \frac{s_3^2}{\left(\gamma N_d s_3^2 + \mathrm{tr}\left(\mathbf{D}^2\right)\right)^2} - \cdots - \frac{{s_N}_d^2}{\left(\gamma N_d {s_N}_d^2 + \mathrm{tr}\left(\mathbf{D}^2\right)\right)^2} \right] \tag{C.7}$$

$$\begin{aligned} = -2\sigma_d^2 \gamma N_d s_1 \Bigg[& \frac{s_2^2}{\left(\gamma N_d s_1^2 + \mathrm{tr}\left(\mathbf{D}^2\right)\right)^2} - \frac{s_2^2}{\left(\gamma N_d s_2^2 + \mathrm{tr}\left(\mathbf{D}^2\right)\right)^2} + \frac{s_3^2}{\left(\gamma N_d s_1^2 + \mathrm{tr}\left(\mathbf{D}^2\right)\right)^2} - \frac{s_3^2}{\left(\gamma N_d s_3^2 + \mathrm{tr}\left(\mathbf{D}^2\right)\right)^2} \\ & + \cdots + \frac{{s_N}_d^2}{\left(\gamma N_d s_1^2 + \mathrm{tr}\left(\mathbf{D}^2\right)\right)^2} - \frac{{s_N}_d^2}{\left(\gamma N_d {s_N}_d^2 + \mathrm{tr}\left(\mathbf{D}^2\right)\right)^2} \Bigg] \end{aligned} \tag{C.8}$$

$$= -2\sigma_d^2 \gamma N_d s_1 \sum_{i=2}^{N_d} \left[\frac{s_i^2}{\left(\gamma N_d s_1^2 + \mathrm{tr}\left(\mathbf{D}^2\right)\right)^2} - \frac{s_i^2}{\left(\gamma N_d s_i^2 + \mathrm{tr}\left(\mathbf{D}^2\right)\right)^2} \right]. \tag{C.9}$$

A generalized expression for the partial derivative consequently follows as

$$\frac{\partial J}{\partial s_k} = -2\sigma_d^2 \gamma N_d s_k \sum_{m=1, m\neq k}^{N_d} \left[\frac{s_m^2}{\left(\gamma N_d s_k^2 + \mathrm{tr}\left(\mathbf{D}^2\right)\right)^2} - \frac{s_m^2}{\left(\gamma N_d s_m^2 + \mathrm{tr}\left(\mathbf{D}^2\right)\right)^2} \right] \tag{C.10}$$

$$= -2\sigma_d^2 \gamma N_d s_k \sum_{m=1, m\neq k}^{N_d} \left[\frac{s_m^2}{\mathrm{f}\left(s_k^2\right)} - \frac{s_m^2}{\mathrm{f}\left(s_m^2\right)} \right], \tag{C.11}$$

whereas the last step with $\mathrm{f}\left(s_k^2\right) := \left(\gamma N_d s_k^2 + \mathrm{tr}\left(\mathbf{D}^2\right)\right)^2$ has been introduced for reasons of compact notation[1]. Setting the gradient to zero and assuming non-zero singular values yields for a single partial derivative

$$0 = \frac{\partial J}{\partial s_k} \quad \Leftrightarrow \quad 0 = \sum_{m=1, m\neq k}^{N_d} \left[\frac{s_m^2}{\mathrm{f}\left(s_k^2\right)} - \frac{s_m^2}{\mathrm{f}\left(s_m^2\right)} \right] \tag{C.12}$$

Carrying out (C.12) for $k \in \{1 \dots N_d\}$ leads to a system of N_d equations similar to the BLUE in appendix B. Subtracting the lth from the kth equation delivers

$$0 = \frac{\partial J}{\partial s_k} - \frac{\partial J}{\partial s_l} \tag{C.13}$$

$$0 = \sum_{m=1, m\neq k}^{N_d} \left[\frac{s_m^2}{\mathrm{f}\left(s_k^2\right)} - \frac{s_m^2}{\mathrm{f}\left(s_m^2\right)} \right] - \sum_{m=1, m\neq l}^{N_d} \left[\frac{s_m^2}{\mathrm{f}\left(s_l^2\right)} - \frac{s_m^2}{\mathrm{f}\left(s_m^2\right)} \right] \tag{C.14}$$

$$0 = \sum_{m=1, m\neq k}^{N_d} \frac{s_m^2}{\mathrm{f}\left(s_k^2\right)} - \sum_{m=1, m\neq l}^{N_d} \frac{s_m^2}{\mathrm{f}\left(s_l^2\right)} + \sum_{m=1, m\neq l}^{N_d} \frac{s_m^2}{\mathrm{f}\left(s_m^2\right)} - \sum_{m=1, m\neq k}^{N_d} \frac{s_m^2}{\mathrm{f}\left(s_m^2\right)} \tag{C.15}$$

$$0 = \sum_{m=1, m\neq k}^{N_d} \frac{s_m^2}{\mathrm{f}\left(s_k^2\right)} - \sum_{m=1, m\neq l}^{N_d} \frac{s_m^2}{\mathrm{f}\left(s_l^2\right)} - \frac{s_l^2}{\mathrm{f}\left(s_l^2\right)} + \frac{s_k^2}{\mathrm{f}\left(s_k^2\right)} \tag{C.16}$$

$$0 = \sum_{m=1}^{N_d} \frac{s_m^2}{\mathrm{f}\left(s_k^2\right)} - \sum_{m=1}^{N_d} \frac{s_m^2}{\mathrm{f}\left(s_l^2\right)} \tag{C.17}$$

$$0 = \frac{\mathrm{f}\left(s_l^2\right) \sum_{m=1}^{N_d} s_m^2 - \mathrm{f}\left(s_k^2\right) \sum_{m=1}^{N_d} s_m^2}{\mathrm{f}\left(s_k^2\right) \mathrm{f}\left(s_l^2\right)}. \tag{C.18}$$

[1] Note that $\mathrm{f}\left(s_k^2\right)$ will depend on all singular values due to $\mathbf{D}$ and not only on s_k. However, for the purpose of compact notation, this is neglected here.

Since the denominator is non-zero, (C.18) can be reformulated as

$$0 = \mathrm{f}\left(s_l^2\right) \sum_{m=1}^{N_d} s_m^2 - \mathrm{f}\left(s_k^2\right) \sum_{m=1}^{N_d} s_m^2 \tag{C.19}$$

$$0 = \underbrace{\sum_{m=1}^{N_d} s_m^2}_{>0} \left(\mathrm{f}\left(s_l^2\right) - \mathrm{f}\left(s_k^2\right)\right) \quad \Leftrightarrow \tag{C.20}$$

$$0 = \left(\mathrm{f}\left(s_l^2\right) - \mathrm{f}\left(s_k^2\right)\right). \tag{C.21}$$

Substituting back $\mathrm{f}\left(s_l^2\right)$ and $\mathrm{f}\left(s_k^2\right)$ yields

$$\begin{aligned} 0 =& \left(\gamma^2 N_d^2 s_l^4 + 2\gamma N_d s_l^2 \mathrm{tr}\left(\mathbf{D}\right) + \left(\mathrm{tr}\left(\mathbf{D}\right)\right)^2\right) \\ & - \left(\gamma^2 N_d^2 s_k^4 + 2\gamma N_d s_k^2 \mathrm{tr}\left(\mathbf{D}\right) + \left(\mathrm{tr}\left(\mathbf{D}\right)\right)^2\right) \end{aligned} \tag{C.22}$$

$$0 = \gamma^2 N_d^2 \left(s_l^4 - s_k^4\right) + 2\gamma N_d \mathrm{tr}\left(\mathbf{D}\right) \left(s_l^2 - s_k^2\right) \tag{C.23}$$

$$0 = \left(s_l^2 - s_k^2\right) \left(\underbrace{\gamma^2 N_d^2 \left(s_l^2 + s_k^2\right)}_{>0} + \underbrace{2\gamma N_d \mathrm{tr}\left(\mathbf{D}\right)}_{>0} \right), \tag{C.24}$$

which is only fulfilled, if and only if $s_l = s_k$. Finally, carrying out

$$0 = \frac{\partial J}{\partial s_k} - \frac{\partial J}{\partial s_l} \quad \text{for} \quad k, l \in \{1, \ldots, N_d\} \text{ with } k \neq l \tag{C.25}$$

immediately proves

$$s_1 = \cdots = s_{N_d} = s \tag{C.26}$$

as the only solution to $\partial J_{\mathrm{LMMSE}} / \partial \mathbf{s} = \mathbf{0}$. The remaining argumentation coincides with the one for the BLUE in appendix B. However, the expression for the global minimum $J_{\mathrm{LMMSE,min}} = \sigma_d^2 N_d / (\gamma + 1)$ obtained by inserting $\mathbf{G}^H \mathbf{G} = s^2 \mathbf{I}$ into (C.1) differs from $J_{\mathrm{BLUE,min}}$.

List of Abbreviations

ASK	amplitude shift keying
BER	bit error ratio
BLUE	best linear unbiased estimator
BMSE	Bayesian mean square error
CDF	cumulative distribution function
CFO	carrier frequency offset
CI	channel inversion
CIR	channel impulse response
CLT	central limit theorem
CP	cyclic prefix
DFE	decision feedback equalization
DFT	discrete Fourier transform
FDE	frequency domain equalization
ICI	inter-carrier interference
IDFT	inverse discrete Fourier transform
iid	independent and identically distributed
ISI	intersymbol interference
LLR	log-likelihood ratio
LMMSE	linear minimum mean square error
MSE	mean square error
MVU	minimum variance unbiased
OFDM	orthogonal frequency division multiplexing
PAPR	peak to average power ratio
PDF	probability density function
PMF	probability mass function
PMR	peak to minimum power ratio
PSD	power spectral density
PSK	phase shift keying

QAM	quadrature amplitude modulation
RS	Reed Solomon
RV	random variable
SC	single carrier
SIC	successive interference cancelation
SNR	signal-to-noise ratio
SVD	singular value decomposition
UW	unique word
ZF	zero forcing

Bibliography

[1] Richard van Nee and Ramjee Prasad. *OFDM for Wireless Multimedia Communications*. Artech House, Inc., Norwood, MA, USA, 1st edition, 2000.

[2] Robert W. Chang and Richard A. Gibby. A Theoretical Study of Performance of an Orthogonal Multiplexing Data Transmission Scheme. *IEEE Transactions on Communication Technology*, 16(4):529–540, 1968.

[3] Stephen B. Weinstein and Paul M. Ebert. Data Transmission by Frequency-Division Multiplexing Using the Discrete Fourier Transform. *IEEE Transactions on Communication Technology*, 19(5):628–634, 1971.

[4] Abraham Peled and Antonio Ruiz. Frequency Domain Data Transmission using Reduced Computational Complexity Algorithms. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 964–967, Denver, CO, USA, 1980.

[5] Botaro Hirosaki. An Orthogonally Multiplexed QAM System Using the Discrete Fourier Transform. *IEEE Transactions on Communications*, 29(7):982–989, 1981.

[6] John A. C. Bingham. Multicarrier Modulation for Data Transmission: An Idea Whose Time Has Come. *IEEE Communications Magazine*, 28(5):5–14, May 1990.

[7] European Telecommunications Standards Institute (ETSI). Radio Broadcasting Systems; Digital Audio Broadcasting (DAB) to Mobile, Portable, and Fixed Receivers. European Telecommunication Standard ETS 300 401, 1995. 1st edition, reference DE/JTC-DAB.

[8] European Telecommunications Standards Institute (ETSI). Digital Video Broadcasting (DVB); Framing Structure, Channel Coding and Modulation for Digital Terrestrial Television. European Telecommunication Standard ETS 300 744, 1997. 1st edition, reference DE/JTC-DVB-8.

[9] Stefano Galli and Oleg Logvinov. Recent Developments in the Standardization of Power Line Communications within the IEEE. *IEEE Communications Magazine*, 46(7):64–71, 2008.

[10] Network and Customer Installation Interfaces — Asymmetric Digital Subscriber Line (ADSL) Metallic Interface. ANSI standard T1.413-1995, 1995.

[11] IEEE Standard for Information Technology–Telecommunications and Information Exchange between Systems; Local and Metropolitan Area Networks–Specific Requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, 2012.

[12] Erik Dahlman, Stefan Parkvall, and Johan Skold. *4G: LTE/LTE-Advanced for Mobile Broadband*. Elsevier Science, 1st edition, 2011.

[13] Bertrand Muquet, Zhengdao Wang, Georgios B. Giannakis, Marc de Courville, and Pierre Duhamel. Cyclic Prefixing or Zero Padding for Wireless Multicarrier Transmissions? *IEEE Transactions on Communications*, 50(12):2136–2148, December 2002.

[14] Framing Structure, Channel Coding and Modulation for Digital Television Terrestrial Broadcasting System. Chinese National Standard, Std. GB 20 600-2006, 2006.

[15] Chung yen Ong, Jian Song, Changyong Pan, and Yangang Li. Technology and Standards of Digital Television Terrestrial Multimedia Broadcasting [Topics in Wireless Communications]. *IEEE Communications Magazine*, 48(5):119–127, May 2010.

[16] Shigang Tang, Fang Yang, Kewu Peng, Changyong Pan, Ke Gong, and Zhixing Yang. Iterative Channel Estimation for Block Transmission with Known Symbol Padding — A New Look at TDS-OFDM. In *Proceedings of the IEEE Global Communications Conference (GLOBECOM)*, pages 4269–4273, Washington, DC, USA, 2007.

[17] Markus Muck, Marc de Courville, and Pierre Duhamel. A Pseudorandom Postfix OFDM Modulator—Semi-Blind Channel Estimation and Equalization. *IEEE Transactions on Signal Processing*, 54(3):1005–1017, 2006.

[18] Liao Jingyi, Wang Hai, Panyuh Joo, and Jungmin Ro. The Effect of Filling Unique Words to Guard Interval for OFDM System. C802.16a-02/87, IEEE 802.16 Broadband Wireless Access Working Group, 2002.

[19] Richard E. Blahut. *Algebraic Codes for Data Transmission*. Cambridge University Press, 1st edition, 2003.

[20] Werner Henkel, Fangning Hu, and Ina Kodrasi. Inherent Time-Frequency Coding in OFDM and ISI Correction without a Cyclic Prefix. In *Proceedings of the 14th International OFDM Workshop*, Hamburg, Germany, 2009.

[21] Christian Hofbauer, Mario Huemer, and Johannes B. Huber. Coded OFDM by Unique Word Prefix. In *Proceedings of the IEEE International Conference on Communication Systems (ICCS)*, pages 426–430, Singapore, 2010.

[22] Harald Witschnig, Thomas Mayer, Andreas Springer, Alois Koppler, Linus Maurer, Mario Huemer, and Robert Weigel. A Different Look on Cyclic Prefix for SC/FDE. In *Proceedings of the 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pages 824–828, Lisbon, Portugal, 2002.

[23] Harald Witschnig, Thomas Mayer, Andreas Springer, Linus Maurer, Mario Huemer, and Robert Weigel. The Advantages of a Known Sequence versus Cyclic Prefix in a SC/FDE System. In *Proceedings of the 5th International Symposium on Wireless Personal Multimedia Communications (WPMC)*, volume 3, pages 1328–1332, Honolulu, HI, USA, 2002.

[24] Mario Huemer, Harald Witschnig, and Josef Hausner. Unique Word Based Phase Tracking Algorithms for SC/FDE-Systems. In *Proceedings of the IEEE Global Communications Conference (GLOBECOM)*, pages 70–74, San Francisco, CA, USA, 2003.

[25] Harald Witschnig. *Frequency Domain Equalization for Broadband Wireless Communication — With Special Reference to Single Carrier Transmission Based on Known Pilot Sequences*. Dissertation, Institute for Communications and Information Engineering, University of Linz, 2004.

[26] Luc Deneire, Bert Gyselinckx, and Marc Engels. Training Sequence versus Cyclic Prefix—A New Look on Single Carrier Communication. *IEEE Communications Letters*, 5(7):292–294, 2001.

[27] Mario Huemer. *Frequenzbereichsentzerrung für hochratige Einträger-Übertragungssysteme in Umgebungen mit ausgeprägter Mehrwegeausbreitung*. Dissertation, Institute for Communications and Information Engineering, University of Linz, 1999.

[28] Mario Huemer, Alois Koppler, Robert Weigel, and Leonhard Reindl. A Review of Cyclically Extended Single Carrier Transmission with Frequency Domain Equalization for Broadband Wireless Transmission. *European Transactions on Telecommunications*, 14(4):329–341, 2003.

[29] Mario Huemer, Christian Hofbauer, and Johannes B. Huber. Unique Word Prefix in SC/FDE and OFDM: A Comparison. In *Proceedings of the IEEE Global Communications Conference (GLOBECOM) Workshops*, pages 1296–1301, Miami, FL, USA, 2010.

[30] Christian Hofbauer and Mario Huemer. A Study of Data Rate Equivalent UW-OFDM and CP-OFDM Concepts. In *Proceedings of the 46th Asilomar Conference on Signals, Systems and Computers*, pages 173–177, Pacific Grove, CA, USA, 2012.

[31] Mario Huemer, Christian Hofbauer, and Johannes B. Huber. Non-Systematic Complex Number RS Coded OFDM by Unique Word Prefix. *IEEE Transactions on Signal Processing*, 60(1):285–299, 2012.

[32] Raphael Cendrillon and Marc Moonen. Efficient Equalizers for Single and Multi-Carrier Environments with Known Symbol Padding. In *Proceedings of the 6th International Symposium on Signal Processing and its Applications*, volume 2, pages 607–610, Kuala Lumpur, Malaysia, 2001.

[33] Weimin Zhang and Michael J. Miller. Baseband Equivalents in Digital Communication System Simulation. *IEEE Transactions on Education*, 35(4):376–382, November 1992.

[34] Bernd Girod, Rudolf Rabenstein, and Alexander Stenger. *Signals & Systems*. John Wiley & Sons, 1st edition, 2001.

[35] Johannes B. Huber, Mario Huemer, and Christian Hofbauer. scientific discussions. Klagenfurt, June 2008.

[36] IEEE Std 802.11a-1999, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: High-Speed Physical Layer in the 5 GHz Band, 1999.

[37] Steven Kay. *Intuitive Probability and Random Processes using MATLAB*. Springer US, 1st edition, 2006.

[38] Mario Huemer, Christian Hofbauer, Alexander Onic, and Johannes B. Huber. On the Exploitation of the Redundant Energy in UW-OFDM: LMMSE Versus Sphere Detection. *IEEE Signal Processing Letters*, 19(6):340–343, 2012.

[39] Alexander Onic and Mario Huemer. Noise Interpolation for Unique Word OFDM. *Signal Processing Letters, IEEE*, 21(7):814–818, July 2014.

[40] Alexander Onic. *Receiver Concepts for Unique Word OFDM*. PhD thesis, Institute of Networked and Embedded Systems, Alpen-Adria-Universität Klagenfurt, 2013.

[41] Mario Huemer, Alexander Onic, and Christian Hofbauer. Classical and Bayesian Linear Data Estimators for Unique Word OFDM. *IEEE Transactions on Signal Processing*, 59(12):6073–6085, 2011.

[42] Steven M. Kay. *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. Prentice Hall, 1st edition, 1993.

[43] John Proakis. *Digital Communications*. McGraw-Hill Science/Engineering/Math, 4th edition, 2000.

[44] Werner Haselmayr, Christian Hofbauer, Bernhard Etzlinger, Andreas Springer, and Mario Huemer. Iterative Detection for Unique Word OFDM. In *Proceedings of the Global Communications Conference (GLOBECOM)*, pages 3261–3266, Austin, TX, USA, December 2014.

[45] Carla Savage. A Survey of Combinatorial Gray Codes. *SIAM Review*, 39(4):605–629, 1997.

[46] John Fakatselis. *Criteria for 2.4 GHz PHY Comparison of Modulation*. IEEE Document, 1997. P802.11-97/157r1.

[47] Andrew J. Viterbi. Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm. *IEEE Transactions on Information Theory*, 13(2):260–269, April 1967.

[48] Theodore S. Rappaport. *Wireless Communications: Principles and Practice*. Prentice Hall, Upper Saddle River, NJ, USA, 2nd edition, 2001.

[49] Peyton Peebles. *Probability, Random Variables, and Random Signal Principles*. McGraw-Hill Science, 4th edition, 2001.

[50] Khalid Hassan, Tharek Abd Rahman, Ramlee Kamarudin, and Faizah Nor. The Mathematical Relationship Between Maximum Access Delay and the R.M.S. Delay Spread. In *Proceedings of the 7th International Conference on Wireless and Mobile Communications*, pages 18–23, Luxembourg, June 2011.

[51] Heidi Steendam. On the Redundant Carrier Distribution for UW-OFDM. In *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)*, pages 32–36, Paris, France, 2012.

[52] Yuan-Pei Lin, See-May Phoong, and P. P. Vaidyanathan. *Filter Bank Transceivers for OFDM and DMT Systems*. Cambridge University Press, 1st edition, 2010.

[53] Heidi Steendam. Analysis of the Redundant Energy in UW-OFDM. *IEEE Transactions on Communications*, 60(6):1692–1701, 2012.

[54] Peter J. Schreier and Louis L. Scharf. *Statistical Signal Processing of Complex-Valued Data*. Cambridge University Press, 1st edition, 2010. Cambridge Books Online.

[55] Hisashi Kobayashi, Brian L. Mark, and William Turin. *Probability, Random Processes, and Statsitical Analysis*. Cambridge University Press, 1st edition, 2012.

[56] Johannes B. Huber, Jakob Rettelbach, Mathis Seidl, and Mario Huemer. Signal Shaping for Unique-Word OFDM by Selected Mapping. In *Proceedings of the 18th European Wireless Conference*, page 8, Poznan, Poland, April 2012.

[57] Mohamed Akkouchi. On the Convolution of Exponential Distributions. *Journal of the Chungcheong Mathematical Society*, 21(4):501–510, 2008.

[58] Mario Huemer, Christian Hofbauer, Alexander Onic, and Johannes B. Huber. Design and Analysis of UW-OFDM Signals. *AEU - International Journal of Electronics and Communications*, 68(10):958–968, 2014.

[59] Mario Huemer, Christian Hofbauer, and Johannes B. Huber. The Potential of Unique Words in OFDM. In *Proceedings of the 15th International OFDM Workshop*, pages 140–144, Hamburg, 2010.

[60] Alexander Onic and Mario Huemer. Direct vs. Two-Step Approach for Unique Word Generation in UW-OFDM. In *Proceedings of the 15th International OFDM Workshop*, pages 145–149, Hamburg, 2010.

[61] Christian Hofbauer, Mario Huemer, and Johannes B. Huber. On the Impact of Redundant Subcarrier Energy Optimization in UW-OFDM. In *Proceedings of the 4th International Conference on Signal Processing and Communication Systems (ICSPCS)*, page 6, Gold Coast, Australia, 2010.

[62] Richard Bronson. *Schaum's Outline of Theory and Problems of Matrix Operations*. Schaum's Outline Series. McGraw-Hill, 1st edition, 1986.

[63] Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, 3rd edition, 1996.

[64] Solomon W. Golomb and Robert A. Scholtz. Generalized Barker Sequences. *IEEE Transactions on Information Theory*, 11(4):533–537, 1965.

[65] Branislav M. Popovič. Generalized Chirp-Like Polyphase Sequences with Optimum Correlation Properties. *IEEE Transactions on Information Theory*, 38(4):1406–1409, 1992.

[66] IEEE Std 802.16-2004, Part 16: Air Interface for Fixed Broadband Wireless Access Systems, 2004.

[67] Heidi Steendam. The Quasi-Uniform Redundant Carrier Placement for UW-OFDM. In *Proceedings of the IEEE Vehicular Technology Conference (VTC Fall)*, Quebec City, Canada, 2012.

[68] Heidi Steendam. On the Selection of the Redundant Carrier Positions in UW-OFDM. *IEEE Transactions on Signal Processing*, 61(5):1112–1120, 2013.

[69] Gerard J. Foschini. Layered Space-Time Architecture for Wireless Communication in a Fading Environment When Using Multi-Element Antennas. *Bell Labs Technical Journal*, 1(2):41–59, 1996.

[70] Peter W. Wolniansky, Gerard J. Foschini, Glen D. Golden, and Reinaldo A. Valenzuela. V-BLAST: An Architecture for Realizing Very High Data Rates Over the Rich-Scattering Wireless Channel. In *Proceedings of the URSI International Symposium on Signals, Systems, and Electronics, (ISSSE)*, pages 295–300, Pisa, Italy, 1998.

[71] Andrey N. Tikhonov and Vasiliy Y. Arsenin. *Solutions of Ill-Posed Problems*. Winston & Sons, Washingthon, D.C., USA, 1977.

[72] Morteza Rajabzadeh, Heidi Steendam, and Hossein Khoshbin. Power Spectrum Characterization of Systematic Coded UW-OFDM Systems. In *Proceedings of the 78th IEEE Vehicular Technology Conference (VTC Fall)*, page 5, Las Vegas, NV, USA, September 2013.

[73] Mario Huemer, Christian Hofbauer, and Johannes B. Huber. Complex Number RS Coded OFDM with Systematic Noise in the Guard Interval. In *Proceedings of the 44th Asilomar Conference on Signals, Systems and Computers*, pages 1023–1028, Pacific Grove, CA, USA, November 2010.

[74] Heidi Steendam. On the Diversity Order of UW-OFDM (invited paper). In *Proceedings of the IFIP TC7 2013 Conference on System Modeling and Optimization*, Klagenfurt, Austria, September 2013.

[75] Heidi Steendam. Design and Analysis of the UW-OFDM Signal. In *Proceedings of the 6th International Symposium on Communications, Control and Signal Processing (ISCCSP)*, pages 470–474, Athens, Greece, May 2014.

[76] Heidi Steendam. Theoretical Performance Evaluation and Optimization of UW-OFDM. *IEEE Transactions on Communications*, 64(4):1739–1750, April 2016.

[77] Morteza Rajabzadeh, Hossein Khoshbin, and Heidi Steendam. Sidelobe Suppression for Non-Systematic Coded UW-OFDM in Cognitive Radio Networks. In *Proceedings of the 20th European Wireless Conference*, pages 826–831, Barcelona, Spain, May 2014.

[78] Mario Huemer, Alexander Onic, Christian Hofbauer, and Stefan Trampitsch. Widely Linear Data Estimation for Unique Word OFDM. In *Proceedings of the 47th Asilomar Conference on Signals, Systems and Computers*, pages 1934–1938, Pacific Grove, CA, USA, November 2013.

[79] Jacob Wolfowitz. Memory Increases Capacity. *Information and Control*, 11(4):423–428, 1967.

[80] Jakob Rettelbach and Johannes B. Huber. PMR-Reduction for Continuous Time OFDM Transmit Signals by Selected Mapping. In *Proceedings of the International Symposium on Signals, Systems and Electronics (ISSSE)*, Potsdam, Germany, 2012.

[81] Xiaodong Cai and Georgios B. Giannakis. Error Probability Minimizing Pilots for OFDM with M-PSK Modulation over Rayleigh-Fading Channels. *IEEE Transactions on Vehicular Technology*, 53(1):146–155, January 2004.

[82] Ferdinand Classen and Heinrich Meyr. Frequency Synchronization Algorithms for OFDM Systems Suitable for Communication over Frequency Selective Fading Channels. In *Proceedings of the 44th Vehicular Technology Conference (VTC Spring)*, pages 1655–1659, Stockholm, Sweden, June 1994.

[83] Ram Sunil Kanumalli. Investigation of Carrier Frequency Offset Estimation Techniques for Unique Word (UW) - OFDM Systems. Master's thesis, Institute of Networked and Embedded Systems, Alpen-Adria-Universität Klagenfurt, Austria, 2012.

[84] Kai Shi, Erchin Serpedin, and Philippe Ciblat. Decision-Directed Fine Synchronization in OFDM Systems. *IEEE Transactions on Communications*, 53(3):408–412, March 2005.

[85] Yuh-Ren Tsai, Xiu-Sheng Li, and Chih-Yu Wei. Data-Carrier Aided Carrier Frequency Offset Estimation for OFDM Systems. In *Proceedings of the 67th Vehicular Technology Conference (VTC Spring)*, pages 898–902, May 2008.

[86] Thomas Keller, Lorenzo Piazzo, Paolo Mandarini, and Lajos Hanzo. Orthogonal Frequency Division Multiplex Synchronization Techniques for Frequency-Selective Fading Channels. *IEEE Journal on Selected Areas in Communications*, 19(6):999–1008, June 2001.

[87] Navid Lashkarian and Sayfe Kiaei. Class of Cyclic-Based Estimators for Frequency-Offset Estimation of OFDM Systems. *IEEE Transactions on Communications*, 48(12):2139–2149, December 2000.

[88] Johannes B. Huber. Grundlagen der Wahrscheinlichkeitsrechnung für iterative Decodierverfahren. *Elektrotechnik und Informationstechnik*, 119(11):386ff, 2002.

[89] Mario Huemer and Oliver Lang. On Component-Wise Conditionally Unbiased Linear Bayesian Estimation. In *Proceedings of the 48th Asilomar Conference on Signals, Systems and Computers*, pages 879–885, Pacific Grove, CA, USA, November 2014.

[90] Oliver Lang and Mario Huemer. On the Log-Likelihood Ratio evaluation of CWCU Linear and Widely Linear MMSE Data Estimators. In *submitted to IEEE Workshop on Statistical Signal Processing (SSP)*, February 2016.

[91] Nicholas Higham. *Accuracy and Stability of Numerical Algorithms*. SIAM, 2nd edition, 2002.