A theoretical framework for the peak-to-average power control problem in OFDM transmission

vorgelegt von
Diplom-Ingenieur
Gerhard Wunder

von der Fakultät IV- Elektrotechnik und Informatik
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften
- Dr.-Ing.-

genehmigte Dissertation

Berichter: Prof. Dr. Dr. Holger Boche
Berichter: Prof. Dr. techn. Helmut Bölcskei (ETH Zürich)

Tag der wissenschaftlichen Aussprache: 04. September 2003

Berlin 2004
D83

Contents

Zusammenfassung

Diese Promotionsarbeit beschäftigt sich mit dem Spitzenwertproblem in orthogonal frequency division multiplexing (OFDM-) Mehrträgersystemen. Das Mehrträgerverfahren OFDM hat sich als eine effiziente Übertragungstechnik in frequenz-selektiven, drahtlosen Schwundkanälen herausgestellt und sich auch bereits in mehreren technischen Standards etabliert. Ein gravierender Nachteil des Verfahrens ist allerdings die nicht konstante Hüllkurve des Sendesignals, die zu einer erhöhten Empfindlichkeit gegenüber nichtlinearen Systemkomponenten im Übertragungspfad führt und insbesondere die Leistungseffizienz des Systems nachhaltig beeinträchtigt. Eine grundlegende Lösung des Problems konnte bisher nicht angegeben werden.

Das Ziel dieser Arbeit ist es, einen theoretischen Rahmen für das Spitzenwertproblem zu liefern. In diesem Zusammenhang werden deterministische und statistische Eigenschaften für das zeitkontinuierliche und zeitdiskrete Sendesignal angegeben. Insbesondere die Beziehung zwischen zeitkontinuierlichem und zeitdiskretem Sendesignal wird dabei grundlegend untersucht. Weiterhin wird der Einfluss von Nichtlinearitäten auf die Performanz analysiert und ein allgemeines Konstruktionsprinzip für Mehrträgersignale mit geringer Dynamik abgeleitet. Mit den Ergebnissen dieser Arbeit ist ein effizienteres Systemdesign möglich.

Abstract

This thesis is devoted to the power control problem in orthogonal frequency division multiplexing (OFDM) multicarrier transmission. Multicarrier transmission has proved to be an efficient transmission technique in frequency-selective channels and has already been adopted to several standards. A major obstacle, however, is its highly non-constant signal envelope making OFDM very sensitive to non-linear components in the transmission path and leading to a severe power efficiency penalty. The problem still defies a fundamental solution.

The goal of this thesis is to provide a mathematical framework for the power control problem. In particular, deterministic and probabilistic characteristics of the continuous-time and discrete-time signals are investigated. The relationship between the continuous-time and discrete-time signals is thoroughly analyzed. The impact on performance is investigated and a general design principle for OFDM signals with low dynamics given.

Acknowledgement

This thesis was written at the Heinrich-Hertz-Institut in Berlin under the supervision of Prof. Boche to which I am most indebted. Prof. Boche was a constant source of ideas and, most importantly, motivation.

In addition, I would like to thank my colleagues S. Stanczak, P. Jung and the whole department for discussions and providing a calm and supportive atmosphere during my work.

I am also grateful to my parents and also my girl friend Stephie for distracting me away from the world of communication engineering (the e-world, as she termed it) back to the real world and its more convenient aspects which was surely not always a simple task.

Chapter 1

Introduction

The appearance of wireless communication networks has significantly changed in the past two decades. Due to the evolution of the Internet, demand for information access as well as personal communication, new services such as short text and advertising messages, email, Internet access and even multimedia applications etc. rather than plain voice communication have soared. The deployment of the emerging universal mobile telecommunications standard (UMTS) in Europe will further sustain and corroborate this development. Future wireless networks must operate under the paradigm that information access can take place at any time and at any place making a form of ubiquitious communication possible.

In order to meet those demands efficient wireless communication networks have to be set up that are capable of supporting high data rates under challenging conditions. The physical wireless link constitutes the bottleneck in this scenario due to multipath fading, mobility and the limited availability of bandwidth. Recently, there have been significant research activities to overcome its limitations by exploring new air-interfaces employing adaptive modulation, multiple antennas and advanced medium access control (MAC) techniques.

In the physical layer several modulation concepts compete for new solutions in next generation systems (or the evolution of existing systems): single-carrier systems like wideband code division multiple access (W-CDMA) or multicarrier systems such as orthogonal frequency division multiplexing (OFDM) or multicarrier (MC-) CDMA. W-CDMA has been adopted by UMTS supporting roughly up to 2 [Mbit/s] at high mobility and large range. Multicarrier techniques are already effectively being used in broadcast systems like digital audio/video broadcasting (DAB/DVB) or wireless local area network (WLAN) standards such as the IEEE 802.11a or the ETSI HIPERLAN 2. Developed by Weinstein and Ebert in the early seventies [1] OFDM has gained great popularity in high data rate (broadband) communications since it can be efficiently realized and channel equalization reduces to a simple equalization of flat (narrowband) subchannels. Another advantage is that OFDM is bandwidth-efficient since the signal spectrum is almost rectangular for a large number of subcarriers. Furthermore, channel adaption can be realized in discrete multitone systems (DMT) using bitloading strategies. The potential of this transmission technique including multiple antennas has been shown recently in [2].

On the other hand, several disadvantages arise with this concept, the most severe of which is the highly non-constant envelope of the transmit signal, making OFDM very sensitive to non-linear components in the transmission path. A key component is the high power amplifier (HPA). Due to cost, design and most importantly power efficiency considerations the HPA cannot resolve the dynamics of the transmit signal and inevitably cuts off the signal at some point causing additional in-band distortion and adjacent channel interference. The power efficiency penalty is certainly the major obstacle to implementing OFDM in low-cost applications. Moreover, in power-limited regimes determined by regulatory bodies the average power is reduced compared to single-carrier systems reducing in turn the range of transmission. The power control problem motivates further research since it touches on many of the advantages that originally made multicarrier transmission popular, i.e. spectral efficiency and implementation issues.

The goal of this thesis is to analyze the peak power problem. We concentrate on a single-user point-to-point communication link and investigate the impact of the non-constant signal envelope on system performance. We also present solutions to alleviate the problem. We would like to emphasize that there are several other disadvantages such as synchronization effects, frequency offsets and channel estimation etc. that we do not comment on here. Also the multiuser or broadcast case, and multiple antennas will not be considered in this thesis.

The work is divided into three main parts. The first part examines the properties of OFDM signals providing considerable insight into the power control problem and the consequences for practical reduction techniques. The second part examines statistical properties and illuminates the effect of clipping on the performance as well as fundamental limits for coding schemes leading to interesting design consequences. The third part introduces new coding strategies to alleviate the problem.

1.1 Related work and general outline

In **Chapter 3** we derive the signal-theoretic properties of the OFDM signal so as to obtain estimates of the peak value of the continuous-time signal from the discrete-time signal. These estimates are of fundamental importance for the OFDM system design. In mathematics such estimates enjoy a long history (Lebesgue constants, Marcinkiewicz type inequalities). For trigonometric polynomials the problem was investigated by the mathematicians Riesz and Bernstein [3, 4]. Part of this work was rediscovered by Ehlich [5]. Jetter et al generalized the result in [6]. In the context of OFDM, as far as we are aware the problem was introduced by Tellado in his thesis [7]. Further contributors were Tarokh and Paterson [8, 9], Tellambura [10, 11] and Wulich [12]. Here we generalize the results and derive the fundamental estimates for band-limited signals. From a practical viewpoint this gives the best possible estimates. In addition, we put the problem in a rather general framework and also derive estimates for the disturbed data [13, 14, 15]. We refer to this as the noise-enhancement problem and improve on results by [16, 17]. Additionally, we highlight the importance of this investigation for the design of appropriate coding schemes. The results of this chapter have been published so far in [18, 19, 20, 21, 22].

In **Chapter 4** we investigate the statistical properties of the OFDM signal and derive upper bounds on the complementary distribution of the peak power. The distribution is a key parameter effecting all relevant performance measures. In mathematical literature the OFDM signals are referred to as random polynomials. The work on random polynomials goes back to Salem and Zygmund [23], Halasz [24], and Gersho et al [25]. Halasz is possibly awarded the most significant contribution. In the context of OFDM the distribution was introduced by Mestagh et al [26]. Müller et al [27] Friese, Ochiai et al, and Dinur et al developed it further in [28, 29, 30]. We derive the first real upper bounds for practical modulation schemes used in OFDM. In addition, we derive an equivalent result for the coded schemes. Futhermore, we present asymptotic results improving on the results in [25]. We use these results for bounding the symbol error rate evoked by clipping, thereby improving on a recent approach by Bahai [31]. The results of this chapter have been published so far in [32, 33, 34, 35, 36, 37].

In **Chapter 5** we introduce the coding approach to alleviate the power control problem. Using coding across the subcarriers is an attractive idea since codes are a natural sequence selector and OFDM systems are coded to increase diversity. Thus, early attempts were made to solve the peak power problem elegantly en passant with the appropriate codes. Pioneering work was done by [38, 39, 40] where codes were found by computer searches for small numbers of subcarriers. Ochiai et al [41, 42] and van Nee [43] introduced complementary sequences for the code design. A systematic approach to the design problem was given by Davis and Jedwab capitalizing on properties of certain cosets of classical Reed-Muller codes and their generalizations [44, 45]. This work was generalized by Paterson in [46]. In [9] a geometric approach was used to derive fundamental limits on the triplet rate, peak power and minimum euclidean distance for coding schemes. We also pursued a more systematic approach leading to a constructive way to codes with low peak power. In particular, we show how to characterize and design constellations with low peak power. We call the outcome generalized constellations and analyze their overall performance. Furthermore, we explain how codes can be constructed so that the peak power is uniformly bounded thereby solving the problem of [9] and generalize these codes to set up a new system concept. The results of this chapter have been published so far in [47, 48].

1.2 Notations

Sets and events are given by calligraphic letters. The cardinality of a set is denoted as $|\cdot|$ (this is not be confused with the magnitude of complex number) and the set function by I. The probability of a set is denoted as $\Pr(\cdot)$, $\sim$ means ,,is distributed" and the expectation operator is given by $E(\cdot)$. For simplicity, we do not distinguish in our notation between random variables and their realization.

The imaginary unit is j. Real and imaginary parts operators are represented by $\Re(\cdot)$ and $\Im(\cdot)$, respectively, $\arg(\cdot)$ is the angle of a complex number (in radiants) and $\log(\cdot)$ denotes the natural logarithm. Real, rational and integer numbers are denoted as $\mathbb{R}, \mathbb{Q}$ and $\mathbb{Z}$. $\mathcal{O}(f(\cdot))$ means that the magnitude of a quantity grows no faster than f.

Chapter 2

The signal model

The idea of OFDM transmission or more general multicarrier transmission is connected to the ideas of channel partioning, i.e. given a certain (probabilistic) channel operator, to transmit the eigenfunctions of the operator so that information passes through the channel without much distortion. If the radio channel is time-varying either the transmitted eigenfunctions must be reconfigured and adapted, which is hard to implement, or the chosen eigenfunctions must operate well even under different conditions. The basic mathematical theory to construct and design those eigenfunctions is represented by time-frequency analysis [49]. Fundamental work on this topic has been done by [50]. In this regard, OFDM represents a compromise for linear, time-invariant channels with respect to implementation and spectral efficiency. The signal is composed of a number of modulated versions of a prototype function (so-called Gabor-system [51]). It has recently been shown that the underlying structure of OFDM is matched to the linear, time-invariant channel in a deeper mathematical sense [52]. The connection to filter bank theory is described in [53].

In this chapter we describe the communication model and the basic OFDM system functionality. Furthermore, we introduce the necessary quantities relevant to our problem, such as the average power, crest-factor etc. Since the OFDM system has been described in various papers we assume basic knowledge of the system concept and for a more detailed description refer to [7, 54].

2.1 The system set up

Suppose a communication source emits an information symbol every T seconds. The stream of information symbols can be modeled as a sequence of independent, identically distributed (IID) random variables (RV) $\ldots, c_{-2}, c_1, c_0, c_1, c_2 \ldots$ which map into a subset of the complex plane called constellation or modulation scheme $\mathcal{Q}$. We shall assume in the sequel that the $c_k, k \in \mathbb{Z}$, have zero mean and a finite variance, i.e. $E(c_0) = 0$ and $E\left(|c_0|^2\right) = \sigma_s^2$. Common constellations used in practice are binary phase-shift keying $\mathcal{Q} = \text{BPSK} := \{A, -A\}$, quadrature phase-shift keying $\mathcal{Q} = \text{QPSK} := \{\pm A \pm jA\}$, M-ary quadrature amplitude modulation $\mathcal{Q} =$ M-QAM $:= \left\{A\left((2m_1 - 1) + j(2m_2 - 1)\right), m_1, m_2 \in \left\{-\frac{m}{2} + 1, \ldots, \frac{m}{2}\right\}\right\}$ for natural numbers m, M such that $M = m^2$ and M-ary phase-shift keying $\mathcal{Q} =$ M-PSK $:= \left\{A, Ae^{\frac{2\pi j}{M}}, \ldots, Ae^{\frac{2\pi j(M-1)}{M}}\right\}$ where $A > 0$ is a real constant normalizing the average power. A constellation is called equal-energy if $|c_k|$ is constant independent of the constellation points. BPSK and more generally M-ary PSK are equal-energy constellations.

Ordering the complex data into subsets of length N, i.e. $(c_{kN}, c_{kN+1}, \ldots, c_{kN+N-1})$, the OFDM baseband signal can be described by

$$s(t) = \sum_{m=-\infty}^{m=+\infty} w_{[-T_g, T_s]}\left(t - m\left(T_s + T_g\right)\right) \sum_{k=0}^{N-1} c_{mN+k}\, e^{2\pi j\left(k - \frac{N-1}{2}\right)\Delta f\left(t - m\left(T_s + T_g\right)\right)}, \tag{2.1}$$

where N is the number of subcarriers, T_s the symbol duration, $\Delta f = \frac{1}{T_s}$ the frequency spacing and $w_{[-T_g, T_s]}(t)$ denotes a standard rectangular window function of duration $NT = T_s + T_g$. The term T_g is the so-called guard interval. The rectangular window can be generally assumed when the number of subcarriers is large. For a smaller number of subcarriers, pulse shaping can be used in order to combat poor delay in the frequency domain. We have also assumed that the total power is uniformly spread over the subcarriers. A generalization to the non-equal power distribution for bitloading strategies in DMT systems is immediate.

For transmission the signal is passed to the RF chain to generate the passband signal $s^{(p)}(t) := \Re\left(s(t)\, e^{2\pi j f_c}\right)$ where f_c is the carrier frequency and is then transmitted over a channel with channel impulse response (CIR) $h^{(p)}$ that we assume to have support only in the interval

$[0, T_g]$. The guard interval prevents intersymbol interference (ISI) between consecutive OFDM symbols, provided that the length of the CIR does not exceed the guard interval. In other words, the transmit signal appears periodic to the channel.

At the receiver the RF-signal is down-converted, the guard interval is removed and the baseband signal is processed by matched filters. Let h be the equivalent baseband CIR of $h^{(p)}$ and $\widehat{h}$ the frequency response (see next chapter for definitions). Then due to the assumptions we have for $m = 0$

$$c'_k = \beta_k c_k + n_k, \quad k = 0, \ldots, N-1.$$

where $\beta_k := \widehat{h}\left(2\pi\left(k - \frac{N-1}{2}\right)\Delta f\right)$ are the complex channel attenuations and $n_k, k = \ldots, N-1$ is additive white Gaussian noise (AWGN). Likewise, the transmission can be written for arbitrary m. The AWGN can be modeled as zero mean, circular symmetric, complex Gaussian RV's denoted by $n_k \sim \mathcal{CN}(0, 1)$, i.e. real and imaginary parts are Gaussian distributed and independent. Finally, frequency equalization is performed and the symbols are passed to the decision unit in order to retrieve the information. To give an example: the HYPERLAN 2 standard is a 5 [GHz] OFDM system with 52 subcarriers. The symbol duration is 3.2 [μs], the subcarrier spacing is 312.5 [kHz]. The guard interval is 800 [ns] in order to provide robustness in indoor environment. Different modulation schemes such as BPSK, QPSK, 16-QAM and 64-QAM can be used.

A common model for the wireless channel is the multipath channel, i.e.

$$\widehat{h}(t) = \sum_{l=0}^{L_m-1} h_l \delta(t - \tau_l)$$

where L_m is the number of multipaths, δ is the Dirac delta ,,function" and $h_l, \tau_l, l = 0, \ldots, L_m - 1$ are the complex path attenuations and path delays, respectively. If the channel is time-varying different fading models are employed. We adopt the model with $h_l \sim \mathcal{CN}(0, r_l)$, where $r_l, l = 0, \ldots, L_m - 1$, is the power-delay profile, and fixed path delays. The channel is assumed to be constant over one or several symbols. Other channel models assume the path attenuations to be a wide-sense stationary random process with common Jakes power spectrum.

So far, we have introduced the conceptual OFDM system where the transceiver must realize a continuous-time Fourier transform. In practice, the signal s is generated at $\frac{lT_s}{L_T N}, l = 0, \ldots, L_T N$, by an $L_T N$-point inverse discrete Fourier transform (IDFT) of the complex data where L_T is the oversampling factor at the transmitter (in the case of $L = 1$ we speak of Nyquist-rate sampling and the Nyquist-rate samples) [7]. Then, a copy of the last L_1 samples is prepended (the guard interval) the samples are lowpass filtered to obtain the continuous-time baseband signal. The receiver performs reverse operations. The signal is sampled at $t = \frac{lT_s}{L_R N}, l = 0, \ldots, L_R N$ where L_R is the oversampling factor at the receiver, the guard interval is removed and a DFT carried out. For convenience we set $L_R = L_T = L \geq 1$. We point out that due to continuous-time filtering there will be a small amount of ISI in the transmission which is neglected here. In fact, in [7] different filters have been tested and it has been found that the model is appropriate for practical configurations. Clearly the chain of continuous-time filtering and channel can be replaced by an appropriate discrete-time channel based on the sampling theorem [55].

2.2 Average power and crest-factor

Let us now introduce some important quantities for the continuous-time and discrete-time signal that are related to the power control problem. For these purposes, it is sufficient to consider a single symbol, say $m = 0$, and to leave out the guard interval since the symbols do not overlap in time.

We interpret the sequence $c_0, c_1, \ldots, c_{N-1}$ as coordinates of a (random) vector. The square length of a vector c is given by

$$\|c\|^2 := \sum_{k=0}^{N-1} |c_k|^2 .$$

After multiplying by the complex factor $e^{\frac{j(N-1)}{2}}$ an OFDM baseband symbol can be described by

$$s_c(t) = \sum_{k=0}^{N-1} c_k e^{2\pi jk\Delta ft}, \quad 0 \leq t \leq T_s. \tag{2.2}$$

In the following, the time axis is normalized by T_s, i.e. we substitute $\theta(t) = \frac{2\pi t}{T_s}$ and write

$s_c(\theta), \theta \in [0, 2\pi]$, instead of $s_c(t)$ without explicitly stating the dependence of θ on t. For convenience we define $\theta_{l,L} := \frac{2\pi l}{LN}, 0 \leq l < LN$, where L is the oversampling factor introduced in the last section.

For any sequence $c_0, c_1, \ldots, c_{N-1}$ the instantaneous power of the transmit signal $\Re^2\left(s_c(\theta) e^{j\zeta\theta}\right)$ is lower or equal to the envelope power $\left|s_c(\theta) e^{j\zeta\theta}\right|^2$. Due to Parseval's theorem the time average envelope power is given by

$$\overline{P}_c := \frac{1}{2\pi} \int_0^{2\pi} |s_c(\theta)|^2 \, d\theta = \|c\|_2^2 .$$

The time average is taken over the symbol because the power is not constant over time. It is well-known that the time average power of the actual OFDM symbol is approximately equal to $\frac{1}{2}\|c\|_2^2$ if the ratio $\zeta := \frac{f_c}{\Delta f}$ is large. This is true for typical OFDM applications. For example in a HIPERLAN 2 system we have $\zeta \simeq 10^4$.

Except for BPSK, QPSK and M-ary PSK, $\overline{P}_c$ is a random variable. Assuming data symbols that occur with equal probability the variance σ_s^2 for an M-ary QAM constellation is given by

$$\sigma_s^2 = \frac{2A^2(M-1)}{3},$$

and the ensemble average is

$$\begin{aligned} P_{\text{av}} &= E\left(\overline{P}_c\right) \\ &= \frac{2NA^2(M-1)}{3}. \end{aligned}$$

The peak-to-average power ratio (PAPR) of (2.2) is now defined by

$$\text{PAPR}(s_c, \zeta) := \max_{0 \leq \theta < 2\pi} \frac{\left[\Re\left(s_c(\theta) e^{j\zeta\theta}\right)\right]^2}{P_{\text{av}}}.$$

The PAPR is also often called peak-to-mean power ratio (PMPR). The peak-to-mean envelope power (PMEPR) is defined by

$$\text{PMEPR}(s_c) := \max_{0 \leq \theta < 2\pi} \frac{|s_c(\theta)|^2}{P_{\text{av}}}$$

and clearly we have

$$\text{PAPR}(s_c, \zeta) \leq \text{PMEPR}(s_c).$$

The crest-factor (CF) of (2.2) is defined as

$$\begin{aligned}\mathrm{CF}\left(s_c\right) &:= \frac{1}{\sqrt{P_{\mathrm{av}}}} \max_{0\le\theta<2\pi} \left|s_c\left(\theta\right)\right| \\ &= \frac{1}{\sigma_s\sqrt{N}} \max_{0\le\theta<2\pi} \left|s_c\left(\theta\right)\right| .\end{aligned} \tag{2.3}$$

In the following we use CF instead of PAPR or PMEPR. Note that in the definition of CF the time average is replaced by its ensemble average. This is the relevant measure for power efficiency concerns because power consumption only depends on the ratio of peak power to average transmitted power. OFDM signals with small time average power are in this sense uncritical although their actual CF might be large. On the other hand from both the theoretical and practical viewpoint there are situations where the actual CF, given by

$$\mathrm{CF}'\left(s_c\right) := \frac{1}{\sqrt{\overline{P}_c}} \max_{0\le\theta<2\pi} \left|s_c\left(\theta\right)\right| ,$$

is considered.

Note that, clearly, $\mathrm{CF}'\left(s_c\right) \ge 1$. However, due to the statistical definition of CF apart from BPSK and QPSK we cannot assume $\mathrm{CF}\left(s_c\right) \ge 1$ for all signals s_c. Given σ_s^2, i.e. $A = \sqrt{\frac{3\sigma_s^2}{2(M-1)}}$, the minimal envelope power $\overline{P}_{\mathrm{min}}$ is

$$\begin{aligned}\overline{P}_{\mathrm{min}} &= 2NA^2 \\ &= \frac{3\sigma_s^2 N}{M-1}\end{aligned}$$

and since for all signals

$$\begin{aligned}\max_{0\le\theta<2\pi} \left|s_c\left(\theta\right)\right|^2 &\ge \overline{P}_c \\ &\ge \overline{P}_{\mathrm{min}}\end{aligned}$$

it follows

$$\begin{aligned}\max_{0\le\theta<2\pi} \left|s_c\left(\theta\right)\right| &\ge \sqrt{\overline{P}_{\mathrm{min}}} \\ &= \sqrt{\frac{3\sigma^2 N}{M-1}}.\end{aligned}$$

On the other hand an upper bound is given by

$$\max_{0\le\theta<2\pi} |s_c(\theta)| \le \sqrt{2}NA(M-1)$$
$$= \sigma N \sqrt{\frac{3\left(\sqrt{M}-1\right)}{\sqrt{M}+1}}$$

and thus

$$\sqrt{\frac{3}{M-1}} \le \mathrm{CF}(s_c) \le \sqrt{\frac{3N\left(\sqrt{M}-1\right)}{\sqrt{M}+1}}.$$

The signals that achieved these maximums have been identified in [55]. Note that the maximum CF, i.e. the CF of (2.1), is an appropriate quantity only if the probability of its occurrence is sufficiently high. However, it has only limited impact if its occurrence is very unlikely. Then, a better description of the sensitivity to non-linear distortion is given by the complementary distribution of (2.3) defined by

$$F(\lambda) := \Pr\left(\mathrm{CF}(s_c) > \lambda\right) \tag{2.4}$$

where $\lambda \ge 0$ is a real parameter. Clearly $F(\lambda) = 1$ if $\lambda \le \sqrt{\frac{3}{M-1}}$ and $F(\lambda) = 0$ if $\lambda \ge \sqrt{\frac{3N\left(\sqrt{M}-1\right)}{\sqrt{M}+1}}$ for M-ary QAM. For the other values the term (2.4) is difficult to analyze. The statistical distribution of the CF is a key parameter in OFDM system design providing both practical and information-theoretic insights, and we analyze this parameter in Chapter 3. The exact evaluation is still an on-going research problem.

Using (2.4) in [29, 30] an effective CF was defined by

$$\Pr\left(\mathrm{CF}_{\mathrm{eff}}(s_c) > \varepsilon\right) \tag{2.5}$$

where ε is a small probability that may depend on the application so that the transmission can be considered almost distortion-free.

In order to investigate (2.4) for practical systems usually a discrete version of (2.2) is introduced, i.e.

$$\mathrm{CF}_L(s_c) := \max_{0\le l<LN} \frac{|s_c(\theta_{l,L})|}{\left(\frac{1}{LN}\sum_{l=0}^{LN-1} E\left(|s_c(\theta_{l,L})|^2\right)\right)^{\frac{1}{2}}} \tag{2.6}$$

In the sequel the term (2.6) is called the discrete-time CF. For the particular case $L = 1$ we call $\mathrm{CF}_1(s_c)$ the Nyquist-rate CF.

The complementary distribution of (2.6) is denoted as

$$F_L(\lambda) := \Pr\left(\mathrm{CF}_L(s_c) > \lambda\right).$$

Note again that due to Parsevals theorem [18]

$$\frac{1}{LN}\sum_{l=0}^{LN-1}\left|s_c(\theta_{l,L})\right|^2 = \frac{1}{2\pi}\int_0^{2\pi}\left|s_c(\theta)\right|^2 d\theta$$

the root mean square (RMS) value is preserved and the CF is

$$\mathrm{CF}_L(s_c) = \frac{1}{\sigma_s\sqrt{N}}\max_{0\le l<LN}\left|s_c(\theta_{l,L})\right|.$$

Given these definitions we now turn to the central power control problem.

2.3 The power control problem

So far, the system concept appears to be rather appealing due to the simple equalization and efficient signal processing. We also pointed out that it is very much adapted to the linear, time-invariant channel. However, the advantages are accompanied by the disadvantage that the system is very sensitive to non-linear components in the transmitter path, such as digital analog (DA) converter, mixers and high power amplifiers (HPA) since many subcarriers (constructively or destructively) add up at a time causing large fluctuations of the signal envelope. A transmission which is free from any distortion requires linear operation over a range N times the average power. As practical values of subcarriers are large the high dynamics prevents power efficient operation of the HPA so that most of the DC power is wasted (see the model below) with deleterious effect on battery life time in mobile applications. In practice, these values are not tolerable (from a technology viewpoint it may even be challenging to provide such a large linear range) and the signal is cut off at some point leading to in-band distortion in form of intermodulation among subcarriers and spectral widening of the transmit signal into adjacent channels. Thus, some form of peak power reduction is desirable.

A common baseband representation for the HPA is the soft-envelope limiter model, i.e.

$$\Phi_{\mathrm{SL}}\left(s\left(t\right)\right):=\begin{cases} s\left(t\right), & \left|s\left(t\right)\right|\leq \mathrm{SL} \\ \mathrm{SL}e^{j\arg[s(t)]}, & \left|s\left(t\right)\right|>\mathrm{SL} \end{cases},$$

Here, SL is the saturation level (or clipping level). The event $\{|s(t)| > \mathrm{SL}\}$ is commonly described as clipping, i.e. the signal envelope is clipped („clipping") at a certain level. The model is clearly somehow idealized since a linear region is only possible by taking additional measures such as pre-distortion techniques [56]. Nevertheless, it is sufficient for characterizing the behavior of standard systems. It accounts as well for the limiting effect of other non-linearities in the signal path that may act only on the discrete-time signal such as the DA converter. Other models are for example the solid-state HPA and the traveling-wave tube HPA which we do not consider here.

The ratio

$$\mathrm{CR}:=\frac{\mathrm{SL}}{\sqrt{P_{\mathrm{av}}}}$$

is called the clipping ratio (CR). The square of CR is the input back-off factor (IBU). The output back-off (OBO) is defined by the ratio of SL and average output power, i.e. it is generally different from IBU. However if the SL is not too small IBU is approximately equal to OBO. The power efficiency is approximately the inverse of the OBO for a class A amplifier [57].

A common performance measure to assess the communication link is the symbol error rate (SER). Denoting the estimated information symbols $\widehat{c}_k$ the average symbol error probability is given by

$$\overline{\mathrm{SER}}=\frac{1}{N}\sum_{k=0}^{N-1}\mathrm{SER}_k,$$

where

$$\mathrm{SER}_k=\Pr\left(c_k\neq\widehat{c}_k\right)$$

is the SER on subcarrier k. For standard channels like AWGN or Rayleigh fading multipath fading channels explicit formulas of the SER are known [58]. Note that in practical systems a trade-off between SER and power efficiency has to be found. A common measure for trading SER due to clipping and power efficiency is total degradation (TD), i.e. minimizing the sum of OBO and energy per bit degradation due to clipping that is needed to meet a certain target SER [59][60].

Since we introduced continuous-time and discrete-time criteria it is natural to ask for their relationships. The next chapter is dedicated to a fundamental analysis of this problem.

Chapter 3

The discrete and continuous-time crest-factor

An OFDM signal can be regarded as a complex-valued trigonometric polynomial and throughout this thesis we can benefit from the rich mathematical literature in this field. In particular, the problem of estimating the peak value of a trigonometric polynomial from its samples with and without oversampling has been considered for a long period of time. Early results date back to the work of [3] stating already a fundamental relation in this regard. The work was rediscovered by [5]. More recent results are due to [8, 9] providing much weaker results while introducing the problem in the context of OFDM. In [9] a bound for Nyquist-rate sampling was derived based on Lagrange interpolation. In [18] a bound for complex trigonometric polynomials was derived and improved in [6]. It was particularly emphasized in [18] that the bound for Nyquist-rate sampling is too weak.

In this chapter, we will adopt a rather general viewpoint to this problem in the sense that we do not only consider polynomials but band-limited signals. This has the advantage that we have a general framework and can use the same results again when using window functions other than the standard rectangular one, leading to several important design aspects even for more general multicarrier systems than those considered in this thesis. In particular, an upper bound on the peak value is established given the peak value of the samples and the oversampling rate. Moreover, it is shown that the bounds are sharp for all practical rates by constructing band-limited signals taking on this bound. The proof also provides a local characterization of band-limited signals in the neighborhood of an extremum. We point out that the results have the particular advantage that the peak can be locally calculated by the samples in the neighborhood. Due to the band-limited context this is obviously more suited to real systems. Furthermore, we also distinguish between the case where the samples are disturbed and an ideal situation. A different analysis examines the effect of small errors in the samples. It is shown that in this situation the peaks strongly depend on the interpolation filter used. It is further shown that oversampling can provide robust recovery in the sense that small errors in the samples lead to small errors in the reconstructed signal. Again, an upper bound is derived relating the peak error in the samples and the peak error in the signals. Furthermore, both problems are shown to be coupled and put in a unifying context. The bounds are compared and applied to bound the CF of OFDM signals. Apart from the filtering issues the analysis has some strong coding implications as we show in Chapter 5. Although not treated here we also mention the influence on interpolation techniques in pilot-based channel estimation [61]. Note that the chapter provides the necessary background for the analysis in the subsequent chapters.

3.1 Problem statement

Let us first provide some notations and thereby rigorously formulize the problem. The Fourier transform of a signal f is denoted as $\widehat{f}$ where $\widehat{f}$ is to be understood in the distributional sense. A signal is called band-limited with B if the Fourier transform is supported on $[-B, B]$. The collection of signals whose p-th power is integrable is denoted by $\mathcal{L}^p(\mathbb{R})$ with the common norm $\|\cdot\|_p$. For $p = \infty$ the norm is given by the supremum norm. The set of band-limited signals with bandwidth B in $\mathcal{L}^p(\mathbb{R})$ form the Paley-Wiener[1] space $\mathcal{PW}_B^p$. Note that the inclusions $\mathcal{PW}_B^1 \subset \mathcal{PW}_B^2 \subset \ldots \subset \mathcal{PW}_B^\infty$ hold for Paley-Wiener spaces [63]. For further purposes let us also introduce the space of bounded, continuous signals over $\mathbb{R}$, denoted by $C(\mathbb{R})$ and endowed with the supremum norm, its subspace $C_{2\pi}(\mathbb{R})$ of all bounded and continuous signals over $\mathbb{R}$ that are periodic with period 2π.

[1] The literature seems not to be clear here. We follow [62] in our definition of the Paley-Wiener spaces but for example in [63] these spaces are called Bernstein spaces.

It is well-known that a signal $f \in \mathcal{PW}_B^p, 1 \leq p < \infty$, can be recovered by its samples by applying the Shannon sampling series

$$f(\theta) = \sum_{l=-\infty}^{\infty} f\left(\frac{\pi l}{B}\right) \frac{\sin\left[B\left(\theta - \frac{\pi l}{B}\right)\right]}{B\left(\theta - \frac{\pi l}{B}\right)} \tag{3.1}$$

which is a simple consequence of the Plancherel-Polya inequality

$$C'(p)\left(\frac{1}{B}\sum_{l=-\infty}^{\infty}\left|f\left(\frac{\pi l}{B}\right)\right|^p\right)^{\frac{1}{p}} \leq \|f\|_p \leq C''(p)\left(\frac{1}{B}\sum_{l=-\infty}^{\infty}\left|f\left(\frac{\pi l}{B}\right)\right|^p\right)^{\frac{1}{p}},\ C'(p), C''(p) > 0,$$

for $1 < p < \infty$ and the inclusion $\mathcal{PW}_B^1 \subset \mathcal{PW}_B^2$. Indeed, let $n > 0$. Then for every $f \in \mathcal{PW}_B^p, 1 < p < \infty$ we have $\left\|f - \sum_{l=-n}^{n} f\left(\frac{\pi l}{B}\right)\right\|_p \leq C''(p)\sum_{|l|>n} f\left(\frac{\pi l}{B}\right)$ and thus $\left\|f - \sum_{l=-n}^{n} f\left(\frac{\pi l}{B}\right)\right\|_p \to 0$ as $n \to \infty$.

The series (3.1) fails to converge in general for $p = \infty$. In this case, the sampling series due to Schönhage [64]

$$f(\theta) = f'(0)\frac{\sin(B\theta)}{B} + f(0)\frac{\sin(B\theta)}{B\theta} + \theta\sum_{l=-\infty, l\neq 0}^{\infty}\frac{f\left(\frac{\pi l}{B}\right)}{l}\frac{\sin\left[B\left(\theta - \frac{\pi l}{B}\right)\right]}{B\left(\theta - \frac{\pi l}{B}\right)} \tag{3.2}$$

can be applied converging uniformly on compact subsets of $\mathbb{R}$. The representation (3.2) is in fact the key to our derivations regarding the space $\mathcal{PW}_B^\infty$ in the next section.

On the other hand, divergence can also be circumvented if oversampling is employed and the $\frac{\sin(B\theta)}{B\theta}$-kernel in (3.1) is replaced by an absolute integrable kernel. Introducing the oversampling factor $L > 1, L \in \mathbb{R}$, the peak value of an oversampling set is given by

$$\|f\|_{\frac{\pi}{LB},\infty} := \sup_{l\in\mathbb{Z}}\left|f\left(\frac{\pi l}{LB}\right)\right|.$$

We will now see that it is an equivalent norm on the space $\mathcal{PW}_B^p$. Defining the set

$$\mathcal{M}_L^B := \left\{ g \in L^1(\mathbb{R}), \widehat{g}(\omega) = \left\{\begin{array}{cc} 1 & |\omega| \leq B \\ \widehat{g}_d(\omega) & B \leq |\omega| \leq LB \\ 0 & \text{elsewhere} \end{array}\right.\right\}, \tag{3.3}$$

where $\widehat{g}_d(\omega)$ is a real function with $0 \leq \widehat{g}_d(\omega) \leq 1, \widehat{g}_d(B) = 1, \widehat{g}_d(LB) = 0$. For some reasons that will become clear later on we assume $\widehat{g}_d(\omega)$ to be a non-increasing function. Now, for all

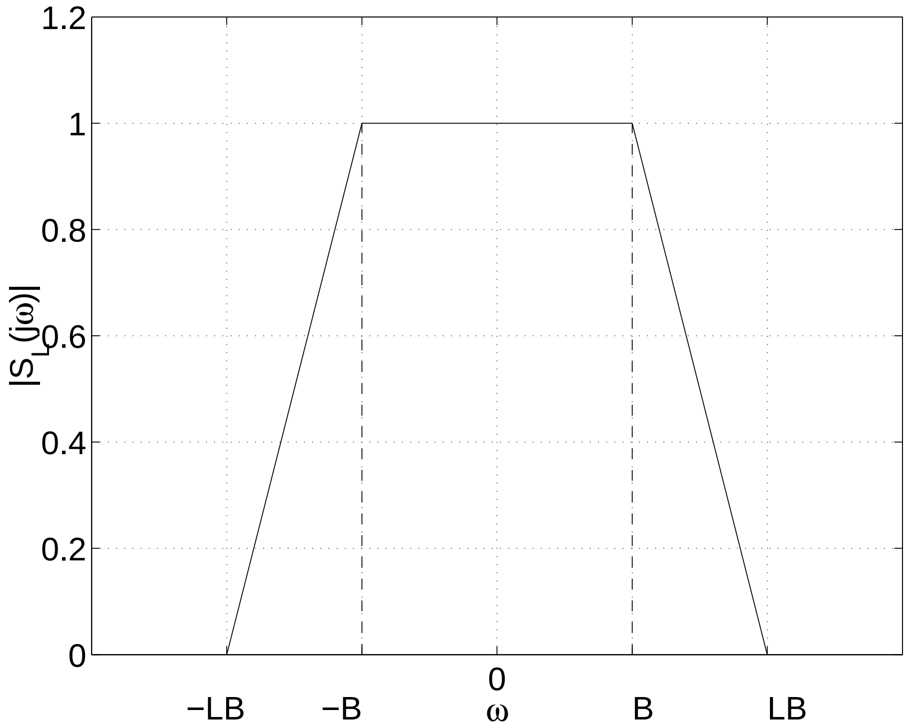

Figure 3.1: The Fourier transform of the trapezoidal kernel.

$f \in \mathcal{PW}_B^p, 1 \leq p < \infty$, we have [20, 20, 65]

$$f(\theta) = \frac{\pi}{LB} \sum_{l=-\infty}^{\infty} f\left(\frac{\pi l}{LB}\right) g\left(\theta - \frac{\pi l}{LB}\right) \tag{3.4}$$

and

$$\sup_{\theta \in \mathbb{R}} |f(\theta)| \leq \|f\|_{\frac{\pi}{LB},\infty} \cdot \sup_{\theta \in \left[0, \frac{\pi}{LB}\right]} \frac{\pi}{LB} \sum_{l=-\infty}^{\infty} \left| g\left(\theta - \frac{\pi l}{LB}\right)\right| < \infty,$$

i.e. every kernel $g \in \mathcal{M}_L$ defines a bounded, linear operator $T_g : \mathcal{PW}_B^\infty \rightarrow \mathcal{PW}_B^\infty$. Thus, equation (3.4) can be extended to the space $\mathcal{PW}_B^\infty$. An example of a kernel is given by the trapezoidal kernel

$$S_L(\theta) = \frac{2\sin\left(\frac{(L+1)B\theta}{2}\right)\sin\left(\frac{(L-1)B\theta}{2}\right)}{\pi(L-1)B\theta^2}, \tag{3.5}$$

of which the Fourier transform is depicted in Fig. 3.1. Using the convolution theorem of the Fourier transform it can be verified that this kernel satisfies the assumptions made in (3.3).

We are now in position to state the first problem:

1) Peak Value Problem: Find a ,,good" (i.e. tight) upper bound on the constant

$$C_1(L) := \sup_{\|f\|_{\frac{\pi}{LB},\infty} \le 1, f \in \mathcal{PW}_B^\infty} \|f\|_\infty .$$

It is interesting to note that $C_1(L)$ is independent of B. This can be seen as follows: For any $f \in \mathcal{PW}_B^\infty$ define $f'(\theta) := f\left(\frac{\pi\theta}{B}\right)$ with $f' \in \mathcal{PW}_\pi^\infty$ and hence

$$\sup_{\|f'\|_{\frac{1}{L},\infty} \le 1, f' \in \mathcal{PW}_\pi^\infty} \|f'\|_{\frac{1}{L},\infty} = \sup_{\|f\|_{\frac{\pi}{LB},\infty} \le 1, f \in \mathcal{PW}_B^\infty} \|f\|_{\frac{\pi}{LB},\infty} .$$

We further note that $C_1(L)$ is not defined for $L = 1$ (i.e. $C_1(1) = +\infty$). Consider for example the sequence

$$f_n(\theta) = \sum_{l=1}^{n} (-1)^l \frac{\sin\left[B\left(\theta - \frac{\pi l}{B}\right)\right]}{B\left(\theta - \frac{\pi l}{B}\right)}$$

where $f_n \in \mathcal{PW}_B^\infty, n \in \mathbb{N}$, and $\|f_n\|_\infty \to \infty$ as $n \to \infty$. The following theorem ensures the existence of a function f such that $C_1(L) = \|f_L\|_\infty$. We will call these signals extremal. Observe that in the proof we explicitly construct these signals.

Lemma 1 *There is a function $f_L \in \mathcal{PW}_\pi^\infty$ and a time instance $\theta_L \in \left(0, \frac{1}{2L}\right]$ so that $\|f_L\|_{\frac{1}{L},\infty} = 1$ and*

$$C_1(L) = f_L(\theta_L) .$$

Proof. We can without loss of generality assume that $B = \pi$. For any arbitrary $\varepsilon > 0$ there is $f^{(\varepsilon)}, \theta^{(\varepsilon)}$ with $\left\|f^{(\varepsilon)}\right\|_{\frac{1}{L},\infty} = 1$ and $f^{(\varepsilon)}\left(\theta^{(\varepsilon)}\right) \ge C_1(L) - \varepsilon$. We let $l^{(\varepsilon)}$ be the integer number so that $\frac{l^{(\varepsilon)}}{L} < \theta^{(\varepsilon)} \le \frac{l^{(\varepsilon)}}{L} + \frac{l^{(\varepsilon)}}{2L}$. Then

$$g^{(\varepsilon)}(\theta) := f^{(\varepsilon)}\left(\theta + \frac{l^{(\varepsilon)}}{L}\right)$$

where $g^{(\varepsilon)}\left(\theta^{(\varepsilon)}\right) \ge C_1(L) - \varepsilon$ and $0 < \theta^{(\varepsilon)} \le \frac{1}{2L}$. Assuming ε_n to be a zero sequence then it follows that the sequence $\theta^{(\varepsilon_n)}$ is bounded and it is possible to select a subsequence $\theta^{\left(\varepsilon_{n_k}\right)} =: \theta^{(k)}$ so that

$$\left\|g^{\left(\varepsilon_{n_k}\right)}\right\|_{\frac{1}{L},\infty} = 1, \quad k \in \mathbb{N},$$

and

$$\theta^{(\infty)} = \lim_{k \to \infty} \theta^{\left(\varepsilon_{n_k}\right)}.$$

Denoting the corresponding sequence of functions by $g^{\left(\varepsilon_{n_k}\right)} =: g^{(k)}$ we can again select a subsequence $g^{(k')}$ that converges due to a theorem of Vitali [66] locally uniformly to an entire function $g^{(\infty)} \in \mathcal{PW}_\pi^\infty$ [64, (2.19)]. It is left to show that

$$g^{(\infty)}\left(\theta^{(\infty)}\right) = C_1(L)$$

We have $\left|g^{(k')}\left(\theta^{(\infty)}\right) - g^{(\infty)}\left(\theta^{(\infty)}\right)\right| \leq \varepsilon$ and $\left|\theta^{(\infty)} - \theta^{(k')}\right| \leq \varepsilon$. Using Bernstein's inequality [67] $\left|g^{(k')\prime}(\theta)\right| \leq C < \infty$ uniformly for all θ and hence

$$\begin{aligned}
\left|g^{(k')}\left(\theta^{(k')}\right) - g^{(\infty)}\left(\theta^{(\infty)}\right)\right| &\leq \left|g^{(k')}\left(\theta^{(k')}\right) - g^{(k')}\left(\theta^{(\infty)}\right) + g^{(k')}\left(\theta^{(\infty)}\right) - g^{(\infty)}\left(\theta^{(\infty)}\right)\right| \\
&\leq \left|g^{(k')}\left(\theta^{(k')}\right) - g^{(k')}\left(\theta^{(\infty)}\right)\right| + \left|g^{(k')}\left(\theta^{(\infty)}\right) - g^{(\infty)}\left(\theta^{(\infty)}\right)\right| \\
&\leq \left|g^{(k')\prime}(\theta)\right| \left|\theta^{(k')} - \theta^{(\infty)}\right| + \varepsilon, \quad \theta \in \left[\theta^{(k')}, \theta^{(\infty)}\right] \\
&\leq (1 + C)\,\varepsilon
\end{aligned}$$

yielding the theorem. ■

Let us introduce the constant

$$C_1^{\mathbb{R}}(L) := \sup_{\|f\|_{\frac{1}{L},\infty} \leq 1, f \in \mathcal{PW}_\pi^\infty, f(\theta) \in \mathbb{R}} \|f\|_\infty,$$

i.e. $C_1^{\mathbb{R}}(L)$ describes the constant $C_1(L)$ if f is confined to be a real-valued signal of $\mathcal{PW}_\pi^\infty$. Clearly, $C_1^{\mathbb{R}}(L) \leq C_1(L)$ and obviously taking squared terms $C_1(L) \leq \sqrt{C_1^{\mathbb{R}}\left(\frac{L}{2}\right)}, L > 2$. We now prove $C_1(L) \leq C_1^{\mathbb{R}}(L)$ and thus $C_1(L) = C_1^{\mathbb{R}}(L)$.

Lemma 2 *We have* $C_1(L) = C_1^{\mathbb{R}}(L)$.

Proof. By definition we have $C_1^{\mathbb{R}}(L) \leq C_1(L)$. According to Lemma 1 there is a complex signal $f_L \in \mathcal{PW}_\pi^\infty$ and a time instance θ_L so that

$$C_1(L) = |f_L(\theta_L)|$$

with $\|f_L\|_{\frac{1}{L},\infty} = 1$ and $f_L(\theta_L) = C_1(L)\, e^{j \arg[f_L(\theta_L)]}$. Multiplying f_L by $e^{-j \arg[f_L(\theta_L)]}$ yields a signal with complex samples and (at least one) real maximum. Taking the real part of this signal

$$f_L^{\mathbb{R}}(\theta) := \Re\left(f_L(\theta)\, e^{-j \arg[f_L(\theta_L)]}\right)$$

gives a real signal $f_L^{\mathbb{R}}(\theta)$ with

$$\left\|f_L^{\mathbb{R}}\right\|_\infty = C_1(L)$$

and

$$\begin{aligned}
\left|f_L^{\mathbb{R}}\left(\frac{l}{L}\right)\right| &= \left|\Re\left(f_L\left(\frac{l}{L}\right) e^{-j \arg[f_L(\theta_L)]}\right)\right| \\
&\leq \left|f_L\left(\frac{l}{L}\right) e^{-j \arg[f_L(\theta_L)]}\right| \\
&= \left|f_L\left(\frac{l}{L}\right)\right| \\
&\leq 1.
\end{aligned}$$

Thus, there is a real signal so that $C_1^{\mathbb{R}}(L) \geq C_1(L)$. Considering $C_1^{\mathbb{R}}(L) \leq C_1(L)$ it follows that $C_1^{\mathbb{R}}(L) = C_1(L)$. ■

In practical systems, the samples are disturbed and when oversampling is employed they generally do not represent the samples of a band-limited signal with respect to the bandwidth defined by the Nyquist rate. It is therefore reasonable to extend the definition region of the operator T_g to the space $C(\mathbb{R})$, i.e. $T_g : C(\mathbb{R}) \to \mathcal{PW}_{LB}^{\infty}, f \hookrightarrow \frac{\pi}{LB} \sum_{l=-\infty}^{\infty} f\left(\frac{\pi l}{LB}\right) g\left(\theta - \frac{\pi l}{LB}\right)$. The norm of this operator is given by

$$|T_g| = \sup_{\|f\|_\infty \leq 1, f \in C(\mathbb{R})} \|T_g f\|_\infty .$$

The operator norm represents the enhancement of errors in the samples. This leads us to the next problem:

2) Noisy Samples Problem: Find a ,,good” upper bound on the operator norm

$$C_2(L) = \inf_{g \in \mathcal{M}_L^B} |T_g| .$$

Again, note that $C_2(L)$ is independent of B. For any $g \in \mathcal{M}_L^B$ define $\widehat{g}'(\omega) = \widehat{g}\left(\frac{B\omega}{\pi}\right)$, i.e. $g' \in \mathcal{M}_L^\pi$ and $g'(\theta) = g\left(\frac{\pi\theta}{B}\right)$ and $|T_g| = |T_{g'}|$. Further, clearly $C_1(L) \le C_2(L)$, i.e. $C_1(L)$ represents a lower bound on what can be achieved for $C_2(L)$. For the purposes of filter design it is also interesting which kernel actually attains this bound. These filters will be called extremal filters and their existence is established in the next theorem.

Lemma 3 *There is a signal f_L, a time instance θ_L, and a kernel g_L such that $C_2(L) = (T_{g_L} f_L)(\theta_L)$.*

Proof. Without loss of generality let $B = \pi$. Furthermore, fix L and g and consider for some arbitrary $f \in C(\mathbb{R})$ with $\|f\|_\infty \le 1$

$$(T_g f)(\theta) = \sum_{l=-\infty}^{+\infty} f\left(\frac{l}{L}\right) g\left(\theta - \frac{l}{L}\right).$$

We have

$$\|T_g f\|_\infty \le \max_{0 \le \theta < \frac{1}{2L}} \sum_{l=-\infty}^{+\infty} \left| g\left(\theta - \frac{l}{L}\right)\right| \cdot \|f\|_{\frac{1}{L},\infty} = \max_{0 \le \theta < \frac{1}{2L}} \sum_{l=-\infty}^{+\infty} \left| g\left(\theta - \frac{l}{L}\right)\right|.$$

Here, we replaced the supremum by the maximum operator since $\sum_{l=-\infty}^{+\infty} \left|g\left(\theta - \frac{l}{L}\right)\right| \in C(\mathbb{R})$. Hence, the operator norm is upperbounded by

$$|T_g| \le \max_{0 \le \theta < \frac{1}{2L}} \sum_{l=-\infty}^{+\infty} \left| g\left(\theta - \frac{l}{L}\right)\right|.$$

where the maximum is attained for some $\theta_0 \in \left[0, \frac{1}{2L}\right)$. On the other hand, defining a signal $f_0 \in C(\mathbb{R})$ with $f_0\left(\frac{l}{L}\right) := e^{-j \arg\left[g\left(\theta_0 - \frac{l}{L}\right)\right]}$ yields

$$\|T_g f_0\|_\infty = \max_{0 \le \theta < \frac{1}{2L}} \sum_{l=-\infty}^{+\infty} \left| g\left(\theta - \frac{l}{L}\right)\right| = |T_g|.$$

Hence, we conclude that for each g we can find f_0, θ_0 so that $|T_k| = (T_g f_0)(\theta_0)$ where $\theta_0 \in \left[0, \frac{1}{2L}\right)$. Since $C_2(L) \ge C_1(L)$ we can assume that for any arbitrary $\varepsilon > 0$ there is $g^{(\varepsilon)}, \theta^{(\varepsilon)}$ with $g^{(\varepsilon)}\left(\theta^{(\varepsilon)}\right) \ge C_2(L) - \varepsilon$. Let ε_n be a zero sequence. We can select a subsequence $\theta^{\left(\varepsilon_{n_k}\right)} =: \theta^{(k)}$ so that there is $\theta^{(\infty)}$ with

$$\theta^{(\infty)} := \lim_{k\to\infty} \theta^{(k)}.$$

By the same argument as in Lemma 1 we can also find a subsequence $g^{\left(\varepsilon_{n_k}\right)} =: g^{(k)}$ so that there is a $g^{(\infty)} \in \mathcal{PW}_{LB}^{\infty}$ with

$$g^{(\infty)} := \lim_{k\to\infty} g^{(k)}$$

where convergence is again local on compact subsets. We need to show that $g^{(\infty)} \in \mathcal{M}_L^{\pi}$. First observe that

$$\lim_{k\to\infty} \int_{-T}^{T} g^{(k)} dt = \int_{-T}^{T} g^{(\infty)} dt \leq C$$

for any interval $T \in \mathbb{R}$. Thus $g^{(\infty)} \in \mathcal{PW}_{LB}^{1}$ but not necessarily $g^{(\infty)} \in \mathcal{M}_L^{\pi}$. This can be achieved since $\widehat{g}^{(\infty)}$ is a signal with bounded variations by the assumptions on $\widehat{g}_d$ in (3.3), i.e. due to Helly's theorem [68] we have point-wise convergence in the frequency domain, and therefore $g^{(\infty)} \in \mathcal{M}_L^{\pi}$.

Denoting the signal that attains the operator norm as $f^{(\infty)}$ corresponding to $g^{(\infty)}$ it is left to show that

$$C_2(L) = \left(T_{g^{(\infty)}} f^{(\infty)}\right)\left(\theta^{(\infty)}\right).$$

Observe that

$$C_2(L) \leq \sum_{l=-\infty}^{+\infty} \left| g^{(\infty)} \left(\theta^{(\infty)} - \frac{l}{L} \right) \right|. \tag{3.6}$$

For any θ and n consider

$$\begin{aligned}
\sum_{l=-n}^{+n} \left| g^{(\infty)} \left(\theta - \frac{l}{L} \right) \right| &\leq \sum_{l=-n}^{+n} \left| g^{(\infty)} \left(\theta - \frac{l}{L} \right) - g^{(k)} \left(\theta - \frac{l}{L} \right) \right| + \sum_{l=-n}^{+n} \left| g^{(k)} \left(\theta - \frac{l}{L} \right) \right| \\
&\leq \sum_{l=-n}^{+n} \left| g^{(\infty)} \left(\theta - \frac{l}{L} \right) - g^{(k)} \left(\theta - \frac{l}{L} \right) \right| + \max_{0\leq\theta<\frac{1}{2L}} \sum_{l=-n}^{+n} \left| g^{(k)} \left(\theta - \frac{l}{L} \right) \right|
\end{aligned}$$

and

$$\begin{aligned}
&\sum_{l=-n}^{+n} \left| g^{(\infty)} \left(\theta - \frac{l}{L} \right) \right| \\
&\leq \limsup_{k\to\infty} \sum_{l=-n}^{+n} \left| g^{(\infty)} \left(\theta - \frac{l}{L} \right) - g^{(k)} \left(\theta - \frac{l}{L} \right) \right| + \limsup_{k\to\infty} \max_{0\leq\theta<\frac{1}{2L}} \sum_{l=-n}^{+n} \left| g^{(k)} \left(\theta - \frac{l}{L} \right) \right| \\
&= \limsup_{k\to\infty} \max_{0\leq\theta<\frac{1}{2L}} \sum_{l=-n}^{+n} \left| g^{(k)} \left(\theta - \frac{l}{L} \right) \right| \\
&= C_2(L).
\end{aligned}$$

Therefore

$$C_2(L) \geq \sum_{l=-\infty}^{+\infty} \left| g^{(\infty)} \left(\theta^{(\infty)} - \frac{l}{L} \right) \right|$$

and due to (3.6)

$$C_2(L) = \sum_{l=-\infty}^{+\infty} \left| g^{(\infty)} \left(\theta^{(\infty)} - \frac{l}{L} \right) \right|$$

completing the proof. ∎

Note that for pure technical reasons we have assumed (applying Helly's Theorem) that the function $\widehat{g}_d$ is non-increasing (the problem could be circumvented if we were able to show weak convergence instead of point-wise convergence of $g^{(\varepsilon)}$, however, unfortunately the $g^{(k)}$ do not have compact support). However we will see that this is not particularly restrictive since the trapezoidal filter already comes very close to the lower bound on C_2 given by C_1. For the special case of polynomials we derive a general theorem that excludes the existence of better kernels. Since the solution to both the peak value and noisy samples problems does not depend on the bandwidth we will always scale the time or frequency domain accordingly.

3.2 Peak value problem

3.2.1 Trigonometric polynomials

The most early results [3] regarding the peak value problem appear in literature when restricting $\mathcal{PW}_B^\infty$ to the set

$$\mathcal{T}_N := \left\{ f \in \mathcal{PW}_N^\infty;\ f(\theta) = \frac{a_0}{2} + \sum_{k=1}^{N} a_k \cos(k\theta) + b_k \sin(k\theta), \quad a_k, b_k \in \mathbb{R} \right\},$$

containing all degree N trigonometric polynomials. Observe that the signal space $\mathcal{T}_N$ covers the so-called discrete multitone (DMT) signals. Since the signals are periodic and due to the finite dimension of the signal space it is sufficient to consider the interval $[0, 2\pi)$ and a finite set of samples. In this light, a polynomial can be recovered by at least $N_1 = 2N + 1$ samples and we define the oversampling factor as $L = \frac{N_1}{2N}$ where $N_1 > 2N$ is the number of samples, i.e. opposed to general band-limited signals only certain oversampling rates are permitted. Since every polynomial of degree N is also a polynomial of degree $N_2 > N$ the ,,reasonable" set can even be extended to cover all the rational numbers $\mathbb{Q} \cap (1, \infty)$.

Based on a classical result obtained by M. Riesz [3] the authors in [5] proved the following theorem.

Theorem 4 *Let* $f \in \mathcal{T}_N$ *and suppose that* $f(\theta) \neq \|f\|_\infty \cos[N(\theta - \theta_0)], \theta_0 \in \mathbb{R}$. *Then for* $N' > N, N' \in \mathbb{N}$,

$$\|f\|_\infty < \frac{1}{\cos\left(\frac{\pi N}{2N'}\right)} \max_{0 \leq l < 2N'} \left| f\left(\frac{\pi l}{N'}\right)\right|. \tag{3.7}$$

with equality if $\frac{N'}{N} = 2^m, m \in \mathbb{N}$.

In terms of the oversampling factor equation (3.7) reads

$$\|f\|_\infty < \frac{1}{\cos\left(\frac{\pi}{2L}\right)} \max_{0 \leq l < 2LN} \left| f\left(\frac{\pi l}{LN}\right)\right|.$$

where $L = \frac{N'}{N}, N' > N$, i.e.. the number of samples is even and the condition $N' > N$ implies that the number of samples is at least $2N + 2$. However, it can be easily seen (see also Theorem 6) that (3.7) is also valid for arbitrary $N_1 > 2N$, i.e. including odd numbers. Equality holds in (3.7) if and only if $L = \frac{N_1}{2N} \in \mathbb{N}$. This can be seen as follows: Sampling the signal $f_L(\theta) = \cos\left[N\left(\theta - \frac{\pi}{2LN}\right)\right], L > 1, L \in \mathbb{N}$, gives

$$\left| f_L\left(\frac{\pi l}{LN}\right)\right| \leq \cos\left(\frac{\pi}{2L}\right), \quad l \in \mathbb{Z}, \quad \text{and} \quad f_L\left(\frac{\pi}{2LN}\right) = 1,$$

since if $\cos N\theta$ is sampled at rate $\frac{LN}{\pi}$ the samples always meet the points $\frac{l\pi}{N}, l \in \mathbb{Z}$. Thus $C_1(L) \geq \frac{1}{\cos\left(\frac{\pi}{2L}\right)}$ and with $C_1(L) \leq \frac{1}{\cos\left(\frac{\pi}{2L}\right)}$ it follows $C_1(L) = \frac{1}{\cos\left(\frac{\pi}{2L}\right)}$. Thus, for practical oversampling rates it is simply the highest in-band frequency that has the worst-case behavior regarding overshooting between the samples.

Introducing the constant

$$C_1^{\mathcal{T}_N}(N_1) := \max_{\|f\|_{\frac{2\pi}{N_1},\infty} \leq 1, f \in \mathcal{T}_N} \|f\|_\infty \tag{3.8}$$

we have therefore

$$C_1^{\mathcal{T}_N}(2LN) \leq \frac{1}{\cos\left(\frac{\pi}{2L}\right)} \tag{3.9}$$

with equality if $L = \frac{N_1}{2N}$ integer, i.e. N_1 is a multiple of $2N$. The extremal signals of (3.8) are denoted as $\mathcal{T}$-$(N, 2LN)$-extremal. Their existence follows from a standard compactness argument. The problem of constructing these polynomials for arbitrary oversampling was regarded in [6]. It was shown that in general cases too the extremal signals only depend on the ratio $\frac{N_1}{N}$.

We can also consider the degree N complex-valued trigonometric polynomials

$$\mathcal{P}_N := \left\{ f \in \mathcal{PW}_N^\infty;\ f(\theta) = \sum_{k=-N}^{N} c_k\, e^{jk\theta}, \quad c_k \in \mathbb{C} \right\}.$$

and the constant

$$C_1^{\mathcal{P}_N}(N_1) := \max_{\|f\|_{\frac{2\pi}{N_1},\infty} \le 1, f \in \mathcal{P}_N} \|f\|_\infty .$$

By Lemma 2 we have $C_1^{\mathcal{P}_N}(N_1) = C_1^{\mathcal{T}_N}(N_1)$. Finally, we can introduce the signal space covering the OFDM signals

$$\mathcal{P}_N^+ := \left\{ f(\theta) = \sum_{k=0}^{N-1} c_k\, e^{jk\theta}, \quad c_k \in \mathbb{C} \right\}$$

and consider

$$C_1^{\mathcal{P}_N^+}(N_1) := \max_{\|f\|_{\frac{2\pi}{N_1},\infty} \le 1, f \in \mathcal{P}_N^+} \|f\|_\infty .$$

Clearly, we have $C_1^{\mathcal{P}_N^+}(LN) \le \frac{1}{\cos\left(\frac{\pi}{2L}\right)}$, where oversampling is defined as $L = \frac{N_1}{N}$ for this class (we use the same notation for L). On the other hand the result can be improved and, this time, the extremal signals depend on N.

Corollary 5 *Let $f \in \mathcal{P}_N^+$ and $L \ge 1, LN \in \mathbb{N}$, be the oversampling factor. Then, the following holds:*

Let N be odd. *If $T \in \mathcal{T}_{\frac{N-1}{2}}$ is $\mathcal{T}$-$\left(\frac{N-1}{2}, LN\right)$-extremal then $P_N(\theta) := e^{j\frac{N-1}{2}} T(\theta)$ is $\mathcal{P}^+$-(N, LN)-extremal and consequently $C_1^{\mathcal{P}_N^+}(LN) = C_1^{\mathcal{T}_{\frac{N-1}{2}}}(LN)$. Moreover if $\frac{LN}{N-1} \in \mathbb{N}$ then*

$$T(\theta) = \frac{\cos\left[\frac{N-1}{2}\left(\theta - \frac{\pi}{LN}\right)\right]}{\cos\left[\frac{\pi(N-1)}{2LN}\right]}.$$

If N is even, *similarly, if $\frac{LN}{N-1} \in \mathbb{N}$ then $P_N(\theta) = e^{j\frac{N-1}{2}}T(\theta)$ is $\mathcal{P}^+$-(N, LN)-extremal. In both cases $C_1^{\mathcal{P}_N^+}(LN)$ is lower- and upperbounded by*

$$\begin{aligned} C_1^{\mathcal{T}_{\frac{N}{2}-1}}(LN) &\leq C_1^{\mathcal{P}_N^+}(LN) \\ &\leq C_1^{\mathcal{T}_{N-1}}(2LN) \\ &\leq \frac{1}{\cos\left(\frac{\pi(N-1)}{2LN}\right)}. \end{aligned}$$

Proof. Consider first the case N *odd* and suppose that P_N is $\mathcal{P}^+$-(N, LN)-extremal. Then after proper scaling $\|P_N\|_{\frac{2\pi}{LN},\infty} \leq 1$ and $P_N(\theta_0) = \|P_N\|_\infty = C_1^{\mathcal{P}_N^+}(LN)$ for some time instance θ_0. Appealing to Lemma 2 define the trigonometric polynomial

$$T_{\frac{N-1}{2}}(\theta) := \Re\left(e^{-j\arg\ e^{-j\frac{N-1}{2}\theta_0}P_N(\theta_0)}\ e^{-j\frac{N-1}{2}\theta}P_N(\theta)\right),$$

then $\left\|T_{\frac{N-1}{2}}\right\|_{LN,\infty} \leq 1$ and $C_1^{\mathcal{P}_N^+}(LN) = P_N(\theta_0) = T_{\frac{N-1}{2}}(\theta_0) \leq C_1^{\mathcal{T}_{\frac{N-1}{2}}}(LN)$ and with $C_1^{\mathcal{P}_N^+}(LN) \geq C_1^{\mathcal{T}_{\frac{N-1}{2}}}(LN)$ we have $C_1^{\mathcal{P}_N^+}(LN) = C_1^{\mathcal{T}_{\frac{N-1}{2}}}(LN) \leq \frac{1}{\cos\left(\frac{\pi(N-1)}{2NL}\right)}$, i.e. if $T_{\frac{N-1}{2}}$ is $\mathcal{T}$-$\left(\frac{N-1}{2}, LN\right)$-extremal then $P_N(\theta) := e^{j\frac{N-1}{2}\theta}T_{\frac{N-1}{2}}(\theta)$ is $\mathcal{P}^+$-(N, LN)-extremal. Furthermore, if $LN = m(N-1)$ for some $m \in \mathbb{N}, m > 1$ then

$$P_N(\theta) = e^{j\frac{N-1}{2}\theta}\frac{\cos\left[\frac{N-1}{2}\left(\theta - \frac{\pi}{LN}\right)\right]}{\cos\left[\frac{\pi(N-1)}{2LN}\right]}. \tag{3.10}$$

Let N be *even* and assume again that P_N is $\mathcal{P}^+$-(N, LN)-extremal. Setting this time

$$T_{N-1}(\theta) := \Re\left(e^{-j\arg\ e^{-j\frac{N-1}{2}\theta_0}P_N(\theta_0)}\ e^{-j\frac{N-1}{2}\theta}P_N(\theta)\right),$$

where $T_{N-1}(\theta)$ is a degree $N-1$ trigonometric polynomial with the properties $T(\theta) = T(\theta + 4\pi)$, $P_N(\theta_0) = T_{N-1}(\theta_0)$ and $\left|T_{N-1}\left(\frac{4\pi l}{2LN}\right)\right| = \left|T_{N-1}\left(\frac{2\pi l}{LN}\right)\right| \leq \left|P_N\left(\frac{2\pi l}{LN}\right)\right| \leq 1, l = 0, \ldots, 2LN-1$. Consequently

$$\begin{aligned} |T(\theta_0)| &\leq C_1^{\mathcal{T}_{N-1}}(2LN) \max_{0\leq l<2LN}\left|T_{N-1}\left(\frac{2\pi l}{LN}\right)\right| \\ &\leq C_1^{\mathcal{T}_{N-1}}(2LN) \max_{0\leq l<LN}\left|P_N\left(\frac{2\pi l}{LN}\right)\right| \\ &\leq C_1^{\mathcal{T}_{N-1}}(2LN) \end{aligned}$$

and $C_1^{\mathcal{P}_N^+}(LN) \leq C_1^{\mathcal{T}_{N-1}}(2LN) \leq \frac{1}{\cos\left(\frac{\pi(N-1)}{2NL}\right)}$. The bound is tight if $LN = m(N-1)$ for some $m \in \mathbb{N}$ and the extremals are given by (3.10). On the other hand for obvious reasons if T_{N-1} is $\mathcal{T}$-$(N-1, 2LN)$ then not necessarily $P_N(\theta) := e^{j\frac{N-1}{2}} T_{N-1}(\theta)$ $\mathcal{P}^+$-(N, LN)-extremal because this will generally require all coefficients of T_{N-1} to be non-zero, i.e. $P_N \notin \mathcal{P}_N$. A lower bound is given by $C_1^{\mathcal{P}_N^+}(LN) \geq C_1^{\mathcal{T}_{\frac{N}{2}-1}}(LN)$. ■

We can conclude that for the case N odd, $C_1^{\mathcal{P}_N^+}(LN)$ and the corresponding polynomials can be numerically obtained. For the case N even, which is the practical case, good lower and upper bounds can be given. In both cases, the extremal signals are analytically known for certain oversampling rates.

3.2.2 Band-limited signals

Considering the proof technique in [3] that cannot be applied to general band-limited signals it may be surprising that the same estimation (i.e. $C_1(L) \leq \frac{1}{\cos\left(\frac{\pi}{2L}\right)}$) holds for the space $\mathcal{PW}_B^\infty$. This statement is the content of the next theorem.

Theorem 6 *Let $L > 1$. We have*

$$C_1(L) = f_L(\theta_L) \leq \frac{1}{\cos\left(\frac{\pi}{2L}\right)}.$$

Moreover, if $L \in \mathbb{N}$, then $\theta_L = \frac{1}{2L}$, $f_L(\theta) = \frac{\cos\left[\pi\left(\theta - \frac{1}{2L}\right)\right]}{\cos\left[\frac{\pi}{2L}\right]}$ and $C_1(L) = \frac{1}{\cos\left(\frac{\pi}{2L}\right)}$.

Proof. Without loss of generality suppose $B = \pi$ and $0 < \theta_L \leq \frac{1}{2L}$. It is sufficient to show that $C_1^{\mathbb{R}}(L) \leq \frac{1}{\cos\left(\frac{\pi}{2L}\right)}$. Assuming f_L to be real where f_L is such that $f_L(\theta_L) = C_1^{\mathbb{R}}(L)$ and setting $F(\theta) := f_L(\theta + \theta_L) - C_1(L)\cos(\pi\theta)$ it follows that $F(0) = 0$ and $F'(0) = 0$ since F has a local extremum at $\theta = \theta_L$. The samples of F are

$$F(l) = f_L(l) - C_1(L)(-1)^l \geq 0, \quad l = 2n+1, n \in \mathbb{Z},$$

and

$$F(l) \leq 0, \quad l = 2n, n \in \mathbb{Z}\backslash\{0\}.$$

Consequently, the samples can be rewritten as

$$F(l) = C'_l (-1)^{l+1}, \quad C'_l \geq 0, l \in \mathbb{Z} \setminus \{0\}.$$

We will now prove that $F(\theta) \geq 0$ for $\theta \in (-1,1)$. Applying (3.2) to $F \in \mathcal{PW}_\pi^\infty$ yields

$$\begin{aligned}
F(\theta) &= \underbrace{F'(0)\frac{\sin(\pi\theta)}{\pi}}_{=0} + \underbrace{F(0)\frac{\sin(\pi\theta)}{\pi\theta}}_{=0} + \theta \sum_{l=-\infty, l\neq 0}^{\infty} \frac{F(l)}{l} \frac{\sin[\pi(\theta-l)]}{\pi(\theta-l)} \\
&= \theta \sum_{l=-\infty, l\neq 0}^{\infty} \frac{F(l)}{l} \frac{\sin[\pi(\theta-l)]}{\pi(\theta-l)} \\
&= \theta \frac{\sin(\pi\theta)}{\pi} \sum_{l=-\infty, l\neq 0}^{\infty} \frac{F(l)}{l} \frac{(-1)^l}{(\theta-l)} \\
&= \theta \frac{\sin(\pi\theta)}{\pi} \sum_{l=1}^{\infty} \frac{F(l)}{l} \frac{(-1)^l}{(\theta-l)} + \theta \frac{\sin(\pi\theta)}{\pi} \sum_{l=-\infty}^{-1} \frac{F(l)}{l} \frac{(-1)^l}{(\theta-l)}.
\end{aligned}$$

The first sum on the right hand side is

$$\begin{aligned}
\sum_{l=1}^{\infty} \frac{F(l)}{l} \frac{(-1)^l}{(\theta-l)} &= \sum_{l=1}^{\infty} \frac{C'_l (-1)^{l+1}}{l} \frac{(-1)^l}{(\theta-l)} \\
&= \sum_{l=1}^{\infty} \frac{C'_l}{l(-\theta+l)} \geq 0, \quad \theta \in (-1,1),
\end{aligned}$$

and the second term is

$$\begin{aligned}
\sum_{l=-\infty}^{-1} \frac{F(l)}{l} \frac{(-1)^l}{(\theta-l)} &= -\sum_{l=-\infty}^{-1} \frac{C'_l}{l(\theta-l)} \\
&= \sum_{l=1}^{\infty} \frac{C'_l}{l(\theta+l)} \geq 0, \quad \theta \in (-1,1).
\end{aligned}$$

Moreover since $\theta \frac{\sin(\pi\theta)}{\pi} \geq 0, \theta \in (-1,1)$, finally

$$F(\theta) \geq 0, \quad \theta \in (-1,1).$$

Equality is achieved for some $\theta \in (-1,1) \setminus \{0\}$ if and only if $C'_l = 0, l \in \mathbb{Z} \setminus \{0\}$, i.e. $f_L(\theta+\theta_L) = C_1(L)\cos(\pi\theta), \theta \in (-1,1)$. Otherwise $f_L(\theta+\theta_L) > C_1(L)\cos(\pi\theta), \theta \in (-1,1)$ and since $f_L(\theta_L - \theta_0) = 1$ due to $\|f_L\|_{\frac{1}{L},\infty} \leq 1$ for some $0 < \theta_0 \leq \theta_L$ we have

$$C_1(L) \leq \frac{1}{\cos(\pi\theta_0)} \leq \frac{1}{\cos(\pi\theta_L)} = \frac{1}{\cos\left(\frac{\pi}{2L}\right)}$$

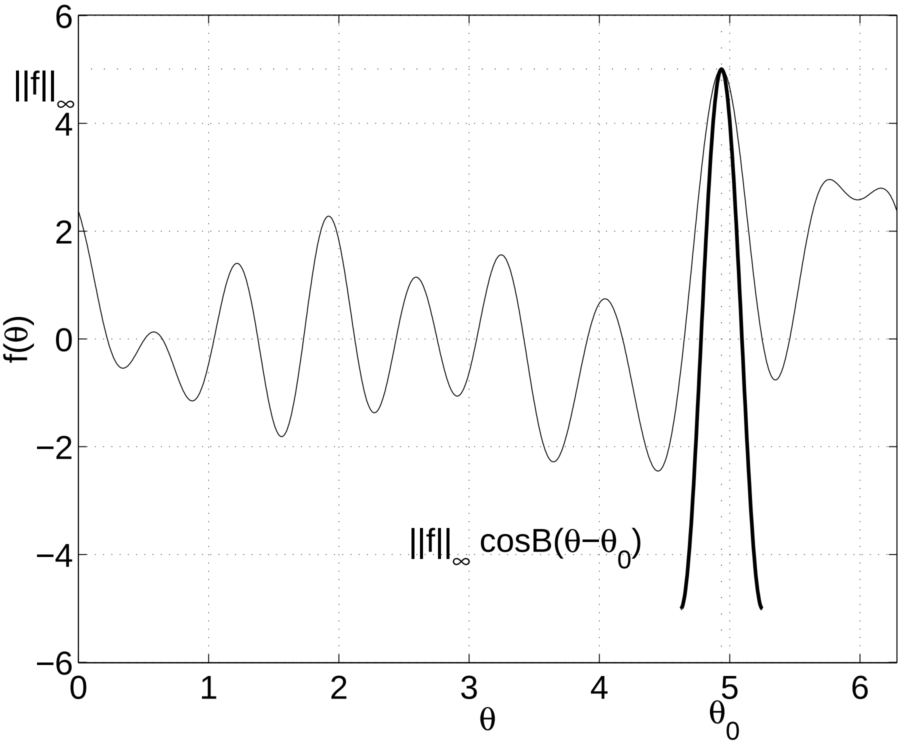

Figure 3.2: Local behavior of real band-limited signals.

which completes the proof. ■

Note that the first part of the proof of Theorem 6 gives a local characterization of a band-limited signal in the neighborhood of an extremum. Indeed, let f be real-valued, band-limited with B and $\|f\|_\infty = M$ and θ_0 a time instance such that $f(\theta_0) = M$, Theorem 6 states that (after some proper denormalization) for $\theta_0 - \frac{\pi}{B} \leq \theta \leq \theta_0 + \frac{\pi}{B}, \theta \neq \theta_0$, the signal f must be strictly greater than $M \cos[B(\theta - \theta_0)]$ with equality only if $f(\theta) = M \cos[B(\theta - \theta_0)]$. This can also be interpreted in the manner that the signal that decreases the fastest around its maximum is $M \cos[B(\theta - \theta_0)]$. An example is given in Fig. 3.2. We can even extend this analysis to complex band-limited signals so that we can characterize the magnitude. Let f be complex and band-limited with $\|f\|_\infty = M$. Then $f' := |f(\theta)|^2$ is real and band-limited with twice the bandwidth of f. Thus, Theorem 6 can be applied to f' and hence $f'(\theta) > M^2 \cos[2B(\theta - \theta_0)]$ and consequently $|f| > M\sqrt{\max\{\cos[2B(\theta - \theta_0)], 0\}}, \theta_0 - \frac{\pi}{2B} \leq \theta \leq \theta_0 + \frac{\pi}{2B}, \theta \neq \theta_0$. On the other hand, equality can never be achieved, since according to Theorem 6 the signal is given by $f(\theta) = M^2 \cos[2B(\theta - \theta_0)]$ which is not always positive.

Theorem 6 can be improved by considering positive signals.

Lemma 7 *Let real f be in $\mathcal{PW}_\pi^\infty$ and $f(\theta) \geq 0, \theta \in \mathbb{R}$. Further $\|f_L\|_{\frac{1}{L},\infty} \leq 1$ then*

$$0 \leq f(\theta) \leq \frac{1}{\cos^2\left(\frac{\pi}{4L}\right)}$$

Proof. Appealing to a theorem of Krein [67] f can be represented by $f(\theta) = |f'(\theta)|^2$ where $f' \in \mathcal{PW}_{\frac{\pi}{2}}^\infty$. Thus

$$f'(l) \leq 1,$$

and

$$f(\theta) = |f'(\theta)|^2 \leq \frac{1}{\cos^2\left(\frac{\pi}{4L}\right)}.$$

■

Thus, if this property is known even critical sampling is sufficient to capture the peaks between the samples. Let us now describe the non-ideal situation where the samples are disturbed.

3.3 Noisy samples problem

3.3.1 Trigonometric polynomials

Let us again provide known results. It was shown in [18] that any signal $f \in \mathcal{P}_N$ can be recovered by its samples $f\left(\frac{2\pi l}{N_1}\right), l = 0, \ldots 2N+1, N_1 > 2N$ applying

$$f(\theta) = \frac{1}{N_1} \sum_{l=0}^{N_1-1} f\left(\frac{2\pi l}{N_1}\right) g\left(\theta - \frac{2\pi l}{N_1}\right)$$

where g is taken from the set

$$\mathcal{M}_N^{N_1} := \left\{ g \in \mathcal{T}_{N_2-1}, g(\theta) = 1 + 2\sum_{k=1}^{N} \cos(k\theta) + 2\sum_{k=N+1}^{N_2-1} c_k \cos(k\theta), \quad c_k \in \mathbb{R}, N_1 - N \geq N_2 > N \right\}.$$

As before we define an operator $T_g : C_{2\pi}(\mathbb{R}) \to \mathcal{P}_{N_2}, f \hookrightarrow \frac{1}{N_1}\sum_{l=0}^{N_1-1} f\left(\frac{2\pi l}{N_1}\right) g\left(\theta - \frac{2\pi l}{N_1}\right)$ with norm

$$|T_g| = \max_{\|f\|_\infty \leq 1, f\in C(\mathbb{R})} \|T_g f\|_\infty$$

and ask for the constant

$$C_2^{\mathcal{P}_N}(N_1) := \inf_{g\in\mathcal{M}_N^{N_1}} |T_g|.$$

Note that for $N_1 = 2N+1$ the set $\mathcal{M}_N^{N_1}$ reduces to the well-known Dirichlet kernel

$$D_N(\theta) = 1 + 2\sum_{k=1}^{N}\cos(k\theta) = \frac{\sin\left(\frac{(2N+1)\theta}{2}\right)}{\sin\left(\frac{\theta}{2}\right)}.$$

It is a well-known result [69] that the operator norm $|T_{D_N}|$ is given by

$$\begin{aligned} |T_{D_N}| &= \frac{1}{2N+1}\max_{-\frac{\pi}{2N+1}\leq\theta\leq\frac{\pi}{2N+1}} \sum_{l=0}^{2N+1}\left|\frac{\sin\left[\frac{(2N+1)}{2}\left(\theta - \frac{2\pi l}{2N+1}\right)\right]}{\sin\left[\frac{1}{2}\left(\theta - \frac{2\pi l}{2N+1}\right)\right]}\right| \\ &\leq \frac{2}{\pi}\log(2N+1) + \mathcal{O}(1). \end{aligned}$$

We obtain from [9][2]

$$|T_{D_N}| \leq \frac{2}{\pi}\log(2N+1) + 2 + \frac{2}{\pi}\log(2).$$

Now, in order to bound $C_2^{\mathcal{P}_N}(N_1)$ for $N_1 > 2N+1$ we can apply the class of the so-called de la Vallée-Poussin polynomials [6]

$$S_{N_2,N}(\theta) = \frac{\sin\left(\frac{(N_2+N)\theta}{2}\right)\sin\left(\frac{(N_2-N)\theta}{2}\right)}{(N_2-N)\sin^2\left(\frac{\theta}{2}\right)}, \quad N_2 > N.$$

In [18] a technique was developed to bound these operator norms efficiently. The approach was refined in [6] where $\left|T_{S_{N_2,N}}\right| \leq \sqrt{\frac{N_2+N}{N_2-N}}$ was proved. On the other hand comparing $\left|T_{S_{N_2,N}}\right|$ and $\left|T_{D_{N_2,N}}\right|$ for $N_2 = N+1$ it is observed that $\left|T_{S_{N+1,N}}\right| = |T_{D_N}| = \sqrt{2N+1}$ opposed to $\log(2N+1)$, i.e. it is not appropriate for $N_2 - N$ small. The following result is a generalization of the bound in [9] on the Dirichlet kernel.

[2]This follows from the equality $\left|\sin\frac{1}{2}\left(t - \frac{2\pi k}{N}\right)\right| = \frac{1}{2}\left|\exp jt - \exp\frac{2\pi jk}{N}\right|$ and the bound from Lagrange interpolation.

Theorem 8 *Let* $N_1 > 2N, N_1, N \in \mathbb{N}$. *Then*

$$C_2^{\mathcal{P}_N}(N_1) \leq \min\left\{\frac{2}{\pi}\log\left(\frac{N_1}{N_1-2N}\right)+c_1, \sqrt{\frac{N_1}{N_1-2N}}\right\},$$

where

$$c_1 = 2 + \frac{2}{\pi}\log(2) + \frac{2\cot\left(\frac{\pi}{2(N_1-2N)}\right)}{\pi(N_1-2N)}.$$

Proof. The operator norm $\left|T_{S_{N_2,N}}\right|$ is given by

$$\left|T_{S_{N_2,N}}\right| = \max_{-\frac{\pi}{N_1}\leq\theta\leq\frac{\pi}{N_1}} \frac{1}{N_1}\sum_{l=0}^{N_1-1}\left|\frac{\sin\left[\frac{(N_2+N)}{2}\left(\theta-\frac{2\pi l}{N_1}\right)\right]\sin\left[\frac{(N_2-N)}{2}\left(\theta-\frac{2\pi l}{N_1}\right)\right]}{(N_2-N)\sin^2\left[\frac{1}{2}\left(\theta-\frac{2\pi l}{N_1}\right)\right]}\right|.$$

Writing

$$\left|\sin\left[\frac{1}{2}\left(\theta-\frac{2\pi l}{N_1}\right)\right]\right| = \frac{1}{2}\left|\exp(j\theta)-\exp\left(\frac{2\pi jl}{N_1}\right)\right|$$

it follows from [9] that

$$\frac{1}{2}\left|\exp(j\theta)-\exp\left(\frac{2\pi jl}{N_1}\right)\right| \geq \sin\left(\frac{\pi l}{2N_1}\right), \quad 0 < l < N_1, -\frac{\pi}{N_1}\leq\theta\leq\frac{\pi}{N_1}.$$

Considering

$$\left|\frac{\sin\left(\frac{(N_2-N)\theta}{2}\right)}{(N_2-N)\sin\left(\frac{\theta}{2}\right)}\right| \leq 1$$

and splitting up the sum yields

$$\left|T_{S_{N_2,N}}\right| = 1 + \frac{1}{N_1}\sum_{l=1}^{\left\lfloor\frac{N_1}{N_2-N}\right\rfloor}\frac{1}{\sin\left(\frac{\pi l}{2N_1}\right)} + \frac{1}{N_1}\sum_{l=\left\lfloor\frac{N_1}{N_2-N}\right\rfloor+1}^{N_1-1}\frac{1}{(N_2-N)\sin^2\left(\frac{\pi l}{2N_1}\right)}$$

where $\left\lfloor\frac{N_1}{N_2-N}\right\rfloor$ denotes the greatest integer smaller than $\frac{N_1}{N_2-N}$. The first term on the right hand side is upperbounded by [9]

$$1 + \frac{1}{N_1}\sum_{l=1}^{\left\lfloor\frac{N_1}{N_2-N}\right\rfloor}\frac{1}{\sin\left(\frac{\pi l}{2N_1}\right)} \leq \frac{2}{\pi}\log\left(\frac{N_1}{N_2-N}\right)+2+\frac{2}{\pi}\log(2)$$

and the second term by

$$\begin{aligned}\frac{1}{N_1\left(N_2-N\right)}\sum_{l=\left\lfloor\frac{N_1}{N_2-N}\right\rfloor+1}^{N_1-1}\frac{1}{\sin^2\left(\frac{\pi l}{2N_1}\right)} &\leq \frac{1}{N_1\left(N_2-N\right)}\sum_{l=\left\lfloor\frac{N_1}{N_2-N}\right\rfloor+1}^{N_1-1}\int_{l-1}^{l}\frac{1}{\sin^2\left(\frac{\pi x}{2N_1}\right)}dx\\ &= \frac{1}{N_1\left(N_2-N\right)}\int_{\left\lfloor\frac{N_1}{N_2-N}\right\rfloor}^{N_1-1}\frac{1}{\sin^2\left(\frac{\pi x}{2N_1}\right)}dx\\ &= \frac{1}{N_1\left(N_2-N\right)}\left[\frac{2N_1}{\pi}\cot\left(\frac{\pi x}{2N_1}\right)\right]_{\left\lfloor\frac{N_1}{N_2-N}\right\rfloor}^{N_1-1}\\ &\leq \frac{2\cot\left(\frac{\pi}{2(N_2-N)}\right)}{\pi\left(N_2-N\right)}.\end{aligned}$$

Therefore

$$\left|T_{S_{N_2,N}}\right| \leq \frac{2}{\pi}\log\left(\frac{N_1}{N_2-N}\right)+2+\frac{2}{\pi}\log\left(2\right)+\frac{2\cot\left(\frac{\pi}{2(N_2-N)}\right)}{\pi\left(N_2-N\right)}$$

and with $N_2 = N_1 - N$ the result follows. The second upper bound follows from [6]. ■

Neglecting the logarithmic part which is only a good bound when $N_1 - 2N$ is small the theorem reads in terms of the oversampling factor $L = \frac{N_1}{2N}$

$$C_2^{\mathcal{P}_N}\left(L\right) \leq \sqrt{\frac{L}{L-1}}.$$

Numerical computations of the operator norm for different $N_1, N, N_2 = N_1 - N$ (i.e. different $L = \frac{N_1}{2N}$), were carried out and shown in Fig. 3.3 along with the bounds on $C_1^{\mathcal{T}_N}\left(N_1\right)$ from Theorem 4 and $C_2^{\mathcal{P}_N}\left(N_1\right)$ from Theorem 8 where, for the sake of simplicity in both figures, the curves are depicted over $\mathbb{R}$. It is observed that the upper bound from Theorem 8 is excellent for $L \geq 2$. Recall also that the bound from Theorem 4 is a lower bound on $C_2^{\mathcal{P}_N}\left(L\right)$ for all $L \in \mathbb{N}$. It is also interesting that the bound in Theorem 8 is better for small oversampling rates $(1 < L < 2)$ than the bound from Theorem 6 and is therefore applicable to the peak value problem as well.

An interesting extension of the theorem is the following result. Note that the signals $f \in \mathcal{P}_N$ can be interpreted as a complex-valued function $f\left(z\right) = \sum_{k=-N}^{N} c_k z^k$, $z \in \mathbb{C}$, evaluated on the unit disc. We denote as $A\left(D\right)$ the signals that are analytic in the interior of the unit disc

$D := \{z : |z| \leq 1\}$ and bounded on the boundary. The norm is given by $\|\cdot\|_{A(D)}$ defined as $\|f\|_{A(D)} := \max_{0\leq\theta<2\pi} f\left(e^{j\theta}\right)$. The following theorem derives a lower bound on the operator norm if the kernel belongs to $A(D)$ and the operator itself is the identity operator on P_N^+.

Theorem 9 *Let $g \in A(D)$ be a kernel so that $T_g f = f$ for all $f \in P_N^+$ then*

$$|T_g f| > \frac{\log(N+1)}{\pi}.$$

Proof. Let $g^{(1)}, \ldots, g^{(N_1)} \in A(D)$ be the interpolating signals of the representation so that $T_g f = f$ for all $f \in P_N^+$ and $\lambda_1, \ldots, \lambda_{N_1}$ are complex numbers. According to Rudin-Carleson's theorem [70] there is a signal $f \in A(D)$ with $f(z_l) = \lambda_l, 1 \leq l \leq N_1$, and $\|f\|_{A(D)} = \max_{1\leq l\leq N_1} |\lambda_l|$. Thus

$$\|T_g\| = \max_{|z|=1} \sum_{l=1}^{N_1} \left| g^{(l)}(z) \right| > \sum_{l=1}^{N_1} \frac{1}{2\pi} \int_{-\pi}^{\pi} |g^{(l)}(e^{it})| \, dt \ .$$

We have for $|z| < 1$

$$g^{(l)}(z) = \sum_{m=0}^{\infty} c_m(l) z^m, \quad 1 \leq l \leq N_1.$$

Using Hardy's inequality [70]

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |g^{(l)}(e^{it})| \, dt \geq \frac{1}{\pi} \sum_{m=1}^{\infty} \frac{1}{m} |c_m(k)|$$

and

$$\|T_g\| \geq \frac{1}{\pi} \sum_{m=1}^{\infty} \frac{1}{m} \left(\sum_{l=1}^{N_1} |c_m(l)| \right) \ .$$

For $0 \leq r \leq N$ we have for $|z| < 1$ with $f_r(e^{it}) = e^{irt}$

$$f_r(z) = \sum_{l=1}^{N_1} f_r(z_l) g^{(l)}(z) = \sum_{m=1}^{\infty} z^m \left(\sum_{l=1}^{N_1} f_r(z_l) c_m(l) \right) = z^r$$

and for $0 \leq r \leq N$

$$\sum_{l=1}^{N_1} f_r(z_l) c_r(l) = 1,$$

or

$$\sum_{l=1}^{N_1} |c_r(l)| \geq 1.$$

Combining

$$\|T\| \geq \frac{1}{\pi}\sum_{m=1}^{N}\frac{1}{m}\left(\sum_{l=1}^{N_1}|c_m(l)|\right) \geq \frac{1}{\pi}\sum_{m=1}^{N}\frac{1}{m} > \frac{1}{\pi}\log(N+1)\ ,$$

completes the proof. ∎

The theorem says that if the operator connected with a (possibly shifted version) kernel that has the reconstruction property for all $f \in P_N^+$ then there is an input so that the input is amplified with $\frac{\log(N+1)}{\pi}$. Thus, even if the inputs are chosen carefully a small disturbance can have a catastrophic effect. The theorem has some interesting conclusions for practical coding schemes as we point out in the last chapter.

3.3.2 Band-limited signals

Concerning results on $\mathcal{PW}_B^\infty$ in this regard we mention the approaches in [16] and [14]. On the other hand using the properties of a sampling series a better result can be obtained. The following theorem improves the result given by [16, Theorem 7.2.5].

Theorem 10 *Let $L > 1$. Then*

$$C_2(L) \leq \min\left\{\frac{2}{\pi}\log\left(\frac{2L}{L-1}\right) + c_2, \sqrt{\frac{L+1}{L-1}}\right\}.$$

where

$$c_2 = \frac{4}{\pi} + \frac{8}{\pi^2} + \frac{L+1}{2L}.$$

Proof. Using the kernel (3.5) the constant $C_2(L)$ is upperbounded by

$$\begin{aligned} C_2(L) &\leq \sup_{-\frac{1}{2L}\leq\theta\leq\frac{1}{2L}} \frac{1}{L}\sum_{l=-\infty}^{\infty}\left|S_L\left(\theta-\frac{l}{L}\right)\right| \\ &= \sup_{-\frac{1}{2L}\leq\theta\leq\frac{1}{2L}} \frac{1}{L}\sum_{l=-\infty}^{\infty}\left|\frac{2\sin\left[\frac{\pi(L+1)}{2}\left(\theta-\frac{\pi l}{L\pi}\right)\right]\sin\left[\frac{\pi(L-1)}{2}\left(\theta-\frac{\pi l}{L\pi}\right)\right]}{\pi(L-1)\pi\left(\theta-\frac{\pi l}{L\pi}\right)^2}\right|. \end{aligned}$$

Let us first prove the more important second inequality. Applying Cauchy-Schwartz's inequality yields

$$C_2(L) \leq \sup_{-\frac{1}{2L}\leq\theta\leq\frac{1}{2L}} \frac{1}{L}\sqrt{\sum_{l=-\infty}^{\infty}\frac{2\sin^2\left[\frac{\pi(L+1)}{2}\left(\theta-\frac{l}{L}\right)\right]}{\pi(L-1)\pi\left(\theta-\frac{l}{L}\right)^2}\sum_{l=-\infty}^{\infty}\frac{2\sin^2\left[\frac{\pi(L-1)}{2}\left(\theta-\frac{l}{L}\right)\right]}{\pi(L-1)\pi\left(\theta-\frac{l}{L}\right)^2}}.$$

Recalling that for all $f \in \mathcal{PW}_\pi^2$

$$\sum_{l=-\infty}^{\infty} |f(l)|^2 = \int_{-\infty}^{\infty} |f(\theta)|^2 \, d\theta = \frac{1}{2\pi} \int_{-\infty}^{\infty} |f(\omega)|^2 \, d\omega$$

results in

$$\begin{aligned}\sum_{l=-\infty}^{\infty} \frac{2\sin^2\left[\frac{\pi(L+1)}{2}\left(\theta - \frac{\pi l}{L\pi}\right)\right]}{\pi(L-1)\pi\left(\theta - \frac{\pi l}{L\pi}\right)^2} &= L\int_{-\infty}^{\infty} \frac{2\sin^2\left(\frac{\pi(L+1)\theta}{2}\right)}{\pi(L-1)\pi\theta^2} d\theta \\ &= \frac{L}{\pi}\frac{1}{L-1}\int_{-\frac{(L+1)}{2}\pi}^{\frac{(L+1)}{2}\pi} 1 \, d\omega \\ &= L\frac{L+1}{L-1}\end{aligned}$$

and

$$\begin{aligned}\sum_{l=-\infty}^{\infty} \frac{2\sin^2\left[\frac{\pi(L-1)}{2}\left(\theta - \frac{\pi l}{L\pi}\right)\right]}{\pi(L-1)\pi\left(\theta - \frac{\pi l}{L\pi}\right)^2} &= \frac{L}{\pi}\frac{1}{L-1}\int_{-\frac{(L-1)}{2}\pi}^{\frac{(L-1)}{2}\pi} 1 \, d\omega \\ &= L\frac{L-1}{L-1} \\ &= L.\end{aligned}$$

The first inequality can be proved along the lines of Theorem 4. Using

$$\left|\frac{2\sin\left(\frac{\pi(L-1)\theta}{2}\right)}{(L-1)\pi\theta}\right| \leq 1$$

and splitting up the sum yields

$$\begin{aligned}C_2(L) \leq |S_L(0)| &+ \sup_{-\frac{1}{2L}\leq\theta\leq\frac{1}{2L}} \frac{2}{L}\sum_{l=1}^{\lceil\frac{L}{L-1}\rceil} \left|S_L\left(\theta - \frac{l}{L}\right)\right| \\ &+ \sup_{-\frac{1}{2L}\leq\theta\leq\frac{1}{2L}} \frac{2}{L}\sum_{l=\lceil\frac{L}{L-1}\rceil+1}^{\infty} \left|S_L\left(\theta - \frac{l}{L}\right)\right|\end{aligned}$$

$$\begin{aligned}
&\leq \frac{L+1}{2L} + \sup_{-\frac{1}{2L}\leq\theta\leq\frac{1}{2L}} \frac{2}{L} \sum_{l=1}^{\lceil \frac{L}{L-1} \rceil} \left| \frac{1}{\pi\left(\theta - \frac{l}{L}\right)} \right| \\
&\qquad\qquad + \sup_{\theta\in\left[-\frac{1}{2L},\frac{1}{2L}\right]} \frac{4}{L} \sum_{l=\lceil \frac{L}{L-1} \rceil+1}^{\infty} \left| \frac{1}{\pi\left(L-1\right)\pi\left(\theta-\frac{l}{L}\right)^2} \right| \\
&= \frac{L+1}{2L} + \frac{2}{\pi} \sum_{l=1}^{\lceil \frac{L}{L-1} \rceil} \frac{1}{\left(l-\frac{1}{2}\right)} + \frac{4L}{\left(L-1\right)\pi^2} \sum_{l=\lceil \frac{L}{L-1} \rceil+1}^{\infty} \frac{1}{\left(l-\frac{1}{2}\right)^2} \\
&\leq \frac{L+1}{2L} + \frac{4}{\pi} + \frac{2}{\pi}\log\left\lceil \frac{L}{L-1} \right\rceil + \frac{4L}{\left(L-1\right)\pi^2} \frac{1}{\lceil \frac{L}{L-1} \rceil - \frac{1}{2}} \\
&\leq \frac{L+1}{2L} + \frac{4}{\pi} + \frac{2}{\pi}\log\left(\frac{2L}{L-1}\right) + \frac{8}{\pi^2}
\end{aligned}$$

where $\lceil \frac{L}{L-1} \rceil$ denotes the smallest integer greater than $\frac{L}{L-1}$. ■

The bound on $C_2(L)$ from Theorem 10 is shown and compared to the other bounds in Fig. 3.3 without the logarithmic part. Again it can be observed that the bound is better for small L than the bound from Theorem 6. We note that in [17] a similar problem regarding discrete-time signals was considered. Moreover, it was shown that the norms of a whole class of filters can be estimated using the trapezoidal filter [17, Prop. 5]. We have not attempted such an analysis.

3.4 Bounding the crest-factor

Since in OFDM systems the signal is digitally generated and processed, the peaks have to be determined on an oversampling grid. Some authors proposed calculating the peaks from Nyquist-rate sampling. However, it was shown that in general peaks of order $\log(N)$ can occur. Using Theorem 4 and 2 this can now be done to any desired accuracy. Defining the constant

$$C_L := \frac{1}{\cos\left(\frac{\pi}{2L}\right)},$$

where $L > 1, L \in \mathbb{N}$, it follows

$$\left| \max_{0\leq\theta<2\pi} |s_c(\theta)| - \max_{0\leq l<LN} \left| s_c\left(\frac{2\pi l}{LN}\right) \right| \right| \leq N\left(C_L - 1\right). \tag{3.11}$$

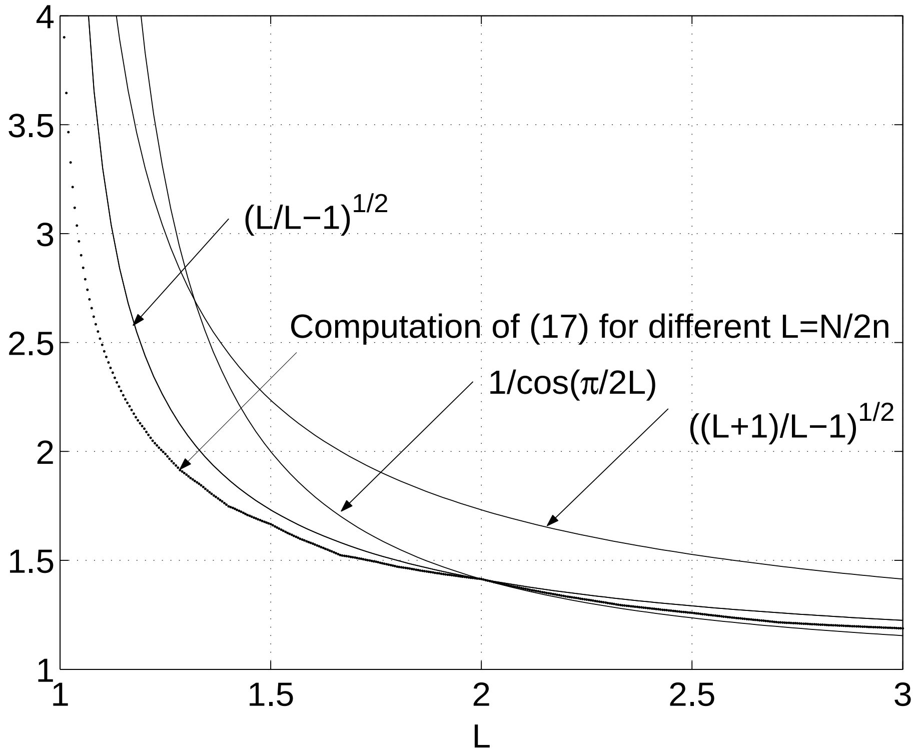

Figure 3.3: Evaluation of operator norm and bounds for different oversampling rates compared to the lower bound.

Thus, when fixing the number of subcarriers N and defining an absolute accuracy $\varepsilon > 0$ the necessary oversampling factor that achieves this accuracy can be calculated. Some examples appear in [47]. Moreover, we point out that we have presented only the simplest OFDM signal model and that practical systems use pulse shaping and consequently the general band-limited framework that is developed here must be invoked in order to obtain error estimates as in (3.11). We already pointed out that the peaks can be locally calculated. Moreover if one guarantees that the signal is always non-negative Lemma 7 shows that even critical sampling is sufficient.

Let us define an important special case where $N = 3$, i.e. $s_c^{(3)}(\alpha) = c_0 + c_1 e^{j\alpha} + c_2 e^{j\alpha}$. Defining $\alpha_{l,K} := \frac{2\pi l}{K}, K > 2, K \in \mathbb{N}$, and

$$C_K := \begin{cases} \frac{1}{\cos\left(\frac{\pi}{2L}\right)} & K > 3 \text{ even} \\ \frac{3-\cos\left(\frac{\pi}{K}\right)}{1+\cos\left(\frac{\pi}{K}\right)} & K \geq 3 \text{ odd} \end{cases}$$

it was shown in [6] that

$$\left| \max_{0 \leq \alpha < 2\pi} \left| s_c^{(3)} (\alpha) \right| - \max_{0 \leq l < K} \left| s_c^{(3)} (\alpha_{l,K}) \right| \right| \leq 3 \left(C_K - 1 \right).$$

The same relation will also be used to estimate the statistical distribution of the CF in Chapter 4 and to bound the CF of codes in Chapter 5. We further note that the technique developed in Theorem 10 can also be used in this context to bound the peak-value of Nyquist filters [58] even without oversampling. An example how to do this is given in [71] in a different context where it was shown that the peak value of the raised cosine kernel is bounded by $\frac{1.2337}{r}$ where r is the roll-off parameter.

The results also suggest consequences for the design of trigonometric polynomials with low CF. Trigonometric polynomials with low CF are effectively being used in radar, sonar, and satellite communications or as test signals for channel identification [72] (in this context they appear as multitone signals). A multitone signal is given by

$$m_N(t) = \sum_{k=1}^{N-1} a_k \cos\left(kt + \phi_k\right),$$

where $a_k, \phi_k, k = 0, \ldots, n-1$ are the real amplitudes and phases. In other notation we also refer to the set $\left\{a_0 e^{j\phi_0}, a_1 e^{j\phi_1}, \ldots, a_{N-1} e^{j\phi_{N-1}}\right\}$ as a (complex-valued) sequence. Fixing the amplitudes the problem is to find appropriate phases so that the CF of the corresponding polynomial is low. For practical purposes it is also often desirable that the phases belong to the set $\{-\pi, \pi\}$.

The classic reference in this regard is that of [73] where an explicit formula for the phases is given by making use of Woodward's theorem on frequency modulated signals. The formula can be applied to arbitrary amplitudes. It is even suited to binary phases. In [72] the Newman phases as well as Rudin-Shapiro sequences were introduced that have CF lower than 3 [dB] [72]. This work was generalized in [74] recognizing that the Rudin-Shapiro sequences belong to the wider set of complementary sequences. A pair of sequences is said to be complementary if their aperiodic autocorrelation sequence adds up to a constant. In [75] good phases that are different from the ones mentioned were given.

The design problem is aggravated if the signals are required to have additional properties for example a small modulation factor that is the ratio of maximum to minimum magnitude. A fundamental result is due to [76] proving that the modulation factor can be made arbitrarily small with increasing N. In [77, 78] a numerical procedure was proposed based on a time-frequency swapping algorithm. This algorithm can be implemented using the DFT. Thus, equation (3.11) can be used for efficient implementation of this algorithm.

The problem can be generalized to find a set of orthogonal signals that have low CF. This problem occurs in OFDM systems employing multiple antennas and pilot-based channel estimation schemes. In [79] a pilot scheme was proposed where the individual pilot signals are obtained by rotating a single sequence that has perfect periodic autocorrelation and low CF. A possible choice for this sequence is a sequence with Newman phases. Other signals can be generated using the time-frequency swapping algorithm. We would also like to mention cross-connections of this topic to sequence design in CDMA context where other performance measures such as the merit-factor are considered [80] which are in some form related to the CF.

3.5 Open problems

We investigated the relationship between the peak value of band-limited signals and the peak value of the samples when oversampling is employed. Such relations are important for efficient implementation of signal processing algorithms in communications systems. We sub-divided the problem into a problem where the samples are disturbed and a ideal situation where they are not. It turned out that with moderate oversampling the peak value between the samples can be effectively bounded which is not possible if the samples are obtained from sampling at Nyquist rate. We have further seen that the bound cannot be improved for practical oversampling values but fails to be tight for oversampling rates below 2. This problem should be further investigated to give better bounds in this region. It is also interesting to extend the analysis done here to the information signals where the coefficients are restricted to belonging to a certain, finite set (i.e. represent a modulation scheme). This would have some interesting implications for the code design.

Research Problem 11 *Compute the good upper bound when the coefficients belong to a finite* *set.*

In a different analysis we examined the impact of noisy samples. We have shown that the noise conveyed to the time-continuous signals can be drastically reduced using oversampling and an appropriate reconstruction filter. We have given an example and provided a good upper bound using new techniques. We have not shown that this is the optimal filter and leave this as an open question. It may also be interesting to give real designs for these filters. Comparing further the results of Theorem 10 and Theorem 8 to Theorem 4 and Theorem 6 shows that apart from applications to filter design the approach is also interesting for the peak value problem for small oversampling rates.

In light of the preceding results we are now able to compute the CF of a single OFDM symbol from its samples. However, as we pointed out we are also interested in the statistical behavior of the CF which is discussed next.

Chapter 4

The distribution of the crest-factor

In order to characterize the impact of non-linear components we introduced the distribution of the CF in Chapter 3. Unfortunately, the exact evaluation is quite difficult and by now still only possible by computing the CF of all signals. Of course, this is impossible for larger numbers of subcarriers. Various papers deal with the problem and propose several approximations relying on the fact that the OFDM signal can be considered as a Gaussian random process [29, 30, 81]. We call this the Gaussian approximation. These approximations are quite useful in practice so as to obtain analytic expressions, but, inherently the scope of this approach is quite unclear in terms of the system parameters and are not suitable for obtaining accurate results in all regions of interest. Indeed, for any fixed N the Gaussian approximation differs from any finite modulation scheme at some point since there is a certain probability of exceeding any finite value. The approach also lacks theoretical justification in the sense that no results exist that the Gaussian approximation becomes better with growing N. Thus, the approach must always be reconfirmed by Monte-Carlo simulations, which are time-consuming in the low probability region, where the most significant differences between the different modulation schemes occur (a quick sketch can be obtained in the higher probability region, though). Moreover, observe that the approximations in [29, 30, 81] only apply to M-ary QAM constellations but not to BPSK (or more generally M-ary PSK constellations). An attempt to describe the distribution for MPSK modulation has been made in [28] using critical sampling and a Gaussian model under a power constraint. However, simulations show that the approach yields poor results in the low probability region.

These arguments actually motivate the derivation of upper bounds on the complementary distribution for the low probability region. In [30] an upper bound has been formally given. However, its computation (apart from the computational effort) involves the computation of Fourier coefficients that can change their signs and the number of coefficients that have to be considered to get accurate results is not clear. Some limit results ($N \to \infty$) can be obtained from [24, 25] for real data. Unfortunately these results do not give any bounds for practical systems. Therefore true upper bounds depending on the different modulation schemes employed in practice are given in this chapter. For example, a straightforward application of our upper bounds would yield an expression of the ,,effective" CF defined in [30, Eqn. (4)]. Another important example is the application for bounding the SER evoked by clipping which we treat in Sec. 4.3. We show that the SER behavior is governed by the CF distribution in the low probability region. The information-theoretic value of these bounds is also obvious so as to provide fundamental limits on the achievable information rate for a constellation carrying M Bits per subcarrier a peak power constraint is imposed on the transmission. Given the distribution F for a given M-ary PSK/M-ary QAM scheme, the achievable information rate is then given as

$$R = \frac{\log_2 \left[1 - F\left(\mathrm{CR}\right)\right]}{N} + \log_2\left(M\right).$$

The achievable rate gives valuable information about the number of subcarriers that are needed to bound the CF within a certain range. Since all redundancy-based CF reduction schemes can be understood as some form of coding the quantity is an upper bound on the number of redundant bits. We can also give more elaborate answers to the questions raised in [82].

The upper bounds are derived for BPSK, M-ary QAM and M-ary PSK constellations. Their scope is confined to the lower probability region and to subcarrier numbers that are not too small. The mathematical thinking is to bound the complementary distribution of the instantaneous magnitude of the OFDM baseband signal and then benefit from an analysis of the relation between the peak value of the discrete-time signal and the peak value of the continuous-time signal. A major advantage is that the bounds are relatively simple to compute and can be easily extended to arbitrary cross-constellations and even codes. A slightly higher computational effort yields even better bounds for lower numbers of subcarriers. In the next step we provide a complete characterization of the distribution for a large number of subcarriers for all practical modulation schemes. We further extend our analysis to dependent subcarriers as for example occurring in coded systems in terms of the distance distribution. Literature seems to be scarce in this field and the only reference we have found in this area is the work in [81] where it is argued that for certain codes (for example convolutional codes) the Gaussian approximation may still be applied. However, for such approximations the same argumentation holds as in the uncoded case. Again, using the bound an effective CF can be defined.

4.1 Independent subcarriers

As we pointed out approximative results for the distribution of the CF assuming independent subcarriers have been derived on the assumptions that the OFDM signal can be modeled as a Gaussian random processes in combination with level-crossing results [29, 30, 81]. In particular, the expression derived in [81] seems to be an appropriate approximation for a wide range. The upper bound given in [30] is slightly better than that of [32]. Here, we take a complete different direction.

4.1.1 The union bound

Theorem 12 *If* $\mathcal{Q} = \mathrm{BPSK}$ *then*

$$F(\lambda) \leq 4B(\lambda) + \min_{L>1,K>2} \sum_{l_1=1}^{\frac{LN}{2}-1} \sum_{l_2=0}^{K-1} e^{-\frac{N\lambda^2}{2\Psi\left(N,\theta_{l_1,L},\alpha_{l_2,K}\right)C_K^2 C_L^2}},$$

where

$$\Psi(N,\theta,\alpha) = \frac{N}{2} + \frac{1}{2}\frac{\sin(N\theta)}{\sin(\theta)}\cos\left[(N-1)\theta + 2\alpha\right]$$

and

$$B(\lambda) = \frac{1}{2^N} \sum_{k=\left\lceil \frac{\lambda\sqrt{N}}{C_L} \right\rceil, k\ even}^{N} \binom{N}{\frac{(N-k)}{2}}.$$

If $\mathcal{Q} = \text{M-QAM}$ *then*

$$F(\lambda) \leq \epsilon_1^{2N}\left(\frac{\lambda}{C_K C_L \sqrt{N}}\sqrt{\frac{6}{M-1}}\right) \min_{L>1,K>2} KLN\, e^{-\frac{\lambda^2}{C_K^2 C_L^2}}$$

where $\epsilon_1(\cdot)$ *is an error term given by (4.2).*

If $\mathcal{Q} = \text{M-PSK}$ *then*

$$F(\lambda) \leq \epsilon_2^{N}\left(\frac{2\lambda}{C_K C_L \sqrt{N}}\right) \min_{L>1,K>2} LNK e^{-\frac{\lambda^2}{C_L^2 C_K^2}}.$$

where $\epsilon_2(\cdot)$ *is an error term given by (4.3).*

Before we prove the theorem let us provide the following lemma that we call the union bound because we estimate the probability of non-distinct subsets by the probability of their union. The lemma will also be interesting for dependent constellations.

Lemma 13 *Let* $\Omega_{L,K} := \{(\theta_{l_1,L}, \alpha_{l_2,K}), 0 \leq l_1 < LN, 0 \leq l_2 < K, L > 1, K > 2\}$ *be a grid of the plane* $[0, 2\pi] \times [0, 2\pi]$ *and suppose that the probabilities*

$$\Pr\left(\Re\left(s_c(\theta)\, e^{j\alpha}\right) > \lambda\right),$$

(θ, α) *running through* $\Omega_{L,K}$*, are given. Then*

$$\Pr(\mathrm{CF}(s_c) > \lambda) \leq \sum_{l_1=0}^{LN-1}\sum_{l_2=0}^{K} \Pr\left(\Re\left(s_c(\theta_{l_1,L})\, e^{j\alpha_{l_2,K}}\right) > \frac{\sigma_s \sqrt{N}\lambda}{C_L C_K}\right)$$

for every L *and* K*. The probabilities can be upperbounded by*

$$\Pr\left(\Re\left(s_c(\theta_{l_1,L})\, e^{j\alpha_{l_2,K}}\right) > \frac{\lambda\sigma_s\sqrt{N}}{C_L C_K}\right) \leq E\left(e^{\varepsilon\ \Re\ s_c(\theta_{l_1,L}) e^{j\alpha_{l_2,K}}\ -\frac{\sigma_s\sqrt{N}\lambda}{C_L C_K}}\right).$$

Proof. The magnitude of s_c can be written as

$$\begin{aligned}
|s_c(\theta)| &= \max_{0\leq\alpha<2\pi} \Re\left(s_c(\theta)\, e^{j\alpha}\right) \\
&= \max_{0\leq\alpha<2\pi} \sum_{k=0}^{N-1} \Re(c_k)\cos(k\theta+\alpha) - \mathrm{Im}(c_k)\sin(k\theta+\alpha) \\
&= \max_{0\leq\alpha<2\pi} \widehat{s}_c(\theta, \alpha)
\end{aligned}$$

where $\widehat{s}_c(\theta,\alpha) := \sum_{k=0}^{N-1} \Re(c_k)\cos(k\theta+\alpha) - \mathrm{Im}(c_k)\sin(k\theta+\alpha)$. Note that $\widehat{s}_c(\theta,\alpha)$ is a degree 1 trigonometric polynomial. Introducing $K > 2, K \in \mathbb{N}$, and referring to Chapter 3 $\widehat{s}_c(\theta,\alpha)$ is upperbounded by

$$\max_{0\le\alpha<2\pi} \widehat{s}_c(\theta,\alpha) \le C_K \cdot \max_{0\le l_2<K} |\widehat{s}_c(\theta,\alpha_{l_2,K})|$$

and thus

$$\begin{aligned}
\Pr(|s_c(\theta)| > \lambda) &= \Pr\left(\max_{0\le\alpha<2\pi} \widehat{s}_c(\theta,\alpha) > \lambda\right) \\
&\le \Pr\left(\max_{0\le l_2<K} |\widehat{s}_c(\theta,\alpha_{l_2,K})| \cdot C_K > \lambda\right) \\
&\le \sum_{l_2=0}^{K-1} \Pr\left(|\widehat{s}_c(\theta,\alpha_{l_2,K})| > \frac{\lambda}{C_K}\right).
\end{aligned}$$

Note that by choosing $K = 4$ the constant C_K equals $\sqrt{2}$. This corresponds to the case that we would estimate the probability that $|s_c(\theta)|$ falls out a circle with radius λ by the probability that $|s_c(\theta)|$ falls out a square with length $\frac{\lambda}{\sqrt{2}}$. Again by oversampling $\theta = \frac{2\pi l_1}{LN}, l_1 = 0,\ldots,LN-1$ with $L > 1, L \in \mathbb{N}$, we obtain

$$\begin{aligned}
\Pr\left(\max_{0\le\theta<2\pi} |s_c(\theta)| > \lambda\right) &\le \Pr\left(C_L \cdot \max_{0\le l_1<LN} |s_c(\theta_{l_1,L})| > \lambda\right) \\
&\le \sum_{l_1=0}^{LN-1} \Pr\left(|s_c(\theta_{l_1,L})| > \frac{\lambda}{C_L}\right) \\
&\le \sum_{l_1=0}^{LN-1} \sum_{l_2=0}^{K-1} \Pr\left(|\widehat{s}_c(\theta_{l_1,L},\alpha_{l_2,K})| > \frac{\lambda}{C_K C_L}\right).
\end{aligned}$$

Let $\varepsilon > 0$ be an arbitrary number. Using the set function

$$I_{\mathcal{A}}(\widehat{s}_c) = \begin{cases} 1 & \widehat{s}_c \in \mathcal{A} \\ 0 & \widehat{s}_c \notin \mathcal{A} \end{cases},$$

where $\mathcal{A}$ denotes any subset of OFDM signals and the Chernoff bounding technique, the number of signals exceeding λ is upperbounded by

$$\Pr(\widehat{s}_c(\theta,\alpha) > \lambda) = E\left(I_{\{\widehat{s}_c(\theta,\alpha)>\lambda\}}(\widehat{s}_c)\right) \le E\left(e^{\varepsilon(\widehat{s}_c(\theta,\alpha)-\lambda)}\right)$$

Since

$$\Pr\left(\mathrm{CF}\left(s_c\right) > \lambda\right) = \Pr\left(\max_{0\leq\theta<2\pi} \left|s_c\left(\theta\right)\right| > \frac{\sigma_s\sqrt{N}\lambda}{C_L C_K}\right)$$

the result follows. ∎

Thus, for the proof of the theorem all we have to do is to bound the distribution of the instantaneous envelope for the different modulation schemes. Generally, we can use numerical methods for computing the instantaneous distribution but due to the large number of subcarriers this will be extensive in computational terms. Analytical methods are also difficult to evaluate since even though the characteristic functions are immediate, the inverse formula is not easy to calculate. Thus, we concentrate on the Chernoff bounding technique next. We need the following lemmas.

Lemma 14 *The following upper bound holds:*

$$\frac{\sinh\left(\frac{\sqrt{M}x}{2}\right)}{\sinh\left(x\right)} \leq \frac{\sqrt{M}}{2} e^{\left(\frac{M}{12}-\frac{1}{3}\right)\frac{x^2}{2}} \epsilon_1\left(x\right). \tag{4.1}$$

Here, $\epsilon_1\left(x\right)$ *is an error term given by*

$$\epsilon_1\left(x\right) := 1 - \frac{\sinh\left(x\right)\left(e^{\frac{Mx^2}{24}} - e^{\left(\frac{M}{12}-\frac{1}{3}\right)\frac{x^2}{2}}\right) - \frac{x^3}{6}}{\sinh\left(x\right) e^{\left(\frac{M}{12}-\frac{1}{3}\right)\frac{x^2}{2}}} \tag{4.2}$$

satisfying $\epsilon_1\left(x\right) \to \left(1 + \mathcal{O}\left(x^4\right)\right)$ *as* $x \to 0$.

Proof. First, we want to find $\gamma > 0$ so that

$$\frac{\sinh\left(\frac{\sqrt{M}x}{2}\right)}{\sinh\left(x\right)} \leq \frac{\sqrt{M}}{2} e^{\frac{\gamma x^2}{2}}.$$

Since $\frac{\sinh\frac{\sqrt{M}x}{2}}{\sinh(x)}$ is symmetric around $x = 0$ it suffices to prove inequality (4.1) for $x \geq 0$. Using $\sinh\left(x\right) \geq x$ we have

$$\frac{\sinh\left(\frac{\sqrt{M}x}{2}\right)}{\sinh\left(x\right)} \leq \frac{\sinh\left(\frac{\sqrt{M}x}{2}\right)}{x}$$

and $\sinh(x) \leq xe^{\frac{x^2}{6}}$ yields the upper bound

$$\frac{\sinh\left(\frac{\sqrt{M}x}{2}\right)}{\sinh(x)} \leq \frac{\frac{\sqrt{M}x}{2}e^{\frac{Mx^2}{24}}}{x} = \frac{\sqrt{M}}{2}e^{\frac{Mx^2}{24}}.$$

Thus $\gamma < \frac{M}{12}$. Since x will be small the inequality will be evaluated at small values and can be improved. Expanding

$$\sinh\left(\frac{\sqrt{M}x}{2}\right) = \frac{\sqrt{M}x}{2} + \frac{\left(\sqrt{M}\right)^3 x^3}{48} + \ldots$$

and

$$e^{\frac{\gamma x^2}{2}} = 1 + \frac{\gamma x^2}{2} + \ldots$$

yields

$$\frac{\sqrt{M}}{2}e^{\frac{\gamma x^2}{2}} \sinh(x) = \frac{\sqrt{M}x}{2} + \left(\frac{\gamma}{2} + \frac{1}{6}\right)\frac{\sqrt{M}x^3}{2} + \ldots$$

Considering the linear and cubic term only yields

$$\frac{\left(\sqrt{M}\right)^3}{48} \leq \left(\frac{\gamma}{2} + \frac{1}{6}\right)\frac{\sqrt{M}}{2}$$

and hence

$$\gamma \geq \frac{M}{12} - \frac{1}{3}.$$

Using the lower bound on γ we have the following inequality chain

$$\sinh\left(\frac{\sqrt{M}x}{2}\right) \leq \frac{\sqrt{M}}{2}\sinh(x)\, e^{\frac{\gamma_1 x^2}{2}} + \epsilon'(x) \leq \frac{\sqrt{M}}{2}\sinh(x)\, e^{\frac{\gamma_2 x^2}{2}}$$

where $\frac{M}{12} - \frac{1}{3} = \gamma_1 \leq \gamma_2 = \frac{M}{12}$ and $\epsilon'(x)$ is an error term. Since the coefficients of the power series are all positive the inequality will hold for each coefficient. Hence, the error will be upperbounded by

$$\epsilon'(x) = \frac{\sqrt{M}}{2}\sinh(x)\, e^{\frac{\gamma_2}{2}x^2} - \frac{\sqrt{M}x}{2} - \left(\frac{\gamma_2}{2} + \frac{1}{6}\right)\frac{\sqrt{M}x^3}{2} - \\ \left(\frac{\sqrt{M}}{2}\sinh(x)\, e^{\frac{\gamma_1}{2}x^2} - \frac{\sqrt{M}x}{2} - \left(\frac{\gamma_1}{2} + \frac{1}{6}\right)\frac{\sqrt{M}x^3}{2}\right).$$

Dividing by $\frac{\sqrt{M}}{2}\sinh(x)$ and $e^{\left(\frac{M}{12}-\frac{1}{3}\right)\frac{x^2}{2}}$ yields

$$\epsilon_1(x) = 1 + \frac{\epsilon'(x)}{\frac{\sqrt{M}}{2}\sinh(x)\, e^{\left(\frac{M}{12}-\frac{1}{3}\right)\frac{x^2}{2}}}.$$

The term $\epsilon'(x)$ satisfies $\epsilon'(x) = \mathcal{O}(x^5)$ as $x \to 0$ and thus $\epsilon_1(x) = 1 + \mathcal{O}\left(\frac{x^5}{x}\right)$ as $x \to 0$. ■

Lemma 15 *The following upper bound holds:*

$$\frac{1}{M}\sum_{m=0}^{M-1} e^{x\cos\left(\alpha+\frac{2\pi m}{M}\right)} \leq e^{\frac{x^2}{4}}\epsilon_2(x),$$

for any $\alpha \in \mathbb{R}$ and $M > 2$ where

$$\epsilon_2(x) = 1 + \frac{e^x - 1 - x - \frac{x^2}{2} - \left(e^{\frac{x^4}{4}} - 1 - \frac{x^4}{4}\right)}{e^{\frac{x^4}{2}}} \tag{4.3}$$

satisfying $\epsilon_2(x) \to (1 + \mathcal{O}(x^3))$ as $x \to 0$.

Proof. Expanding the exponential function yields

$$\begin{aligned}\frac{1}{M}\sum_{m=0}^{M-1} e^{x\cos\left(\alpha+\frac{2\pi m}{M}\right)} &= 1 + \frac{\sum_{m=0}^{M-1}\cos\left(\alpha+\frac{2\pi m}{M}\right)x}{M} + \frac{\sum_{m=0}^{M-1}\cos^2\left(\alpha+\frac{2\pi m}{M}\right)x^2}{2M} + \dots \\ &= 1 + \frac{x^2}{4} + \dots\end{aligned}$$

where we used

$$\sum_{m=0}^{M-1}\cos\left(\alpha+\frac{2\pi m}{M}\right) = 0$$

and

$$\sum_{m=0}^{M-1}\cos^2\left(\alpha+\frac{2\pi m}{M}\right) = \frac{M}{2}.$$

for any $\alpha \in \mathbb{R}$ and $M > 2$. The coefficients are trivially upperbounded by the coefficients of e^x and we can write

$$\frac{1}{M}\sum_{m=0}^{M-1} e^{x\cos\left(\alpha+\frac{2\pi m}{M}\right)} \leq e^{\frac{x^2}{4}} + \epsilon'(x)$$

where the error is upperbounded by

$$\epsilon'(x) = e^x - 1 - x - \frac{x^2}{2} - \left(e^{\frac{x^4}{4}} - 1 - \frac{x^4}{4}\right).$$

The term $\epsilon'(x)$ satisfies $\epsilon'(x) = \mathcal{O}(x^3)$ as $x \to 0$ and thus $\epsilon_2(x) = 1 + \mathcal{O}\left(\frac{x^3}{1+x^2}\right) = 1 + \mathcal{O}(x^3)$ as $x \to 0$. ■

We are now able to prove the theorem.

Proof of theorem. It can be shown that the following bounds are independent of A and hence we set without loss of generality $A = 1$. For $\mathcal{Q} = \text{BPSK}$, i.e. $c_k = \{-1, 1\}$, we have

$$\Pr\left(\widehat{s}_c(\theta,\alpha) > \lambda\right) \le e^{-\lambda\varepsilon} \prod_{k=0}^{N-1} \cosh\left[\varepsilon \cos(k\theta + \alpha)\right]$$

Using $\cosh(x) \le e^{\frac{x^2}{2}}$ yields

$$\Pr\left(\widehat{s}_c(\theta,\alpha) > \lambda\right) \le e^{-\lambda\varepsilon} e^{\frac{\varepsilon^2}{2}\sum_{k=0}^{N-1}\cos^2(k\theta+\alpha)}$$

and

$$\Pr\left(\widehat{s}_c(\theta,\alpha) > \lambda\right) \le e^{-\lambda\varepsilon} e^{\frac{\varepsilon^2}{2} C_1(N,\theta,\alpha)}$$

where

$$\Psi(N,\theta,\alpha) := \sum_{k=0}^{N-1} \cos^2(k\theta + \alpha) = \frac{N}{2} + \frac{1}{2}\frac{\sin(N\theta)}{\sin(\theta)} \cos\left[(N-1)\theta + 2\alpha\right].$$

The last identity is given for purely numerical purposes. Optimizing the right hand side with respect to ε yields

$$\begin{aligned} \frac{d}{d\varepsilon} -\varepsilon\lambda + \frac{1}{2}\varepsilon^2 \Psi(N,\theta,\alpha)\Big|_{\varepsilon=\overline{\varepsilon}} &= -\lambda + \overline{\varepsilon}\Psi(N,\theta,\alpha) = 0 \\ &\Rightarrow \overline{\varepsilon} = \frac{\lambda}{\Psi(N,\theta,\alpha)}. \end{aligned}$$

Thus, we obtain

$$\Pr\left(\widehat{s}_c(\theta,\alpha) > \lambda\right) \le e^{-\frac{\lambda^2}{2\Psi(N,\theta,\alpha)}}$$

and

$$\max_{0\le\alpha<2\pi} \widehat{s}_c(\theta,\alpha) \le C_K \cdot \max_{0\le l_2<K} \left|\widehat{s}_c\left(\theta, \frac{2\pi l_2}{K}\right)\right|, \quad \theta \notin \{0,\pi\}$$

and for the special case $\theta = 0$ or $\theta = \pi$ where $\widehat{s}_c(\theta,\alpha)$ is real

$$|s(\theta)| = \max\{s(\theta), -s(\theta)\}, \quad \theta \in \{0,\pi\}.$$

Thus

$$\Pr\left(|s(\theta)|>\lambda\right)\leq\sum_{l_2=0}^{K-1}\Pr\left(\left|\widehat{s}_c\left(\theta,\frac{2\pi l_2}{K}\right)\right|>\frac{\lambda}{C_K}\right),\quad\theta\notin\{0,\pi\},$$

and

$$\Pr\left(|s(\theta)|>\lambda\right)\leq 2\Pr\left(s(\theta)>\lambda\right),\quad\theta\in\{0,\pi\}.$$

By Lemma 13

$$\Pr\left(\max_{0\leq\theta<2\pi}|s(\theta)|>\lambda\right)\leq 4e^{-\frac{\lambda^2}{2NC_L^2}}+\sum_{l_1=1,l_1\neq\frac{LN}{2}}^{LN-1}\sum_{l_2=0}^{K-1}e^{-\frac{\lambda^2}{2C_1\left(N,\frac{l_1}{LN}2\pi,\frac{l_2}{K}2\pi\right)C_K^2C_L^2}}.$$

Note that owing to the real data we can also multiply the last equation by a factor $\frac{1}{2}$ since $|s(\theta)|$ is symmetric around π, i.e.

$$|s(\theta)|=|s(2\pi-\theta)|.$$

Setting $\theta=\frac{2\pi l}{LN},0<l<\frac{LN}{2}$ (N assumed even) yields

$$\left|s\left(\frac{2\pi l}{LN}\right)\right|=\left|s\left(2\pi-\frac{2\pi l}{LN}\right)\right|=\left|s\left(\frac{2\pi\left(LN-l\right)}{LN}\right)\right|,\quad 0<k<\frac{LN}{2}.$$

Hence if the signal takes on a certain value at $\frac{l}{LN}$ so it does at $\frac{LN-l}{LN}$ for $0<k<\frac{LN}{2}$. On the other hand the argumentation does not hold for $\theta=0$ and $\theta=\pi$. For these samples we can improve our result by incorporating the fact that the distribution of $s(0)$ can be explicitly given in terms of the binomial coefficients, i.e. for N even

$$\Pr\left(s(0)=\lambda\right)=\frac{1}{2^N}\begin{cases}\binom{N}{\frac{(N-\lambda)}{2}} & \lambda\text{ even (integer), }|\lambda|\leq N\\ 0 & \text{else}\end{cases}.$$

Furthermore, since $s(0)$ and $s(\pi)$ have the same distribution, we arrive at

$$\Pr\left(s(0)>\lambda\right)=\Pr\left(s(\pi)>\lambda\right)=\sum_{k=\lceil\lambda\rceil,k\text{ even}}^{N}\Pr\left(s(0)=k\right).$$

Setting $\sigma_s=1$ and optimizing K and L the result follows.

For $\mathcal{Q}=$ M-QAM, i.e. $c_k\in\left\{\left((2m_1-1)+j\left(2m_2-1\right)\right),m_1,m_2\in\left\{-\frac{m}{2}+1,\ldots,\frac{m}{2}\right\}\right\}M=m^2$, we have by independence

$$E\left(e^{\varepsilon\widehat{s}_c(\theta,\alpha)}\right)=E\left\{e^{\varepsilon\sum_{k=0}^{N-1}\Re(c_k)\cos(k\theta+\alpha)}\right\}E\left\{e^{-\varepsilon\sum_{k=0}^{N-1}\Im(c_k)\sin(k\theta+\alpha)}\right\}.\tag{4.4}$$

Defining $x_1 := \cos(k\theta + \alpha)$ and $x_2 := \sin(k\theta + \alpha)$ processing the first term on the right hand side of (4.4) yields

$$E\left(e^{\varepsilon \sum_{k=0}^{N-1} \Re(c_k) x_1}\right) = \prod_{k=0}^{N-1} \frac{1}{\sqrt{M}} \left[e^{-\varepsilon x_1} \sum_{m=1}^{\frac{\sqrt{M}}{2}} e^{2\varepsilon m x_1} + e^{\varepsilon x_1} \sum_{m=1}^{\frac{\sqrt{M}}{2}} e^{-2\varepsilon m x_1} \right]. \tag{4.5}$$

Introducing the series

$$\sum_{m=1}^{\sqrt{M}} q^{m-1} = \frac{1 - q^{\sqrt{M}}}{1 - q}$$

for some real $q \neq 1$ and this time processing the first sum on the right hand side of (4.5) yields

$$\begin{aligned} \sum_{m=1}^{\frac{\sqrt{M}}{2}} e^{2\varepsilon m x_1} &= e^{\varepsilon \ \ 1+\frac{\sqrt{M}}{2} \ \ x_1} \frac{\sinh\left(\frac{\varepsilon\sqrt{M}x_1}{2}\right)}{\sinh(\varepsilon x_1)} \\ &\leq e^{\varepsilon \ \ 1+\frac{\sqrt{M}}{2} \ \ x_1} \frac{\sqrt{M}}{2} e^{\left(\frac{M}{12} - \frac{1}{3}\right)\frac{\varepsilon^2 x_1^2}{2}} \epsilon_1(\varepsilon) \end{aligned}$$

where the inequality is due to Lemma 14.

The second sum is upperbounded by

$$\sum_{m=1}^{\frac{\sqrt{M}}{2}} e^{(-2\varepsilon m x_1)} \leq \frac{\sqrt{M}}{2} e^{-\varepsilon \ \ 1+\frac{\sqrt{M}}{2} \ \ x_1} e^{\left(\frac{M}{12} - \frac{1}{3}\right)\frac{\varepsilon^2 x_1^2}{2}} \epsilon_1(\varepsilon).$$

Setting $\gamma = \left(\frac{M}{12} - \frac{1}{3}\right)$ and inserting both into (4.5) yields

$$\begin{aligned} E\left(e^{\left(\varepsilon \sum_{k=0}^{N-1} \Re(c_k) x_1\right)}\right) &\leq \prod_{k=0}^{N-1} \frac{1}{2} e^{\frac{\gamma\varepsilon^2 x_1^2}{2}} \left[e^{\frac{\varepsilon\sqrt{M}x_1}{2}} + e^{-\frac{\varepsilon\sqrt{M}x_1}{2}} \right] \epsilon_1(\varepsilon) \\ &= \prod_{k=0}^{N-1} e^{\frac{\gamma\varepsilon^2 x_1^2}{2}} \cosh\left(\frac{\varepsilon\sqrt{M}x_1}{2}\right) \epsilon_1(\varepsilon) \\ &\leq \prod_{k=0}^{N-1} e^{\frac{\gamma\varepsilon^2 x_1^2}{2}} e^{\frac{M\varepsilon^2 x_1^2}{8}} \epsilon_1(\varepsilon) \\ &= e^{\frac{\varepsilon^2}{2}\left(\gamma + \frac{M}{4}\right) \sum_{k=0}^{N-1} \cos^2(k\theta+\alpha)} \epsilon_1^N(\varepsilon) \end{aligned}$$

where we again used the inequality $\cosh(x) \leq e^{\frac{x^2}{2}}$. In the same manner, the second term of (4.4) is processed to give

$$E\left(e^{-\varepsilon \sum_{k=0}^{N-1} \Im(c_k) \sin(k\theta+\alpha)}\right) \leq e^{\frac{\varepsilon^2}{2}\left(\gamma + \frac{M}{4}\right) \sum_{k=0}^{N-1} \sin^2(k\theta+\alpha)} \epsilon_1(\varepsilon).$$

In the case of numerical computation we have the upper bound

$$E\left(e^{\varepsilon\widehat{s}_c(\theta,\alpha)}\right) \leq \prod_{m=1}^{2}\prod_{k=0}^{N-1}\frac{2\sinh\left(\frac{\varepsilon\sqrt{M}x_m}{2}\right)\cosh\left(\frac{\varepsilon\sqrt{M}x_m}{2}\right)}{\sqrt{M}\sinh\left(\varepsilon x_m\right)}$$

and otherwise

$$\begin{aligned}\Pr\left(\widehat{s}_c\left(\theta,\alpha\right)>\lambda\right) &\leq e^{-\lambda\varepsilon}e^{\frac{\varepsilon^2}{2}\left(\gamma+\frac{M}{4}\right)\left(\sum_{k=0}^{N-1}\cos^2(k\theta+\alpha)+\sum_{k=0}^{N-1}\sin^2(k\theta+\alpha)\right)}\epsilon_1^{2N}\left(\varepsilon\right)\\ &= e^{-\lambda\varepsilon}e^{\frac{N\varepsilon^2}{2}\left(\gamma+\frac{M}{4}\right)}\epsilon_1^{2N}\left(\varepsilon\right).\end{aligned}$$

Optimizing the right hand side with respect to ε yields

$$\begin{aligned}\frac{d}{d\varepsilon}\, -\varepsilon\lambda+\frac{1}{2}\varepsilon^2N\left(\gamma+\frac{M}{4}\right)\Big|_{\varepsilon=\overline{\varepsilon}} &= -\lambda+\overline{\varepsilon}N\left(\gamma+\frac{M}{4}\right)=0\\ &\Rightarrow \overline{\varepsilon}=\frac{\lambda}{N\left(\gamma+\frac{M}{4}\right)},\end{aligned}$$

and we get

$$\Pr\left(\widehat{s}_c\left(\theta,\alpha\right)\geq\lambda\right)\leq e^{-\frac{\lambda^2}{2N\left(\gamma+\frac{M}{4}\right)}}\epsilon_1^{2N}\left(\overline{\varepsilon}\right).$$

The averaged RMS value of s is

$$\sigma_s=\sqrt{\frac{2\left(M-1\right)}{3}}$$

and therefore

$$\begin{aligned}\Pr\left(\mathrm{CF}\left(s\right)>\lambda\right) &= \Pr\left(\max_{\theta\in[0,2\pi]}\left|s\left(\theta\right)\right|>\frac{\sigma_s\sqrt{N}\lambda}{C_LC_K}\right)\\ &\leq KLN\,e^{-\frac{(M-1)\lambda^2}{3\left(\gamma+\frac{M}{4}\right)C_L^2C_K^2}}\epsilon_1^{2N}\left(\frac{\lambda\sigma_s}{C_LC_K\sqrt{N}\left(\gamma+\frac{M}{4}\right)}\right)\\ &= KLN\,e^{-\frac{\lambda^2}{C_L^2C_K^2}}\epsilon_1^{2N}\left(\frac{\lambda\sigma_s}{C_LC_K\sqrt{N}\left(\gamma+\frac{M}{4}\right)}\right).\end{aligned}$$

For $\mathcal{Q}$ = M-PSK, i.e. $c_k\in\left\{1,e^{\frac{2\pi j}{M}},\ldots,e^{\frac{2\pi j(M-1)}{M}}\right\}$ define RV's $\varphi_k\in\left\{0,\frac{2\pi}{M},\ldots,\frac{2\pi(M-1)}{M}\right\}$ uniformly distributed over their range. The Chernoff bound yields

$$\begin{aligned}\Pr\left(\widehat{s}_c\left(\theta,\alpha\right)>\lambda\right) &\leq e^{-\lambda\varepsilon}Ee^{\varepsilon\sum_{k=0}^{N-1}\cos(k\theta+\alpha+\varphi_k)}\\ &= e^{-\lambda\varepsilon}\prod_{k=0}^{N-1}\frac{1}{M}\sum_{m=0}^{M-1}e^{\varepsilon\cos\left(\frac{2\pi m}{M}\right)}.\end{aligned}$$

Due to Lemma 15

$$\Pr\left(\sum_{k=0}^{N-1}\cos\left(k\theta+\alpha+\varphi_k\right)>\lambda\right)\leq e^{-\lambda\varepsilon}e^{\frac{\varepsilon^2 N}{4}}\epsilon_2^N\left(\varepsilon\right)$$

and with $\overline{\varepsilon}=\frac{2\lambda}{N}$ the result follows. ■

In Fig. 4.1 we compare simulation, some well-known approximations and our bound for BPSK modulation and 48 subcarriers. Due to the (relatively) low number of subcarriers we are able to evaluate the Chernoff bound numerically. The analytical upper bounds from Theorem 12 can be used to obtain optimal values for K and L (by simple search procedures). Then, these optimal values can be taken in a subsequent optimization of ε (however, we simply took the optimal ε from the Chernoff upper bound). In Fig. 4.1 there is also a lower bound indicating that our bound well approximates the complementary distribution in the low probability region. We conclude that for the BPSK case, the standard models are not appropriate.

In Fig. 4.2 we compare simulation, some approximations and the bound for several other modulation schemes and 48 subcarriers. It is observed that this time the approximations are better and the difference is moderate and below 1 [dB]. Note that the difference between the modulation schemes is important for systems that employ adaptive modulation.

Generally we can state that given a certain outage probability the effective CF grows slower than the maximum CF (in fact, it is of order $\log\left(N\right)$ rather than N as we will prove). Clearly the bounds are very weak in the high probability region but the curves always provide the correct trend. It can be seen for example, that the BPSK curve is slightly better in the high probability region compared to, for instance, QPSK while in the low probability region it is worse thereby mimicking the real situation. Furthermore we can now prove that with only one additional bit the CF can be bound within 40% of its maximum value. On the other hand, clearly this bit cannot be assigned to a particular subcarrier.

Why does the union bound work? For higher values of λ the average power constraint is inherently satisfied since if the signal is large at a point it cannot be that large at other points and the sets become distinct as required in the upper bound. Thus the lemma can be used for any constellation or even codes. Furthermore, we may consider numerical evaluation of the Chernoff

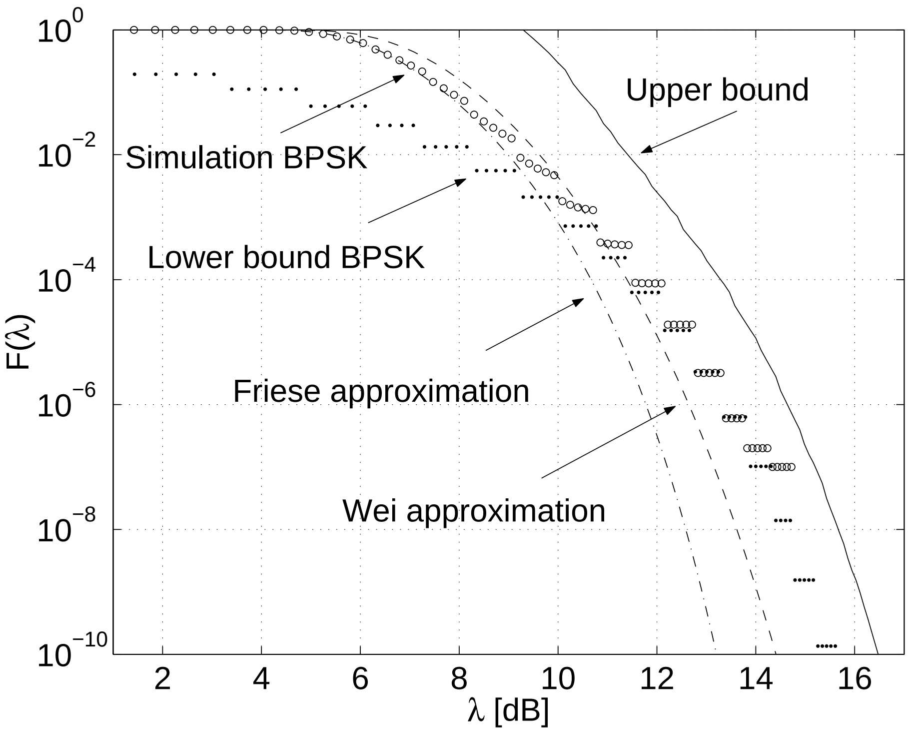

Figure 4.1: Complementary distribution of BPSK modulation and bound for $N = 48$ subcarriers. The reference Wei can be found in [81]. The reference Friese can be found in [28]. The lower bound is due to the binomial distribution for $\theta = 0$. The Monte-Carlo simulation was obtained by simulating 10^7 OFDM symbols using one times oversampling.

bound in general but for large numbers the gain is negligible while the effort is much higher [33]. Additionally, the upper bound approach enabled us to select the optimization parameters optimally.

4.1.2 Asymptotic lower bounds

These bounds shall also be further analyzed to give some interesting information-theoretic results in terms of an asymptotic analysis, i.e. for N large.

Corollary 16 *For the modulation schemes in Theorem 12 we have for any $\varepsilon > 0$*

$$\lim_{N\to\infty} \Pr\left(\mathrm{CF}\left(s_c\right) > \sqrt{(1+\varepsilon)\log\left(N\right)}\right) = 0.$$

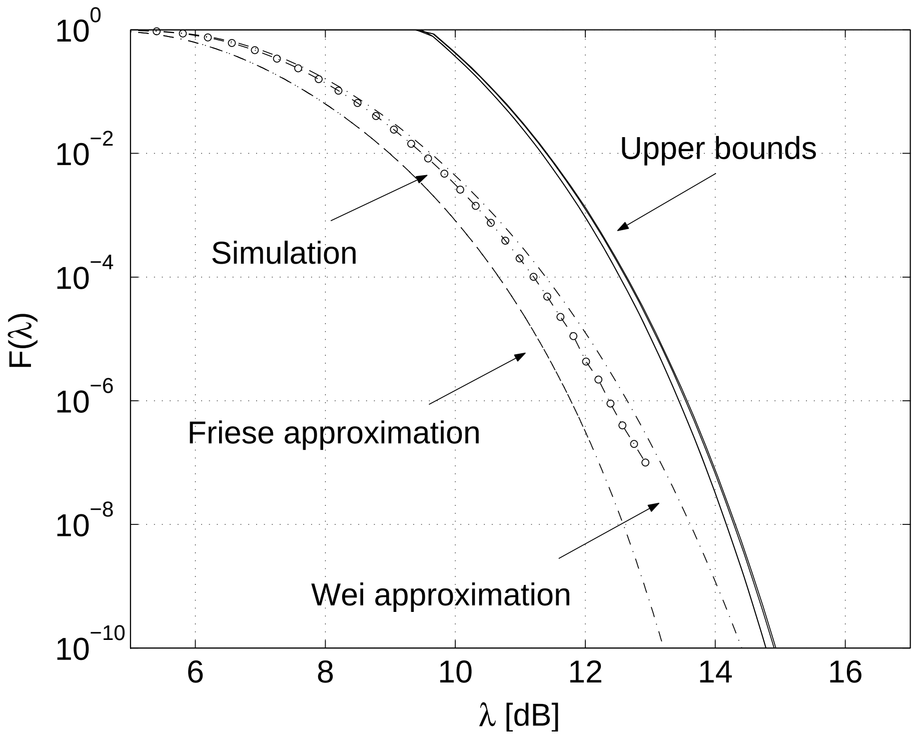

Figure 4.2: Complementary distributions for $N = 48$ subcarriers and several modulation schemes (solid lines): QPSK, 16-QAM, 64-QAM and 16-PSK. Note that QPSK, 16-QAM, 64-QAM are indistinguishable, while 16-PSK is slightly better. The simulation was obtained by simulating 10^6 OFDM symbols using one times oversampling.

Proof. The proof follows immediately from Theorem 12, Lemma 14, 15 and the upper bound $\log(1+x) \leq x$. ∎

As we already implemented a lower bound for the BPSK case, it is interesting to ask for general lower bounds. It is often argued that since the instantaneous distribution of OFDM signals converges in probability to a Gaussian distribution, the Gaussian distribution can account for the asymptotic behavior as suggested in [30]. However, this argumentation is clearly flawed. The following theorem is more general.

Theorem 17 *Let $c_k, k = 0, \ldots, N-1$ be (not necessarily independent) RV's with $\sigma_s^2 = 1$. Suppose that for some non-trivial interval $\omega \in [-d, d]$*

(i) $Ee^{j\omega\Re\left(\sum_{k=0}^{N-1} c_k e^{j(k\theta+\alpha)}\right)} = e^{-\frac{\omega^2 N}{4}+\mathcal{O}\left(N\omega^4\right)}, \quad \alpha, \theta \in [0, 2\pi)\,,$

(ii) $Ee^{j\Re\left[\sum_{k=0}^{N-1} c_k \ \omega_1 e^{jk\theta_{l_1,L}}+\omega_2 e^{jk\theta_{l_2,L}}\ \right]} = e^{-\frac{N\omega_1^2}{4}-\frac{N\omega_2^2}{4}+\mathcal{O}\left(N(|\omega_1|+|\omega_2|)^4\right)}, \quad 0 \le l_1 < N, 0 \le l_2 < N, l_1 \ne l_2.$

Then, for some $\gamma \ge 2$ independent of N

$$\Pr\left(\sqrt{\log(N)} - \gamma\sqrt{\frac{1}{\log(N)}}\log[\log(N)] < \mathrm{CF}(s_c) < \sqrt{\log(N)} + \gamma\sqrt{\frac{1}{\log(N)}}\log[\log(N)]\right)$$
$$= 1 - \mathcal{O}\left(\frac{1}{[\log(N)]^{2\gamma-\frac{5}{2}}}\right)$$

for N large enough.

Proof. We use the Fourier-Stieltjes integral method [24] and define the function

$$u_\lambda(x) = \begin{cases} 1 & |x| \ge \lambda \\ \overline{u}(x) & \lambda - \delta \le |x| \le \lambda \\ 0 & |x| \le \lambda - \delta \end{cases}, \quad \delta > 0,$$

that due to Bochner's theorem [67] can be represented by a Fourier-Stieltjes integral of form

$$u_\lambda(x) = \int_{-\infty}^{+\infty} e^{j\omega x} dU_\lambda(\omega)$$

where U is a function of bounded variation. The function $\overline{u}(x)$ can be chosen so that U is 10 times differentiable. It was shown in [24] that in this case

$$\int_{-\infty}^{+\infty} |dU_\lambda(\omega)| = \mathcal{O}\left(\frac{\lambda}{\delta}\right) \tag{4.6}$$

and

$$\int_{-\infty}^{+\infty} |\omega|^p |dU_\lambda(\omega)| = \mathcal{O}\left(\frac{1}{\delta^p}\right). \tag{4.7}$$

for $1 \le p \le 10$. Introducing the RV

$$\xi_{L,K}(\lambda) = \sum_{l_1=0}^{LN-1}\sum_{l_2=0}^{K-1}\int_{-\infty}^{+\infty} e^{j\omega\Re\left[\sum_{k=0}^{N-1} c_k e^{j\left(k\theta_{l_1,L}+\alpha_{l_2,K}\right)}\right]} dU_\lambda(\omega)$$

we have

$$\Pr\left(\max_{0\le\theta<2\pi}|s_c(\theta)| > \lambda\right) \le \Pr\left(\xi_{L,K}\left(\frac{\lambda}{C_L C_K}\right) > 1\right)$$
$$\le E\left(\xi_{L,K}\left(\frac{\lambda}{C_L C_K}\right)\right), \quad L > 1, K > 2,$$

by Markov's inequality.

Now, by hypothesis

$$\int_{-\infty}^{+\infty} Ee^{j\omega\Re\left[\sum_{k=0}^{N-1} c_k e^{j(k\theta+\alpha)}\right]} dU_\lambda(\omega) = \int_{-d}^{+d} e^{-\frac{N\omega^2}{4}+\mathcal{O}\left(\omega^4 N\right)} dU_\lambda(\omega) + \int_{|\omega|\geq d} |dU_\lambda(\omega)|$$

where we used the trivial bound in the last term outside some interval $[-d, d]$. It is easy to see that $e^{-a} \leq e^{-b} + (b-a)\, e^{-b} + (b-a)^2$ for $a, b > 0$. Setting $b = \frac{N\omega^2}{4}$ and $a = \frac{N\omega^2}{4} - \mathcal{O}\left(\omega^4 N\right)$ we have for the first term

$$\begin{aligned} &\int_{-d}^{+d} e^{-\frac{N\omega^2}{4}+\mathcal{O}\left(\omega^4 N\right)} dU_\lambda(\omega) \\ &= \int_{-d}^{+d} e^{-\frac{N\omega^2}{4}} dU_\lambda(\omega) + \mathcal{O}\left(N \int_{-d}^{+d} \omega^4 e^{-\frac{N\omega^2}{4}} dU_\lambda(\omega)\right) + \mathcal{O}\left(N^2 \int_{-d}^{+d} \omega^8 dU_\lambda(\omega)\right). \end{aligned}$$

Integration in the first term on the right hand side can be extended to infinity making the same error as in the trivial bound on the last term which can in turn be included in the second term. Note that the constants are required to be positive. Since $a = \frac{N\omega^2}{4} + \mathcal{O}\left(\omega^4 N\right)$ this may not be fulfilled for the range $[0, d)$. Then, we need to rescale the range of d somewhat.

Thus, according to Plancherel's Theorem [68]

$$\begin{aligned} \int_{-\infty}^{\infty} e^{-\frac{N\omega^2}{4}} dU_\lambda(\omega) &= \sqrt{\frac{4\pi}{N}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{N}} u(x)\, dx \\ &\leq \sqrt{\frac{4\pi}{N}} \int_{|x|\geq\lambda-\delta} \frac{x}{\lambda-\delta} e^{-\frac{x^2}{N}} dx \\ &= \frac{2N}{(-2)(\lambda-\delta)} \sqrt{\frac{4\pi}{N}} \int_{x\geq\lambda-\delta} -\frac{2x}{N} e^{-\frac{x^2}{N}} dx \\ &= \frac{2\sqrt{\pi N}}{\lambda-\delta} e^{-\frac{(\lambda-\delta)^2}{N}} \end{aligned} \tag{4.8}$$

and by partial integration

$$\int_{-\infty}^{\infty} \omega^4 e^{-\frac{N\omega^2}{4}} dU_\lambda(\omega) = \frac{2}{N} \int_{-\infty}^{\infty} \omega^3 \frac{N\omega}{2} e^{-\frac{N\omega^2}{4}} dU_\lambda(\omega)$$

$$
\begin{aligned}
&= \frac{2}{N}\int_{-\infty}^{\infty} 3\omega^2 e^{-\frac{N\omega^2}{4}} dU_\lambda(\omega) \\
&= \frac{12}{N^2}\int_{-\infty}^{\infty} \omega \frac{N\omega}{2} e^{-\frac{N\omega^2}{4}} dU_\lambda(\omega) \\
&= \frac{12}{N^2}\int_{-\infty}^{\infty} e^{-\frac{N\omega^2}{4}} dU_\lambda(\omega) \\
&\leq \frac{12}{N^2}\frac{2\sqrt{\pi N}}{\lambda-\delta} e^{-\frac{(\lambda-\delta)^2}{N}}.
\end{aligned}
$$

Appealing to (4.7) we have

$$
\begin{aligned}
\int_{|\omega|\geq d} |dU_\lambda(\omega)| &\leq \frac{1}{d^p}\int_{|\omega|\geq d} |\omega|^p |dU_\lambda(\omega)| \\
&\leq \frac{1}{d^p}\int_{-\infty}^{+\infty} |\omega|^p |dU_\lambda(\omega)| \\
&\leq \frac{1}{d^p}\int_{-\infty}^{+\infty} |\omega|^p |dU_\lambda(\omega)| \leq \mathcal{O}\left(\frac{1}{(d\delta)^p}\right)
\end{aligned}
$$

and

$$
\Pr\left(\max_{0\leq\theta<2\pi} |s_c(\theta)| > \lambda\right) \leq \mathcal{O}\left(\frac{2LNK\sqrt{\pi N}\left(1+\frac{12}{N}\right)}{\left(\frac{\lambda}{C_L C_K}-\delta\right)} e^{-\frac{\left(\frac{\lambda}{C_L C_K}-\delta\right)^2}{N}}\right) + \mathcal{O}\left(\frac{LKN^3}{d^8\delta^8}\right).
$$

The asymptotic behavior is governed by choosing the values of δ, λ, L, K accordingly. Choosing $L = K = \sqrt{\log(N)}$ then $\frac{1}{C_L^2 C_K^2} \geq 1 - \frac{2\pi^2}{\log(N)}$ for N large. Taking $\delta = \sqrt{\frac{N}{\log(N)}}$ and $\lambda_N = \sqrt{N\log(N)} + \gamma\sqrt{\frac{N}{\log(N)}}\log[\log(N)]$ yields

$$
\begin{aligned}
\Pr\left(\max_{0\leq\theta<2\pi} |s_c(\theta)| > \lambda_N\right) &= \mathcal{O}\left(\frac{N\sqrt{\log(N)}}{N^{\;1-\frac{2\pi^2}{\log(N)}}\;[\log(N)]^{2\gamma\;1-\frac{2\pi^2}{\log(N)}}}\right) + \mathcal{O}\left(\frac{\log^5(N)}{N}\right) \\
&= \mathcal{O}\left(\frac{1}{[\log(N)]^{2\gamma-\frac{1}{2}}}\right).
\end{aligned}
$$

for N large.

For the lower bound define

$$
\xi_1(\lambda) = \sum_{l=0}^{N-1}\int_{-\infty}^{+\infty} e^{j\omega\Re\left[\sum_{k=0}^{N-1} c_k e^{j\left(k\theta_{l,1}\right)}\right]} dU_\lambda(\omega).
$$

Then

$$\begin{aligned}
\Pr\left(\max_{0\le\theta<2\pi}|s_c(\theta)|>\lambda\right) &= 1-\Pr\left(\max_{0\le\theta<2\pi}|s_c(\theta)|\le\lambda\right)\\
&\ge 1-\Pr\left(\max_{0\le l<N}|s_c(\theta_{l,L})|<\lambda\right)\\
&= 1-\Pr\left(\xi_1(\lambda+\delta)=0\right)\\
&\ge 1-\frac{E\left(\xi_1^2(\lambda+\delta)\right)-E^2\left(\xi_1(\lambda+\delta)\right)}{E^2\left(\xi_1(\lambda+\delta)\right)}
\end{aligned}$$

where the last step is due to Tschebyscheff's inequality.

Since $e^{-a}\le e^{-b}+|(b-a)|$ for $a,b>0$ we have for the lower bound

$$\begin{aligned}
&\int_{-d}^{+d} e^{-\frac{N\omega^2}{4}+\mathcal{O}\left(N\omega^4\right)}dU_\lambda(\omega)\\
&=\int_{-d}^{+d} e^{-\frac{N\omega^2}{4}}dU_\lambda(\omega)+\mathcal{O}\left(N\int_{-d}^{+d}\omega^4 dU_\lambda(\omega)\right)+\mathcal{O}\left(\int_{|\omega|\ge d}|dU_\lambda(\omega)|\right)\\
&=\int_{-\infty}^{+\infty} e^{-\frac{N\omega^2}{4}}dU_\lambda(\omega)+\mathcal{O}\left(N\int_{-\infty}^{+\infty}\omega^4 dU_\lambda(\omega)\right)\\
&=\int_{-\infty}^{+\infty} e^{-\frac{N\omega^2}{4}}dU_\lambda(\omega)+\mathcal{O}\left(\frac{N}{\delta^4}\right).
\end{aligned}$$

The variance is given by

$$\xi_1^2(\lambda)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{j\Re\left[\sum_{k=0}^{N-1}c_k\ \ \omega_1 e^{jk\theta_{l_1,1}}+\omega_2 e^{jk\theta_{l_2,1}}\ \right]}dU_\lambda(\omega_1)\,dU_\lambda(\omega_2)$$

and by hypothesis

$$Ee^{j\Re\left[\sum_{k=0}^{N-1}c_k\ \ \omega_1 e^{jk\theta_{l_1}}+\omega_2 e^{jk\theta_{l_2}}\ \right]}=e^{-\frac{N\omega_1^2}{4}-\frac{N\omega_2^2}{4}+\mathcal{O}\left(N(|\omega_1|+|\omega_2|)^4\right)},\quad 0\le l_1<N, 0\le l_2<N,$$

for $l_1\ne l_2$ and

$$E\left(\xi_1^2(\lambda)\right)\le 2E\left(\xi_1(\lambda)\right).$$

for $l_1=l_2$. Therefore, we have

$$\begin{aligned}
&\int_{-d}^{+d}\int_{-d}^{+d} e^{-\frac{N\omega_1^2}{4}-\frac{N\omega_2^2}{4}+\mathcal{O}\left(N(|\omega_1|+|\omega_2|)^4\right)}dU_\lambda(\omega_1)\,dU_\lambda(\omega_2)\\
&=\left(\int_{-d}^{+d} e^{-\frac{N\omega_1^2}{4}}dU_\lambda(\omega_1)\right)^2+\mathcal{O}\left(N\int_{-d}^{+d}\int_{-d}^{+d}(|\omega_1|+|\omega_2|)^4 e^{-\frac{N\omega_1^2}{4}-\frac{N\omega_2^2}{4}}dU_\lambda(\omega_1)\,dU_\lambda(\omega_2)\right)\\
&+\mathcal{O}\left(N^2\int_{-d}^{+d}\int_{-d}^{+d}(|\omega_1|+|\omega_2|)^8 dU_\lambda(\omega_1)\,dU_\lambda(\omega_2)\right).
\end{aligned}$$

Again, we may extend integration to infinity. The second term is given by

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left(|\omega_1|+|\omega_2|\right)^4 e^{-\frac{N\omega_1^2}{4}-\frac{N\omega_2^2}{4}}\,dU_\lambda(\omega_1)\,dU_\lambda(\omega_2)=\mathcal{O}\left(\left(\frac{12}{N^2}\frac{2\sqrt{\pi N}}{\lambda-\delta}e^{-\frac{(\lambda-\delta)^2}{N}}\right)^2\right)$$

and the third term

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left(|\omega_1|+|\omega_2|\right)^8 dU_\lambda(\omega_1)\,dU_\lambda(\omega_2)\leq\mathcal{O}\left(\frac{\lambda}{\delta^9}\right)$$

Hence, the second moment is given by

$$E\left(\xi_1^2(\lambda)\right)=2E\left(\xi_1(\lambda)\right)+\mathcal{O}\left(\frac{2N^2\sqrt{\pi N}}{\lambda-\delta}e^{-\frac{(\lambda-\delta)^2}{N}}\right)^2+\mathcal{O}\left(\frac{\lambda N^4}{\delta^9}\right).$$

The variance is given by

$$E\left(\xi_1^2(\lambda)\right)-E^2\left(\xi_1(\lambda)\right)=2E\left(\xi_1(\lambda)\right)+\mathcal{O}\left(\frac{N^4\lambda}{\delta^9}\right)+\mathcal{O}\left(\frac{N^2E\left(\xi_1(\lambda)\right)}{\delta^4}\right)+\mathcal{O}\left(\frac{N^4}{\delta^8}\right)$$

and

$$\Pr\left(\max_{0\leq\theta<2\pi}|s_c(\theta)|>\lambda\right)\geq 1-\frac{2E\left(\xi_1(\lambda+\delta)\right)+\mathcal{O}\left(\frac{N^4(\lambda+\delta)}{\delta^9}\right)+\mathcal{O}\left(\frac{N^2E(\xi_1(\lambda+\delta))}{\delta^4}\right)}{E^2\left(\xi_1(\lambda+\delta)\right)}$$

Taking $\delta=\sqrt{\frac{N}{\log(N)}}$ and setting this time $\lambda_N=\sqrt{N\log(N)}-\gamma\sqrt{\frac{N}{\log(N)}}\log\left[\log(N)\right]$ yields

$$\begin{aligned}\Pr\left(\max_{0\leq\theta<2\pi}|s_c(\theta)|>\lambda\right)&\geq 1-\mathcal{O}\left(\frac{E\left(\xi_1(\lambda+\delta)\right)\log^2(N)+\log^5(N)}{E^2\left(\xi_1(\lambda+\delta)\right)}\right)\\&\geq 1-\mathcal{O}\left(\frac{\log^2(N)}{E\left(\xi_1(\lambda+\delta)\right)}\right)\end{aligned}$$

provided that

$$E\left(\xi_1(\lambda+\delta)\right)\log^2(N)\geq\log^5(N).$$

Again by Plancherel's formula

$$\begin{aligned}E\left(\xi_1(\lambda)\right)&\geq N\int_{-\infty}^{+\infty}e^{-\frac{N\omega^2}{4}}dU_\lambda(\omega)\\&=\frac{N}{\sqrt{N\pi}}\int_{-\infty}^{+\infty}e^{-\frac{x^2}{N}}u(x)\,dx\\&\geq\frac{N}{\sqrt{N\pi}}\int_{\lambda+\delta}^{\lambda+2\delta}e^{-\frac{x^2}{N}}dx\\&\geq\frac{2N\delta}{\sqrt{N\pi}}e^{-\frac{(\lambda+2\delta)^2}{N}}\end{aligned}$$

finally

$$E\left(\xi_1\left(\lambda_N\right)\right) \geq \frac{2}{\sqrt{\pi}}\left[\log\left(N\right)\right]^{2\gamma-\frac{1}{2}}.$$

for $\gamma \geq 2$. ∎

It seems difficult to translate this approach into real upper and lower bounds for a specific constellation (and classes of constellations) since the generating function $U_\lambda\left(\omega\right)$ is not monotonously increasing. Splitting up the measure in two positive measures may help if the number of terms in the characteristic function is small. Furthermore, note that it is essential that the third order term in the characteristic function is zero. We also want to mention that the second condition in the theorem can be replaced by any set of cardinality $\mathcal{O}\left(N\right)$.

The theorem will now exclude a number of constellations with asymptotic CF performance better than $\mathcal{O}\left(\sqrt{\log\left(N\right)}\right)$ given by the following corollary.

Corollary 18 *The theorem holds for any modulation scheme where real and imaginary parts are IID and the probability functions are symmetric and have finite support. It is particularly true for the M-ary QAM constellations for arbitrary M and the M-ary PSK schemes for $M = 2^m, m > 1$.*

Proof. By independence we have

$$\begin{aligned} Ee^{j\omega\Re\left(\sum_{k=0}^{N-1} c_k e^{j(k\theta+\alpha)}\right)} &= Ee^{j\omega\sum_{k=0}^{N-1}\Re(c_k)\cos(k\theta+\alpha)}Ee^{-j\omega\sum_{k=0}^{N-1}\Im(c_k)\sin(k\theta+\alpha)} \\ &= \prod_{k=0}^{N-1} Ee^{j\omega\Re(c_k)\cos(k\theta+\alpha)}Ee^{-j\omega\Im(c_k)\sin(k\theta+\alpha)} \end{aligned}$$

Since the probability functions are symmetric and have finite support the characteristic functions define analytic functions in a neighborhood of $\omega = 0$ where they are symmetric and strictly positive [70]. Thus, for any real RV x with these properties the characteristic function can be rewritten as

$$Ee^{j\omega x} = e^{g(\omega)}$$

in a neighborhood of $\omega = 0$ where

$$g\left(\omega\right) = \sum_{k=2}^{\infty} a_k \omega^k$$

is a power series in ω with even terms only and $a_2 = -\frac{E\left(x^2\right)}{2}$. Thus

$$\begin{aligned}
Ee^{j\omega\sum_{k=0}^{N-1}\Re(c_k)\cos(k\theta+\alpha)} &= \prod_{k=0}^{N-1} Ee^{j\omega\Re(c_k)\cos(k\theta+\alpha)} \\
&= \prod_{k=0}^{N-1} e^{\frac{\omega^2\cos^2(k\theta+\alpha)}{4}+\mathcal{O}\left(\omega^4\cos^4(k\theta+\alpha)\right)} \\
&= e^{-\frac{\omega^2\sum_{k=0}^{N-1}\cos^2(k\theta+\alpha)}{4}+\mathcal{O}\left(\omega^4\sum_{k=0}^{N-1}\cos^4(k\theta+\alpha)\right)}.
\end{aligned}$$

since the total power is 1. The second term yields

$$Ee^{j\omega\sum_{k=0}^{N-1}\Im(c_k)\sin(k\theta+\alpha)} = e^{-\frac{\omega^2\sum_{k=0}^{N-1}\sin^2(k\theta+\alpha)}{4}+\mathcal{O}\left(\omega^4\sum_{k=0}^{N-1}\sin^4(k\theta+\alpha)\right)}.$$

Combining both yields the first condition in Theorem 17. The second condition is satisfied because for the Nyquist-rate samples

$$\begin{aligned}
&\sum_{k=0}^{N-1}\left(\omega_1\cos\left(k\theta_{l_1,L}\right)+\omega_2\cos\left(k\theta_{l_2}\right)\right)^2+\sum_{k=0}^{N-1}\left(\omega_1\sin\left(k\theta_{l_1,L}\right)+\omega_2\sin\left(k\theta_{l_2}\right)\right)^2 \\
&= \frac{N\left(\omega_1^2+\omega_2^2\right)}{2}, \quad 0\le l_1<N, 0\le l_2<N, l_1\neq l_2.
\end{aligned} \tag{4.9}$$

This immediately implies the M-ary QAM scheme. For the M-ary PSK scheme we have

$$\begin{aligned}
Ee^{j\omega\Re\left(\sum_{k=0}^{N-1}c_k e^{j(k\theta+\alpha)}\right)} &= \prod_{k=0}^{N-1} Ee^{j\omega\cos(k\theta+\alpha+\varphi_k)} \\
&= \prod_{k=0}^{N-1} Ee^{j\omega\cos(k\theta+\alpha+\varphi_k)}.
\end{aligned}$$

where $\cos\left(k\theta+\alpha+\varphi_k\right)$ (recall that RV $\varphi_k \in \left\{0, \frac{2\pi}{M}, \ldots, \frac{2\pi(M-1)}{M}\right\}$) is symmetric due to the assumption $M = 2^m, m > 1$. Since

$$E\left[\cos^2\left(k\theta+\alpha+\varphi_k\right)\right] = \frac{1}{2} \tag{4.10}$$

for any θ, α the first condition follows. The second condition is

$$\begin{aligned}
Ee^{j\Re\left[\sum_{k=0}^{N-1}c_k\ \ \omega_1 e^{jk\theta_{l_1,L}}+\omega_2 e^{jk\theta_{l_2,L}}\ \ \right]} &= Ee^{j\sum_{k=0}^{N-1}\omega_1\cos\left(k\theta_{l_1,L}+\varphi_k\right)+\omega_2\cos\left(k\theta_{l_2,L}+\varphi_k\right)} \\
&= \prod_{k=0}^{N-1} Ee^{j\left[\omega_1\cos\left(k\theta_{l_1,L}+\varphi_k\right)+\omega_2\cos\left(k\theta_{l_2,L}+\varphi_k\right)\right]}.
\end{aligned}$$

The RV in the exponent is symmetric and can be rewritten as

$$Ee^{j\Re\left[\sum_{k=0}^{N-1} c_k \ \ \omega_1 e^{jk\theta_{l_1,L}}+\omega_2 e^{jk\theta_{l_2,L}} \ \ \right]} = \prod_{k=0}^{N-1} Ee^{a_1\left(\omega_1,\omega_2,\theta_{l_1,L},\theta_{l_2,L}\right)\cos\left[\varphi_k+\phi\left(\omega_1,\omega_2,\theta_{l_1,L},\theta_{l_2,L}\right)\right]}$$

where

$$\begin{aligned} a_1\left(\omega_1,\omega_2,\theta_{l_1,L},\theta_{l_2,L}\right) &= \sum_{k=0}^{N-1}\left(\omega_1\cos\left(k\theta_{l_1,L}\right)+\omega_2\cos\left(k\theta_{l_2}\right)\right)^2+\sum_{k=0}^{N-1}\left(\omega_1\sin\left(k\theta_{l_1,L}\right)+\omega_2\sin\left(k\theta_{l_2}\right)\right)^2 \\ &= N\left(\omega_1^2+\omega_2^2\right) \end{aligned}$$

and due to (4.10) the result follows. ■

The result can also be shown for the case where the modulation scheme is such that the constellation points are uniformly distributed on the unit sphere (see next section). The information-theoretic value of this asymptotic analysis is that, in accordance with the results in [9] where the existence of good BPSK or M-ary PSK codes with low CF was investigated, one can expect that even if the assumptions on the constellation are relaxed that a good code has CF of order $\sqrt{\log(N)}$. Furthermore, it is easy to see that the same limit analysis holds when the bandwidth is kept constant. This would reflect practical design criteria since commonly the bandwidth is always limited and must therefore be rescaled when the number of subcarriers goes to infinity.

In the Gaussian case, $c_k \sim \mathcal{CN}(0,1)$, there is also a limit distribution if Nyquist-rate sampling is assumed. The distribution is explicitly given in [83]. For limit distributions in the general case one could calculate higher moments than just first and second moments. We have not attempted such analysis. Furthermore, the approach can be applied to other constellations like cross-constellations as well [84][85]. The derivations for the standard constellations here can be seen as a baseline for estimating the distribution of more general designs.

In Chapter 3 we investigated peak regrowth between the samples so as to get estimates for the CF of the signal from its samples. We concluded that Nyquist-rate sampling is not sufficient and oversampling must be applied. In the case of Nyquist-rate sampling we found that peak regrowth can be of order $\log(N)$. However, we did not consider the statistical properties which we can do now in terms of the distribution of the CF. We see that the worst case is in fact very unlikely for large N. Thus any analysisinvolving Nyquist-rate sampling can make use of this fact.

Finally, we want to discuss some consequences of this analysis for the system designer. First we can ask how to design constellations. Note that the result excludes a broad class of distributions when designing constellation with better CF performance than $\mathcal{O}\left(\sqrt{\log(N)}\right)$. For example, taking the constellation presented in [86] it is easy to see that in the asymptotic regime there is no gain with respect to the distribution over the standard constellations. On the other hand it is easy to see that simply shifted asymmetric constellations (we think of $\{0,1\}$ constellation, for instance) will not work better. For the BPSK case the bounds can be improved using phase shifts. Analytically this can be done by minimizing $\Psi(N,\theta,\alpha)$ in Theorem 12. For the HPA design, the value $\sqrt{\log(N)}$ can be a rough parameter to adjust the OBO of the HPA since most of the non-linear distortion will be evoked by the part below this value for large N (however, a better analysis is given in Sec. 4.3).

4.2 Dependent subcarriers

Because we saw that for independent subcarriers there is almost identical behavior in the asymptotic case for all practical constellations we now introduce and analyze dependent subcarriers. Since in practical systems the dependence between the subcarriers is due to coding and since in practical situations the overall distribution is not known we need to pursue different avenues. We introduce spherical and binary codes.

4.2.1 Spherical codes

Spherical codes are point sets on a sphere in the euclidean N-space. A code is a spherical code if the points represented by the codewords lie on an N-dimensional complex sphere with $\|c\| = \sqrt{N}$ denoted as $\mathbb{S}^{N-1}$. Note that this is in particular satisfied if the underlying constellation $\mathcal{Q}$ is equal-energy. Spherical codes were investigated in the context of OFDM in [9] where bounds on the achievable region of the triplet rate, minimum distance and crest-factor (which is defined by the maximum of the crest-factor over all transmit signals generated by the codewords after

OFDM modulation, see Chapter 5) were derived. In [28, Eqn. (25)] the case of dependent subcarriers was regarded in the sense that the constellation points are uniformly distributed over the sphere (we speak of a spherical distribution of the N-dimensional constellation here) and only Nyquist-rate sampling is considered. It was argued that the spherical distribution can generally serve as a model for M-ary PSK constellations. An error analysis was not given though. Since it is easy to see that the derivation is independent of any time shift an upper bound can be obtained for the continuous-time CF given by

$$F(\lambda) \leq B_s(\lambda) := L \sum_{1 \leq l < \frac{NC_L^2}{\lambda^2}} \binom{N}{l} (-1)^{l+1} \left(1 - \frac{l\lambda^2}{NC_L^2}\right)^{N-1}. \tag{4.11}$$

Note that in light of our union bound the approach may be indeed better motivated because we showed that the characteristic functions of both the model and the real distribution have only to be point-wise close. We do not give an error analysis here (which is possible in our approach) but we want to use this model for an asymptotic analysis. The following lemma generalizes the Gaussian results to the spherical distribution.

Lemma 19 *Let the cumulative distribution of the $c_k, k = 0, \ldots, N-1$, be so that the codewords lie uniformly distributed on $\mathbb{S}^{N-1}$. Then for any $\varepsilon > 0$*

$$\lim_{N \to \infty} \Pr\left(\sqrt{(1-\varepsilon)\log(N)} < \mathrm{CF}(s_c) < \sqrt{(1+\varepsilon)\log(N)}\right) = 0.$$

Proof. Since c is uniformly distributed on $\mathbb{S}^{N-1}$ it can be represented as

$$c = \frac{\sqrt{N} c'}{\|c'\|}$$

where c' has an N-variate spherical distribution, i.e. that it is invariant under multiplication by a unitary matrix. The distribution of c' is not unique and as a representative we can choose $c'_k \sim \mathcal{CN}(0,1)$. In the sequel let $\varepsilon > 0$ be arbitrary and $\varepsilon_1 > 0$ be such that $(1+\varepsilon_1)\sqrt{(1-\varepsilon)} < 1$. At first, we want to prove

$$\lim_{N \to \infty} \Pr\left(\sqrt{(1-\varepsilon)\log(N)} < \mathrm{CF}(s_{c'}) < \sqrt{(1+\varepsilon)\log(N)}\right).$$

Note that this result can be obtained from literature, however, our approach seems appropriate for tight lower and upper bounds. Next, we show that from this it follows the main result.

It is well known that the magnitude is Rayleigh distributed and the Nyquist-rate samples are independent from which it follows the lower bound

$$\Pr\left(\mathrm{CF}\left(s_{c'}\right) \leq \lambda\right) \leq \left(1 - e^{-\lambda^2}\right)^N .$$

Setting $\lambda = \sqrt{(1-\varepsilon)\log(N)}$ for some $\varepsilon > 0$ we have

$$\begin{aligned}\Pr\left(\mathrm{CF}\left(s_{c'}\right) \leq \sqrt{(1-\varepsilon)\log(N)}\right) &\leq \left(1 - \frac{1}{N^{1-\varepsilon}}\right)^N \\ &= \left(1 - \frac{N^\varepsilon}{N}\right)^N\end{aligned}$$

Since $\log(1-x) \leq -x$ we have $\Pr\left(\mathrm{CF}\left(s_{c'}\right) \leq \sqrt{(1-\varepsilon)\log(N)}\right) = e^{-N^\varepsilon}$ which is the desired lower bound.

The upper bound is obtained by observing that for any $L \geq 1$ we have

$$\Pr\left(\mathrm{CF}_L\left(s_{c'}\right) > \lambda\right) \leq L\left[1 - \left(1 - e^{-\lambda^2}\right)^N\right].$$

This can be concluded from the autocorrelation function [30] which is independent from any shift. Hence

$$\Pr\left(\mathrm{CF}\left(s_{c'}\right) > \frac{\lambda}{C_L}\right) \leq L\left[1 - \left(1 - e^{-\frac{\lambda^2}{C_L^2}}\right)^N\right]$$

for $L > 1$. Setting this time $\lambda = \sqrt{(1+\varepsilon)\log(N)}$ for some $\varepsilon > 0$ yields

$$\Pr\left(\mathrm{CF}\left(s_{c'}\right) > \sqrt{(1+\varepsilon)\log(N)}\right) \leq L\left[1 - \left(1 - \frac{1}{N^{\frac{(1+\varepsilon)}{C_L^2}}}\right)^N\right]$$

Since $\frac{1}{C_L^2} \geq 1 - \frac{2\pi^2}{8L^2}$ we can choose L so that $\frac{(1+\varepsilon)}{C_L^2} = 1 + \varepsilon' > 1$ and

$$\Pr\left(\mathrm{CF}\left(s_{c'}\right) > \sqrt{(1+\varepsilon)\log(N)}\right) \leq L\left[1 - \left(1 - \frac{N^{-\varepsilon'}}{N}\right)^N\right].$$

Thus, for any $\mu > 0$

$$\Pr\left(\mathrm{CF}\left(s_{c}\right) > \sqrt{(1+\varepsilon)\log(N)}\right) \leq L\left[1 - \left(1 - \frac{\mu}{N}\right)^N\right]$$

for N large enough and

$$\limsup_{N\to\infty} \Pr\left(\mathrm{CF}\left(s_{c'}\right) > \sqrt{(1+\varepsilon)\log(N)}\right) \le L\left[1-e^{-\mu}\right].$$

which is the desired upper bound since for fixed L the right hand side can be made arbitrarily small.

We consider now the first inequality of the lemma. By the law of total probability

$$\begin{aligned}
&\Pr\left(\left\{\mathrm{CF}\left(s_c\right) < \sqrt{(1-\varepsilon)\log(N)}\right\}\right)\\
&= \Pr\left(\left\{\mathrm{CF}\left(s_c\right) < \sqrt{(1-\varepsilon)\log(N)}\right\} \cap \left\{\|c\| < \left(1-\varepsilon_1\right)\sqrt{N}\right\}\right)\\
&\quad + \Pr\left(\left\{\mathrm{CF}\left(s_c\right) < \sqrt{(1-\varepsilon)\log(N)}\right\} \cap \left\{\left(1-\varepsilon_1\right)\sqrt{N} < \|c\| < \left(1+\varepsilon_1\right)\sqrt{N}\right\}\right)\\
&\quad + \Pr\left(\left\{\mathrm{CF}\left(s_c\right) < \sqrt{(1-\varepsilon)\log(N)}\right\} \cap \left\{\|c\| > \left(1+\varepsilon_1\right)\sqrt{N}\right\}\right).
\end{aligned}$$

Since both $\lim_{N\to\infty} \Pr\left(\|c\| < \left(1-\varepsilon_1\right)\sqrt{N}\right) = 0$ and $\lim_{N\to\infty} \Pr\left(\|c\| > \left(1+\varepsilon_1\right)\sqrt{N}\right) = 0$ for every $\varepsilon_1 > 0$, it is left to show that

$$\lim_{N\to\infty} \Pr\left(\left\{\mathrm{CF}\left(s_c\right) < \sqrt{(1-\varepsilon)\log(N)}\right\} \cap \left\{\left(1-\varepsilon_1\right)\sqrt{N} < \|c\| < \left(1+\varepsilon_1\right)\sqrt{N}\right\}\right) = 0$$

We have

$$\begin{aligned}
&\Pr\left(\left\{\mathrm{CF}\left(s_c\right) < \sqrt{(1-\varepsilon)\log(N)}\right\} \cap \left\{\left(1-\varepsilon_1\right)\sqrt{N} < \|c\| < \left(1+\varepsilon_1\right)\sqrt{N}\right\}\right)\\
&= \Pr\left(\left\{\max_{0\le\theta<2\pi}\left|\sum_{k=0}^{N-1}\frac{c_k e^{jk\theta}}{\|c\|}\right| < \sqrt{(1-\varepsilon)\log(N)}\right\} \cap \left\{\left(1-\varepsilon_1\right)\sqrt{N} < \|c\| < \left(1+\varepsilon_1\right)\sqrt{N}\right\}\right)\\
&\le \Pr\left(\frac{1}{\left(1+\varepsilon_1\right)\sqrt{N}}\max_{0\le\theta<2\pi}\left|\sum_{k=0}^{N-1} c_k e^{jk\theta}\right| < \sqrt{(1-\varepsilon)\log(N)}\right)\\
&= \Pr\left(\frac{1}{\sqrt{N}}\max_{0\le\theta<2\pi}\left|\sum_{k=0}^{N-1} c_k e^{jk\theta}\right| < \left(1+\varepsilon_1\right)\sqrt{(1-\varepsilon)\log(N)}\right)
\end{aligned}$$

from which the result follows. The proof for the upper bound is similar. ■

The theorem suggests that for high rate spherical codes, asymptotically, one cannot expect better performance than CF of order $\mathcal{O}\left(\sqrt{\log(N)}\right)$. On the other hand, optimistically speaking, the performance will at least be no worse than the performance in the independent case.

The theorem has some interesting implications for an important subset of spherical codes, the so-called group codes. Group codes were introduced by Slepian in [87] and form a subset of spherical codes. They are generated by a group G of $N \times N$ orthogonal matrices and a real initial vector a. Group codes do exist in great abundance and have been shown to possess good communication properties in AWGN [87]. We note that for example a 2^m-ary PSK modulation, $m > 1$, [88] or permutation modulation [89] can be described by group codes. Thus, the reader unfamiliar with group codes may think of 2^m-ary PSK modulation although we have adopted a broader viewpoint.

A group code $\mathcal{C}_a$ is defined by the set of vectors $c^{(m)} = Q^{(m)}a, m = 0, \ldots, M_1 - 1$, where the Q_m constitute the group and a is the initial vector. The characteristics of these codes depend strongly on the choice of the initial vector and there has been a great deal of research into finding algorithms for calculating optimum initial vectors. The next theorem shows that it is also suited to adjusting the CF.

Theorem 20 *Given any group code and arbitrary λ there is an initial vector so that*

$$F(\lambda) \leq B_s(\lambda)$$

where B_s is the distribution from (4.11).

Proof. Define the set

$$\mathcal{A} := \left\{ c \in \mathbb{S}^{N-1} : \mathrm{CF}(s_c) > \lambda \right\}.$$

The term $\Pr(\mathrm{CF}(s_c) > \lambda)$ is then given by

$$\begin{aligned} \Pr(\mathrm{CF}(s_c) > \lambda) &= \Pr(c \in \mathcal{C}_a : \mathrm{CF}(s_c) > \lambda) \\ &= \frac{1}{M_1} \sum_{m=0}^{M_1-1} I_{\mathcal{A}}\left(c^{(m)}\right) \end{aligned}$$

Using this expression we have

$$E_a\left(\Pr(\mathrm{CF}(s_c) > \lambda)\right) = \frac{1}{M_1} \sum_{m=0}^{M_1-1} E_a I_{\mathcal{A}}\left(c^{(m)}\right)$$

$$
\begin{aligned}
&= \frac{1}{M_1} \sum_{m=0}^{M_1-1} E_a I_{\mathcal{A}} \left(Q^{(m)} a \right) \\
&= \frac{1}{M_1} \sum_{m=0}^{M_1-1} B_s (\lambda) \\
&= B_s (\lambda)
\end{aligned}
$$

where E_a means averaging with respect to a. ■

Consequently, the result proves that we can always adjust the initial vector such that the CF is of the same order than in the uncoded case. However, this may affect rate and distance properties (the initial vector may not even be real). This general trade-off is treated in Chapter 5. We now turn our attention to binary codes.

4.2.2 Binary codes

We assume here binary codes with the mapping $0 \to 1$ and $1 \to -1$. We give a bound in terms of the distance distribution

$$B_k := \frac{1}{M_1} \left| \{ (c_1, c_2) : d_H (c_1, c_2) = k, c_1 \in \mathcal{C}, c_2 \in \mathcal{C} \} \right| \tag{4.12}$$

where $\mathcal{C}$ is some arbitrary binary code with M_1 codewords and $d_H(x, y)$ is the Hamming distance between two codewords which is the number of different coordinates of two codewords.

Note that the distance distribution coincides with the weight distribution for linear codes, i.e.

$$W_k := \left| \{ c : w (c) = k, c \in \mathcal{C} \} \right|$$

where $w (c)$ is the weight of the codeword. Furthermore if the code contains the all-one codeword we have $W_k = W_{N-k}$.

Given an integer j let $i = (i_0, i_1, \ldots, i_{N-1})$ be a non-negative, integer-valued vector of length N satisfying $i_0 + i_1 + \ldots + i_{N-1} = j$ and let $\mathcal{I}_j$ denote the set of all such vectors. Define the moments

$$m_i := E \left(c_0^{i_0} c_1^{i_1} \ldots c_{N-1}^{i_{N-1}} \right), \tag{4.13}$$

and for the sake of clarity

$$b_i^{(j)} := \frac{j}{i_0! i_1! \cdots i_{N-1}!}, \tag{4.14}$$

and

$$k_i(\theta, \alpha) := \left(\cos^{i_0}(\alpha) \cos^{i_1}(\theta + \alpha) \cos^{i_2}(2\theta + \alpha) \ldots \cos^{i_{N-1}}[(N-1)\theta + \alpha]\right). \tag{4.15}$$

Expanding the exponential function in the Chernoff bound yields

$$E\left(e^{\varepsilon \Re\ s_c(\theta_{l_1,L}) e^{j\alpha_{l_2,K}}}\right) = \sum_{j=0}^{N_1} \frac{\varepsilon^j}{j!} \sum_{\mathbf{i} \in \mathcal{I}_j} b_i^{(j)} m_i k_i(\theta_{l_1,L}, \alpha_{l_2,K}) + \frac{\varepsilon^{N_1+1} N^{N_1+1} e^{\varepsilon N}}{(N_1+1)!}$$

where the error term is given by Taylor's theorem. Then, exchanging integration and summation, for any $\varepsilon > 0$ the complementary distribution of the CF is upperbounded by

$$F(\lambda) \leq \min_{L>1, K>3} \sum_{l_1=0}^{LN-1} \sum_{l_2=0}^{K-1} e^{-\frac{\varepsilon\sqrt{N}\lambda}{C_L C_K}} \sum_{j=0}^{N_1} \frac{\varepsilon^j}{j!} \sum_{\mathbf{i} \in \mathcal{I}_j} b_i^{(j)} m_i k_i(\theta_{l_1,L}, \alpha_{l_2,K}) + \frac{\varepsilon^{N_1+1} N^{N_1+1} e^{\varepsilon N}}{(N_1+1)!}.$$

However, from a practical point of view the moments bound is rather unwieldy to evaluate. On the other hand, we can obtain simpler expressions for linear, binary codes.

Theorem 21 *Let $\mathcal{C}$ be a linear, binary code containing the all-one codeword. Then the complementary distribution of the CF is upperbounded by*

$$F(\lambda) \leq \min_{\varepsilon>0} \sum_{k=0}^{N} \frac{f_N(\varepsilon, \lambda) W_k \cosh(\varepsilon(N-2k))}{M_1} \tag{4.16}$$

with

$$f_N(\varepsilon, \lambda) := \min_{L>1, K>3} LNKe^{-\frac{\varepsilon\sqrt{N}\lambda}{C_L C_K}} \tag{4.17}$$

Proof. Fixing $\varepsilon > 0$ we know that

$$\begin{aligned} & F(\lambda) \\ & \leq \min_{L>1, K>3} \sum_{l_1=0}^{LN-1} \sum_{l_2=0}^{K-1} e^{-\frac{\varepsilon\sqrt{N}\lambda}{C_L C_K}} \sum_{j=0}^{N_1} \frac{\varepsilon^j}{j!} \sum_{\mathbf{i} \in \mathcal{I}_j} b_i^{(j)} m_i k_i(\theta_{l_1,L}, \alpha_{l_2,K}) + \frac{\varepsilon^{N_1+1} N^{N_1+1} e^{\varepsilon N}}{(N_1+1)!}. \end{aligned}$$

By Cauchy's inequality

$$
\begin{aligned}
&F(\lambda) \\
&\leq \min_{L>1,K>3} \sum_{l_1=0}^{LN-1} \sum_{l_2=0}^{K-1} e^{-\frac{\varepsilon\sqrt{N}\lambda}{C_L C_K}} \sum_{j=0}^{N_1} \frac{\varepsilon^j}{j!} \max_{i \in I_j} \left|k_i\left(\theta_{l_1,L}, \alpha_{l_2,K}\right)\right| \sum_{i \in \mathcal{I}_j} b_i^{(j)} m_i + \frac{\varepsilon^{N_1+1} N^{N_1+1} e^{\varepsilon N}}{(N_1+1)!} \\
&= \sum_{j=0}^{N_1} \frac{f_N(\varepsilon,\lambda)\,\varepsilon^j}{j!} \sum_{i \in \mathcal{I}_j} b_i^{(j)} m_i + \frac{\varepsilon^{N_1+1} N^{N_1+1} e^{\varepsilon N}}{(N_1+1)!}
\end{aligned}
$$

provided that $m_i \geq 0, i \in \mathcal{I}_j$, which is indeed the case for linear codes. This can be seen as follows:

The sum $\sum_{c \in \mathcal{C}} \prod_{k=0}^{N-1} c_k^{i_k}$ is just the sum over all codewords of $(-1)^p$ where p is the parity of the subvectors on the corresponding positions of the codewords. Whatever subsets of positions one picks on these positions the parity of the subvectors is either all of even parity or half the subvectors are of even or half the vectors are of odd parity since the sum of two subvectors with different (equal) parity yields odd (even) parity and since the code is linear. Thus, the sum is either zero or N.

Now, we need to replace the term

$$
m_i = E\left(c_0^{i_0} c_1^{i_1} \ldots c_{N-1}^{i_{N-1}}\right) = \frac{1}{M_1} \sum_{c \in \mathcal{C}} \prod_{k=0}^{N-1} c_k^{i_k}
$$

in terms of the weight distribution of the linear code. Note that for $c \in \{\pm 1\}^N$, we have

$$
\sum_{k=0}^{N-1} c_k = N - 2w(c)
$$

where $w(c)$ is the weight of the codeword c. Consider the following sum

$$
s_j = \sum_{c \in \mathcal{C}} \left(\sum_{k=0}^{N-1} c_k\right)^j .
$$

On the one hand it is

$$
\begin{aligned}
s_j &= \sum_{c \in \mathcal{C}} (N - 2w(c))^j \\
&= \sum_{k=0}^{N} (N - 2k)^j W_k
\end{aligned}
$$

On the other hand

$$
\begin{aligned}
s_j &= \sum_{c\in\mathcal{C}} \left(\sum_{k=0}^{N-1} c_k\right)^j \\
&= \sum_{c\in\mathcal{C}} \sum_{i\in I_j} b_i^{(j)} \prod_{k=0}^{N-1} c_k^{i_k} \\
&= \sum_{i\in\mathcal{I}_j} b_i^{(j)} \sum_{c\in\mathcal{C}} \prod_{k=0}^{N-1} c_k^{i_k}.
\end{aligned}
$$

Thus

$$\sum_{i\in\mathcal{I}_j} b_i^{(j)} \sum_{c\in\mathcal{C}} \prod_{k=0}^{N-1} c_k^{i_k} = \sum_{k=0}^{N} (N-2k)^j W_k$$

and

$$
\begin{aligned}
F(\lambda) &\le \sum_{j=0}^{N_1} \frac{f_N(\varepsilon,\lambda)\,\varepsilon^j}{j!} \sum_{i\in\mathcal{I}_j} b_i^{(j)} m_i + \frac{\varepsilon^{N_1+1} N^{N_1+1} e^{\varepsilon N}}{(N_1+1)!} \\
&= \sum_{j=0}^{N_1} \frac{f_N(\varepsilon,\lambda)\,\varepsilon^j}{M_1 j!} \sum_{k=0}^{N} (N-2k)^j W_k + \frac{\varepsilon^{N_1+1} N^{N_1+1} e^{\varepsilon N}}{(N_1+1)!}
\end{aligned}
$$

Observing that the sum is zero for odd j and exchanging the order of summation

$$
\begin{aligned}
F(\lambda) &\le \sum_{k=0}^{N} \frac{f_N(\varepsilon,\lambda)\,W_k}{M_1} \sum_{j=0}^{N_1} \frac{\varepsilon^{2j}(N-2k)^{2j}}{(2j)!} + \frac{\varepsilon^{N_1+1} N^{N_1+1} e^{\varepsilon N}}{(N_1+1)!} \\
&\le \limsup_{N\to\infty} \sum_{k=0}^{N} \frac{f_N(\varepsilon,\lambda)\,W_k}{M_1} \sum_{j=0}^{N_1} \frac{\varepsilon^{2j}(N-2k)^{2j}}{(2j)!} + \frac{\varepsilon^{N_1+1} N^{N_1+1} e^{\varepsilon N}}{(N_1+1)!}
\end{aligned}
$$

Since N_1 is arbitrary and

$$\sum_{j=0}^{N_1} \frac{\varepsilon^{2j}(N-2k)^{2j}}{(2j)!} \to \cosh\left[\varepsilon(N-2k)\right]$$

and

$$\frac{\varepsilon^{N_1+1} N^{N_1+1} e^{\varepsilon N}}{(N_1+1)!} \to 0$$

as $N_1 \to \infty$ yields the desired result. ∎

The optimization problem in (4.17) can be easily solved using simple line search procedures. The optimization parameter in (4.16) must be also numerically computed. Here, we choose a good starting point from the BPSK case where the optimal ε can be analytically obtained.

Since the proof of the bound relies on the fact that the moments are non-negative, the bound does not apply to general non-linear codes. However, a simple calculation shows that the bound can also be written in terms of distance distribution.

Corollary 22 *For any binary code the (complementary) distribution of the CF is upperbounded by*

$$F(\lambda) \leq \min_{\varepsilon>0} \sum_{j=0}^{N_1} \frac{f_N(\varepsilon,\lambda)\,\varepsilon^j N^{\frac{j}{2}}}{M_1 j!} \left(\sum_{k=0}^{N} (N-2k)^j B_k\right)^{\frac{1}{2}} + \frac{\varepsilon^{N_1+1} N^{N_1+1} e^{\varepsilon N}}{(N_1+1)!}.$$

Proof. Observe that generally

$$\begin{aligned} s_j &= \sum_{c_1\in\mathcal{C}} \sum_{c_2\in\mathcal{C}} (N-2d_H(c_1,c_2))^j \\ &= \sum_{k=0}^{N} (N-2k)^j B_k \\ &= \sum_{i\in I_j} b_i^{(j)} \left(\sum_{c\in\mathcal{C}} \prod_{k=0}^{N-1} c_k^{i_k}\right)^2. \end{aligned}$$

Applying Cauchy's inequality as in Theorem 21 and observing that

$$\sum_{\mathbf{i}\in\mathcal{I}_j} k_i^2\left(\theta_{l_1,L},\alpha_{l_2,K}\right) \leq N^j$$

yields the result. ∎

Observe that the error term in (4.16) can be made arbitrarily small by simply considering more terms in the sum. If $B_k = B_{N-k}$ then the second sum in (4.16) is zero for odd k. However, we point out that a numerical evaluation is difficult for $N \geq 512$.

The bound can also be generalized with respect to other modulation schemes like QPSK. Obviously, to bound the distribution expressions or bounds on the weight distribution must be known. In the past many authors have dealt with the topic (see for example [90, 91, 92]) and

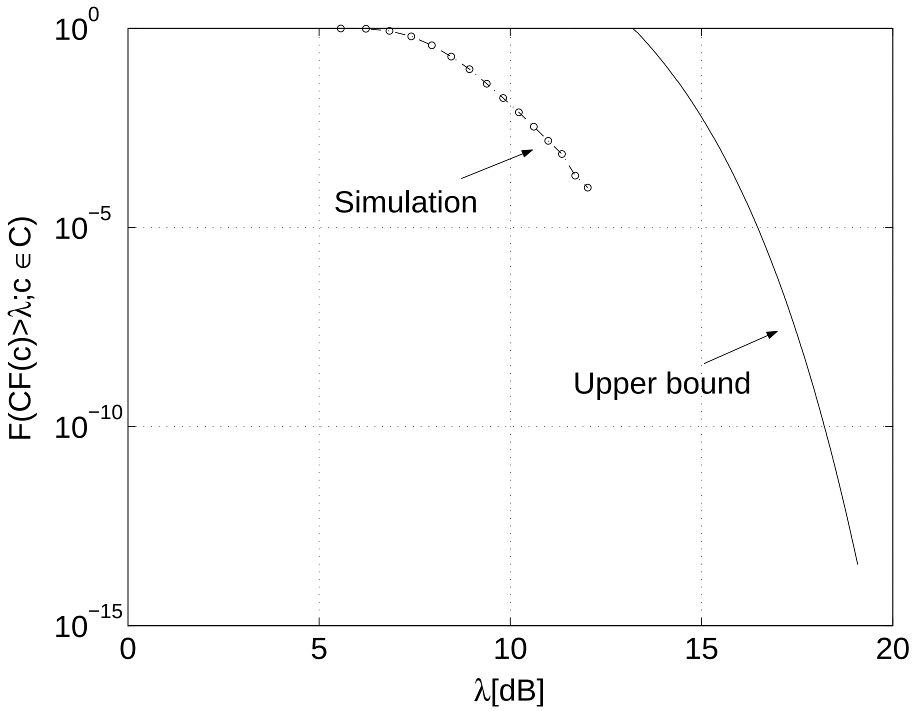

Figure 4.3: Complementary distribution and upper bound of BCH code. The simulation is based on 10^5 OFDM symbols. Note that the simulation is extremely time-consuming.

many expressions are known. Let us illustrate the results with an example. In Fig. 4.3 the complementary CF distribution of the BCH code of length $N = 255$ and 215 information bits is shown. The two curves correspond to the simulation of 10^5 OFDM symbols along with a bound computed from the known distance distribution. Note that here $B_0 = 1, B_1 = \ldots = B_{10} = 0, B_{11} \ldots B_{18} \lesssim 2^{76}$ and $B_i \approx 2^{-40}\binom{N}{i}$ [93]. The bound is relatively loose but note that N is small in the OFDM context and due to the Chernoff bounding technique the bound will improve for larger N as in the independent case. On the other hand the example proves that the effective CF is roughly 5.5 [dB] below the maximum possible value $\sqrt{N} \approx 24$ [dB] given an outage probability of $\epsilon = 10^{-12}$. Moreover, we observe from (4.16) that an exponential decay is only possible if the number of codewords grows exponentially with N.

The bound suggests some asymptotic conclusions regarding Sidelnikov's bound for some t-error correcting BCH codes of lenght N with m information bits. Sidelnikov proved the following result. For $0 < t < \sqrt{\frac{N}{10}}$ and

$$2t + v_N(t) \leq j \leq N - 2t - v_N(t),$$

where $v_N(t) \to 1$ as $N \to \infty$ the number of codewords of weight j is given by

$$W_j = 2^{mt} \binom{N}{j} (1 + E_j)$$

with

$$|E_j| < KN^{-0.1}, \quad K > 0,$$

indicating that, in the asymptotic case, the CF of BCH codes is lower than $\sqrt{2\log(N)}$ by comparing the Chernoff bound with that of BPSK modulation.

4.3 Bounding the symbol error rate

The analysis in the last section showed that a certain CF cannot be avoided. Assuming the soft limiter model from Chapter 1 then, if the OBO of the power amplifier is lower than the maximum CF, an additional SER is evoked by clipping. In practice, the OFDM system parameters are adjusted by simulations, as for example the OBO of the HPA minimizing the total degradation trading off power efficiency for SER [59][60]. Obviously, analytical expressions would permit simplifying the system design procedure. The standard approach for evaluating the SER evoked by clipping assumes Gaussian distribution both of the complex envelope and of the distortion terms after the DFT at the receiver [94, 95, 96, 60] (the Gaussian approach hereafter). However, this model neglects the fact that clipping is a rare event for higher OBO's (i.e. when the desired error probability is low) occurring more infrequently than once every symbol interval. In [31, 36, 35, 37] it is shown that under these circumstances the SER is significantly underestimated. A further disadvantage is that the scope of this approach is not clear in terms of the number of subcarriers, the constellation size [36]. In fact, the SER is independent of N and a theoretical justification that the approximation improves when N is large is still missing. Furthermore, real systems may use oversampling, discrete-time or continuous-time filtering in order to reduce noise power [97]. Note that if oversampling is applied at the receiver the Gaussian approximation fails to hold since the samples fed into the DFT are correlated. Oversampling is therefore not included in the analysis of [94, 95, 96, 60] (but it may be considered by using an appropriate HPA model [98]). The effect is usually taken advantage of by deliberately clipping and filtering in the discrete-time domain, since filtering after the HPA is costly. It is thus the simplest method to reduce clipping distortion.

By maintaining the Gaussian distribution of the complex envelope in [36] a Chernoff upper bound on the SER was derived. Although this will give ensured upper bounds on the SER (including asymptotic results) the method is generally confined to the low probability region for large N and approximative SER curves cannot be obtained. Moreover, the possibility of extensions to true upper bounds on the SER, dependent on the different constellations that are used in OFDM transmission seems questionable due to the computational effort. Another problem is that the approach does not give much insight into the mechanism of the clipping effect. For example, it is difficult to relate the bounds in [36] to the distribution of the CF. An interesting approach to the clipping phenomenon was presented in [31] elaborating more on the impulsive nature of the clipping process. The envelope of the (distortion) pulse above the clipping level is described by a parametrized parabolic arc and the effect on the SER is studied by its spectral properties. However, the description of the pulse is simplifying and insufficient as the shape is described by only one random parameter of which the distribution is obtained from an asymptotic analysis of Gaussian random processes (even though the underlying process is Rayleigh). Thus, the approach lacks of some theoretical justification and, furthermore, simulations are not very encouraging. We would like to point out that our new approach can be interpreted as putting the approach in [31] on a firm mathematical basis. Generally, we note that basically as in Chapter 4 we have to bound sums of RV's but this time (except for Gaussian) they are dependent and its distribution can only be numerically calculated. Moreover, the numerical effort to bound these expressions will very quickly become large. Thus, it is of most interest to extract the quantities that contribute most to the terms in question. The following section gives a partial answer to this question in terms of the distribution of the CF. In fact, the approach is much better than that of [31] and can substitute the Gaussian approach for small numbers of subcarriers and larger constellations. For larger numbers of subcarriers we propose the Chernoff bound approach in [36] (in combination with the results here). It is worth mentioning that we include oversampling in our analysis thereby approximating discrete-time clipping and filtering (as well as approximately continuous-time filtering for large oversampling rates). Note that an appropriate model for the clipping analysis is crucial.

Note that we do not comment on the effects of clipping on the power density spectrum since even though the problems discussed in this chapter are in principle the same, they turn out to be of less importance when applying the Gausssian approach. For details see [99, 100, 101, 94, 102].

4.3.1 An overall SER bound

The following theorem combines the distribution of the CF and the distribution of the number of samples exceeding the clipping level to give an overall bound on the SER due to clipping. We assume the receiver to operate with oversampling rate L. The power of the M-ary QAM constellation and the transmit signal is normalized to one.

Theorem 23 *Given a clipped OFDM signal then the average SER is upperbounded by*

$$\begin{aligned}\overline{\mathrm{SER}} &\leq \frac{1}{N}\sum_{k=0}^{N-1}\min_{\mu_k>\lambda}\left[F_L(\mu_k)+4\Pr\left(\xi_c^{(k)}(\lambda,\{\theta_{l,L}\})\geq N_L(\lambda,\mu_k)\right)\right]\\ &\leq \frac{1}{N}\sum_{k=0}^{N-1}\min_{\mu_k>\lambda}\left[F_L(\mu_k)+4\sum_{m=0}^{L-1}\Pr\left(\xi_c^{(k)}\left(\lambda,\left\{\theta_{l,L}^{(m)}\right\}\right)\geq\left\lceil\frac{N_L(\lambda,\mu_k)}{L}\right\rceil\right)\right]\end{aligned}$$

where $\theta_{l,L}^{(m)}$ is defined below in (4.18),

$$\xi_c^{(k)}(\lambda,\{\theta_{l,L}\}):=\left|\left\{\theta_{l,L}:|s_c(\theta_{l,L})|>\lambda;\arg\left[d_c(\theta_{l,L})\right]+\frac{2\pi k}{N}\in\left[-\frac{\pi}{2},\frac{\pi}{2}\right),l=0,\ldots,LN-1\right\}\right|$$

and

$$N_L(\lambda,\mu)=\left\lceil\frac{L\sqrt{N}A}{\mu-\lambda}\right\rceil.$$

Proof. The event ,,symbol error on subcarrier 0", i.e. $\{c_0 \neq \widehat{c}_0\}$ can be partioned into disjoint events $\{c_0 \neq \widehat{c}_0\} \cap \{\mathrm{CF}_L(s_c) > \mu\}$ or $\{c_0 \neq \widehat{c}_0\} \cap \{\mathrm{CF}_L(s_c) \leq \mu\}$. Thus

$$\Pr(c_0 \neq \widehat{c}_0) \leq F_L(\mu) + \Pr(\{c_0 \neq \widehat{c}_0\} \cap \{\mathrm{CF}_L(s_c) \leq \mu\}).$$

Next we need to calculate the term $\Pr(\{c_0 \neq \widehat{c}_0\} \cap \{\mathrm{CF}(s_c) \leq \mu\})$. The samples after the non-linearity can be decomposed into $\Phi_{\mathrm{SL}}(s_c(\theta)) = s_c(\theta) + d_c(\theta)$ (we call $d_c(\theta)$ the distortion samples) and clearly the event $\{c_0 \neq c_0\}$ after the DFT using oversampling L is contained in the event

$$\left\{ \pm \sum_{l=0}^{LN-1} |d_c(\theta_{l,L})| \cos(\arg[d_c(\theta_{l,L})]) \geq L\sqrt{N}A \right\}$$

and

$$\left\{ \pm \sum_{l=0}^{LN-1} |d_c(\theta_{l,L})| \sin(\arg[d_c(\theta_{l,L})]) \geq L\sqrt{N}A \right\}$$

for the M-ary QAM constellations [35] where

$$A = \sqrt{\frac{3}{2(M-1)}}.$$

Since the events differ only in a phase shift of $\frac{\pi}{2}$ and π, and the phase shifts caused by the different subcarriers is a multiple of $\frac{\pi}{2}$, it is sufficient to consider the cosine term and incorporate the others simply by the union bound. We have

$$\begin{aligned}
&\left\{ \sum_{l=0}^{LN-1} |d_c(\theta_{l,L})| \cos(\arg[d_c(\theta_{l,L})]) > L\sqrt{N}A \right\} \cap \{\mathrm{CF}_L(s_c) \leq \mu\} \\
&\subseteq \left\{ \sum_{l=0}^{LN-1} \min\{|d_c(\theta_{l,L})|, \mu - \lambda\} \cos(\arg[d_c(\theta_{l,L})]) > L\sqrt{N}A \right\} \\
&\equiv \bigcup_{m=1}^{LN-1} \left\{ \sum_{l=0}^{LN-1} \min\{|d_c(\theta_{l,L})|, \mu - \lambda\} \cos(\arg[d_c(\theta_{l,L})]) > L\sqrt{N}A \right\} \cap \left\{\xi_c^{(0)}(\lambda, \{\theta_{l,L}\}) = m\right\} \\
&\equiv \bigcup_{m=1}^{LN-1} \left\{ c \in \mathcal{A}_m : \sum_{l=0}^{LN-1} \min\{|d_c(\theta_{l,L})|, \mu - \lambda\} \cos(\arg[d_c(\theta_{l,L})]) > L\sqrt{N}A \right\}
\end{aligned}$$

where

$$\mathcal{A}_m := \left\{c : \xi_c^{(0)}(\lambda, \{\theta_{l,L}\}) = m\right\}.$$

For any $m > 0$ and $c \in \mathcal{A}_m$ define the set

$$\mathcal{M}_m^{(c)} := \left\{\theta_{l,L} : \left|d_c\left(\theta_{l_j}\right)\right| > 0; \arg\left[d_c\left(\theta_{l_j}\right)\right] \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right), j = 0, \ldots, m-1\right\}$$

Then

$$\begin{aligned}
&\bigcup_{m=1}^{LN-1}\left\{c \in \mathcal{A}_m : \sum_{l=0}^{LN-1} \min\left\{\left|d_c\left(\theta_{l,L}\right)\right|, \mu-\lambda\right\}\cos\left(\arg\left[d_c\left(\theta_{l,L}\right)\right]\right) > L\sqrt{N}A\right\} \\
&\equiv \bigcup_{m=1}^{LN-1}\left\{c \in \mathcal{A}_m : \sum_{\theta_{l,L} \in \mathcal{M}_m^{(c)}} \min\left\{\left|d_c\left(\theta_{l,L}\right)\right|, \mu-\lambda\right\} > L\sqrt{N}A\right\} \\
&\subseteq \bigcup_{m=1}^{LN-1}\left\{c \in \mathcal{A}_m : \sum_{\theta_{l,L} \in \mathcal{M}_m^{(c)}} \left(\mu-\lambda\right) > L\sqrt{N}A\right\} \\
&\equiv \bigcup_{m=1}^{LN-1}\left\{c \in \mathcal{A}_m : \left|\left\{\mathcal{M}_m^{(c)}\right\}\right|\left(\mu-\lambda\right) > L\sqrt{N}A\right\} \\
&\equiv \bigcup_{m=\left\lceil \frac{L\sqrt{N}A}{\mu-\lambda}\right\rceil}^{LN-1}\left\{c \in \mathcal{A}_m\right\}
\end{aligned}$$

The probability measure of the quantity on the right hand side is just the probability of the event that the number of samples above the clipping level exceeds $\left\lceil \frac{L\sqrt{N}A}{\mu-\lambda}\right\rceil$.

Define

$$\left\{\theta_{l,L}^{(m)}\right\} := \left\{\theta_{0\cdot L+m,L}, \theta_{1\cdot L+m,L}, \ldots, \theta_{(N-1)\cdot L+m,L}\right\} \tag{4.18}$$

for $m = 0, \ldots, L-1$, then

$$\begin{aligned}
\left\{\xi_c^{(0)}\left(\lambda, \left\{\theta_{l,L}\right\}\right) \geq N_L\left(\lambda, \mu\right)\right\} &= \left\{\sum_{m=0}^{L-1} \xi_c^{(0)}\left(\lambda, \left\{\theta_{l,L}^{(m)}\right\}\right) \geq N_L\left(\lambda, \mu\right)\right\} \\
&\subseteq \bigcup_{m=0}^{L-1}\left\{\xi_c^{(0)}\left(\lambda, \left\{\theta_{l,1}^{(m)}\right\}\right) \geq \left\lceil \frac{N_L\left(\lambda, \mu\right)}{L}\right\rceil\right\}
\end{aligned}$$

which is the second inequality. The theorem follows by introducing the appropriate phase shifts for the other subcarriers. ∎

Obviously, we apply a balancing method between the distribution of the CF and the number of samples that exceed a given level. We note that the theorem can easily be generalized to the case where the non-linearity comes after the linear channel.

Obviously, in order to make this theorem workable an upper bound on $\xi_c^{(k)}(\lambda, \{\theta_{l,L}\})$ is needed. Generally the problem of bounding $\xi_c^{(k)}(\lambda, \{\theta_{l,L}\})$ is related to extreme value and large deviations theory in probability theory [83]. Although extensively studied, results for arbitrary processes are still rare and we cannot hope for a complete solution to this problem. A famous result goes back to Rice for stationary processes where it was shown that the level-crossing density, i.e. the average number of crossings per unit time, decreases exponentially with the crossing level. However, regarding the instationary processes considered here no general results are known. We can at least obtain the following upper bound.

Lemma 24 *The complementary distribution of* $\xi_c^{(k)}(\lambda, \{\theta_{l,L}\})$ *is upperbounded by*

$$\Pr\left(\xi_c^{(k)}(\lambda, \{\theta_{l,L}\}) > x\right) \leq \sum_{l=0}^{LN-1} \frac{\Pr\left(|s_c(\theta_{l,L})| > \lambda\right)}{x}, \quad x > 0.$$

Proof. Define the event

$$\mathcal{M}(\lambda, \theta) := \{c : |s_c(\theta)| > \lambda\}.$$

Then by using the set function I

$$\xi_c^{(k)}(\lambda, \{\theta_{l,L}\}) \leq \sum_{l=0}^{LN-1} I_{\mathcal{M}\left(\lambda, \theta_{l,L}\right)}(c)$$

and taking expectation E of this quantity yields

$$E\left(\xi_c^{(k)}(\lambda), \{\theta_{l,L}\}\right) = \sum_{l=0}^{LN-1} E\left(I_{\mathcal{M}\left(\lambda, \theta_{l,L}\right)}(c)\right) = \sum_{l=0}^{LN-1} \Pr\left(|s_c(\theta_{l,L})| > \lambda\right).$$

The upper bound follows from Markov's inequality. ■

Bounds on the instantaneous distribution of the signal envelope are given in Theorem 12. Both can now be combined to give an upper bound on the average SER for arbitrary OFDM signals. Although the bound is extremely weak (the distribution is estimated by its mean only) we found it interesting enough to include it in here since we regard it as a first non-trivial upper bound on the SER and gives interesting asymptotic conclusions, for example about the clipped energy rather than the peak value (see next section). Furthermore, the approach gives (a posteriori) a kind of justification for methods alleviating the impact of clipping. The assessment of these methods is mostly given in terms of the (complementary) distribution of the CF which is regarded as a key parameter in OFDM system design. Now, given a certain CF reduction Theorem guarantees a certain reduction with respect to the SER even in coded systems because no further assumptions regarding the OFDM signal have been made. For a better bound, higher moments must be considered. This is, in principle, possible by introducing the RV from Theorem 17.

We can significantly tighten our results by assuming that the Nyquist-rate samples are independent and the phases are approximately uniformly distributed. Then

$$\overline{\mathrm{SER}} \leq \min_{\mu \geq \lambda} \left[F_L(\mu) + 4L \sum_{l=\left\lceil \frac{N_L(\lambda,\mu)}{L} \right\rceil}^{N} B_{p(\lambda)}(N,l)\, B_{\frac{1}{2}}\left(l, \left\lceil \frac{N_L(\lambda,\mu)}{L} \right\rceil \right) \right]$$

where $p(\lambda) := e^{-\lambda^2}$and $B_p(N,l) := \binom{N}{l} p^{N-l}(1-p)^l$ is the Binomial distribution. The bound may also work well for all constellations where at least the decorrelation of the Nyquist-rate samples (rather than independence) is fulfilled and the phases are approximately uniformly distributed. In the simulation in Fig. 4.4 we assume that the channel does not inflict any distortion on the OFDM symbol and also that the receiver operates at high SNR. Another line of thinking is that we compute the SER at the transmitter. Furthermore, the non-linearity acts on the Nyquist-rate samples. Note that the bound is still relatively weak but it is much better than the bound in [31]. Moreover the effects that describe the upper bound seem to explain the characteristic behavior, i.e. the converging behavior in the low probability region. In order to incorporate thermal noise the bound can either be described for fixed noise realization and averaging or the clipping noise can be replaced with an equivalent noise floor which is used in the standard SER formula. An annoying feature is that the bound becomes looser with larger N. The reason is that for large N the impact of small peaks is overestimated. We were not able to give a reasonable upper bound using for example Bernstein's inequality that could enable us to better assess the influence of small peaks. Clearly, the impact of oversampling is overestimated for it is more a stability result in the sense that the impact of oversampling does not adversely

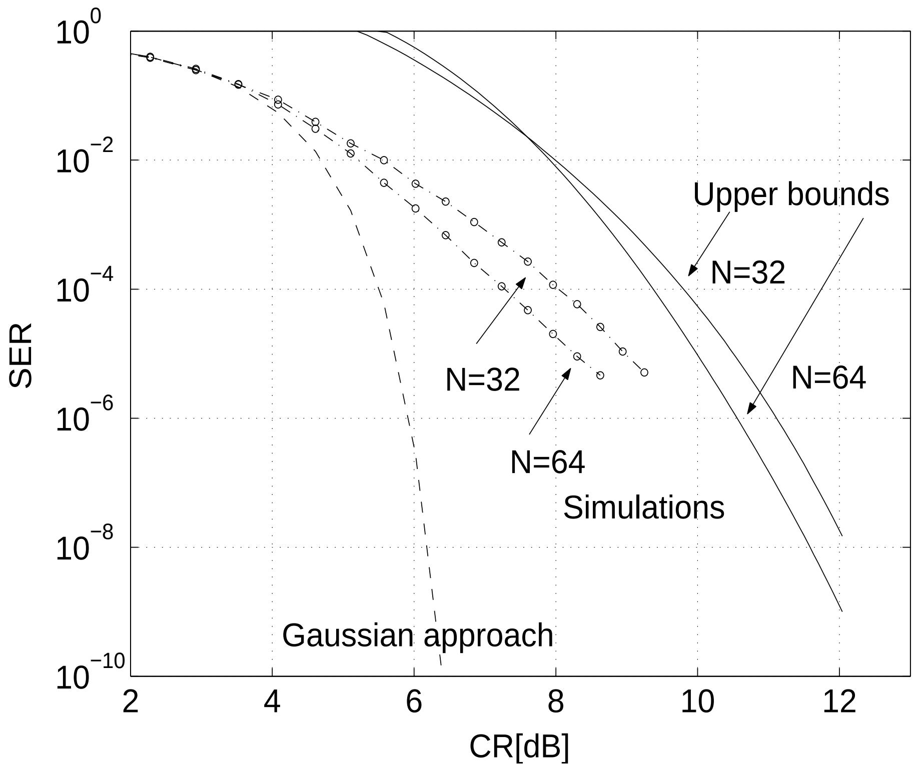

Figure 4.4: Symbol error rates and upper bounds for the 64-QAM constellation and different subcarrier numbers. A minimum of 20 events were observed for each simulation point. The maximum number of trials was $5 \cdot 10^6$.

affect the results. However, note that so far we have not introduced any heuristic arguments about the shape of the excursion above the clipping level as done, for example, in [31, 103, 104]. Following intuitive arguments we can now start shaping the pulse, for example by assuming that the pulse is well approximated by a parabolic arc. Approximative curves obtained from this approach appear in [37].

4.3.2 Asymptotic upper bounds

It is interesting to investigate the behavior of the average SER for $N \rightarrow \infty$. This can provide insights for adjusting system parameters in high order OFDM. In [105] the asymptotic SER was investigated assuming critical sampling, Gaussian distribution and a SL non-linearity. It was shown that if the CR grows as fast as $\sqrt{\log(N)}$ the probability of a peak above CR becomes zero and the SER is governed by noise and fading only. In Theorem 12 the result derived in [105] for the Gaussian case was proved for BPSK, MQAM and MPSK constellations and non-linearities that act on the time-continuous signals. As expected, it turned out that the same order for the clipping level holds as in the Gaussian case. From a practical viewpoint it may be interesting to ask whether the converse holds, i.e. if the CR grows slower than $\sqrt{\log(N)}$ the SER monotonously increases. The answer is negated by the following corollary.

Corollary 25 *Setting* $\lambda_N = \sqrt{\frac{(1+\varepsilon)\log(N)}{2}}$ *then for any* $\varepsilon > 0$

$$\lim_{N\to\infty} \left(\overline{\mathrm{SER}}\right) = 0.$$

Proof. This follows from setting $\lambda_N = \sqrt{\frac{(1+\varepsilon)\log(N)}{2}}$ and choosing $\mu_N = \sqrt{(1+\varepsilon)\log(N)}$ [36]. ■

We can deduce from this result that not only the peak values are of interest but also the clipped energy which decreases faster than the single peaks. However, it is much slower compared to the Gaussian approximation.

Note that since we have put the oversampling case down to the Nyquist-rate sampling case (up to a constant) it suffices to consider the Nyquist-rate sampling case. Assuming the input process to be Gaussian and exploiting the independence it follows that the distortion terms converge in distribution to a Gausssian random variable and thus the SER can be made arbitrarily small even if the CR is bounded for all N. It may now be interesting to make this statement mathematically strong for the standard constellations [106].

4.4 Open problems

We have derived bounds on the complementary distribution of the CF for both dependent and independent subcarriers. The bounds are sharp in the low probability region and large numbers of subcarriers. We have argued that this is a reasonable approach since in the high probability region a quick sketch can be obtained by simulations. The bounds can be improved by replacing the Chernoff technique by the approach in [107]. The asymptotic behavior was characterized and concluded that a certain CF cannot be avoided which has impact on constellation and HPA design. We have presented a SER bound but we were not able to appropriately consider the impact of oversampling and leave this as an open problem.

Research Problem 26 *Find improved bounds on the SER if oversampling is employed.*

We emphasize that the preceding coding results are fundamental since in principle all methods for tackling the CF problem can be identified as some form of coding. In the next chapter we examine the case where we use both error control coding and coding with low CF.

Chapter 5

Coded OFDM systems

In light of previous results it is clear that in one form or another the impact of clipping can be a serious concern and must, therefore, be tackled. Indeed, the search for methods in order to lower the CF in OFDM transmission is still an on-going research topic and although the CF reduction problem is a relatively recent problem, literature has been extremely prolific since the occurrence of the problem in the mid-nineties.

The CF reduction techniques can be divided into two main categories: reduction techniques that create additional distortion and distortionless reduction techniques. In the first class is for example clipping and filtering [108, 109] which has been considered in the last section. Distortionless techniques are redundancy-based and might generate side-information or not. Several methods such as partial transmit sequences [110, 27, 111, 112], selected mapping [113, 114, 26], iterative decoding [115], extended constellations [116], and trellis shaping [117] etc. have been proposed.

A more elegant method is to use error control coding in order to prevent peaks before transmission and enjoy the twin benefit of error correction and power control. In particular, coding methods are implementation-efficient for there is no need to detect peaks after modulation. The first approaches with respect to codes with low CF are due to Jones and Wilkenson [38, 39] where a block coding scheme with reduced CF was presented. This work was extended in [40] to combine error control coding and power control using linear offsets. There are also a couple of other interesting approaches for a small number of subcarriers [118, 119, 120, 121]. A major theoretical advance was made in recognizing the connection between Golay complementary pairs and first order cosets of second order Reed-Muller codes in [45]. The CF of these cosets have CF at most 3 [dB]. Special classes of these codes were found by [43] and [41]. The decoding these codes was considered in [122, 123, 124]. Further simple extensions of these codes can be obtained from [125] and [126]. In [9] new classes of codes with small Nyquist-rate CF such as trace codes were introduced.

A systematic approach to studying this coding problem was pursued in [9] where the trade-off between rate, minimum euclidean distance and CF of a code was investigated. A main result is that good codes exist which have CF of order $\sqrt{\log(N)}$. However, no codes that have the postulated properties could be designed (even when the additional constraint of a low CF is omitted, the design of asymptotically good codes is a challenging task). Another general outcome of this work is that not all the redundancy introduced by the coder can be exploited for CF reduction thereby giving up this paradigm of coding theorists.

The design problem is exacerbated by the fact that the minimum distance is usually not the criterion that counts and must be replaced by a different criterion [127] with respect to maximizing diversity. So far, no attempt has been made to investigate the fundamental trade-off between these generalized performance measures. A first step was taken in [82] where the trade-off of CF and Hamming distance was investigated. This combined problem was also considered in [128] where, however, the Nyquist-rate CF instead of the real CF was optimized.

We would like to point out that these design rules are usually not followed in existing systems where interleaving is used to increase the diversity order. Thus, although coding is used the symbols passed to the DFT are nearly independent implying the CF behavior derived in the last chapter. By contrast, we use a general approach and adopt the model in [9] where the constellation symbols are passed rawly to the DFT. It is worth noting that a careful analysis constellation must decide which model is appropriate, i.e. whether it is better to use interleaving and exploiting the additional time diversity or to benefit from lowering the impact of large CF (in terms of lower SER and lower adjacent channel interference). This may depend on a couple of practical side constraints for the applications and is beyond our scope.

In this chapter we examine the coding approach to lowering the crest-factor. First, we introduce a coded OFDM system and derive methods to determine the CF of the coded system and its error control properties. We provide a systematic approach to designing point sets in N-space with low crest-factor (corresponding to the number of subcarriers) that obey both criteria, i.e. they have low CF and large diversity order. We will call these sets generalized constellations. We do not need any numerical optimization and can analytically tightly bound the CF. We can also trade off the CF for spectral efficiency and complexity. The line of thinking is to design codes that yield good Nyquist-rate CF and then to modify the codewords systematically so that the real CF is given by the discrete CF and a small constant dependent on the scheme used. Modulation and demodulation of these constellations will be discussed to set up a new system concept. Finally, we discuss the consequences of the design method for the trade-off between rate, minimum distance and CF. It is shown how to design general constellations that perform well in this sense. Moreover, we can easily design constellations that have good distance properties and uniformly bounded CF. These codes will solve the problem stated in [9]. On the other hand, we will point out that these codes need not perform well in a coded OFDM system which simply shows that the considered minimum distance is not appropriate.

5.1 Design criteria for coded OFDM systems

Let $\mathcal{C}$ be a code that maps k input bits into blocks of n constellation symbols $c_0, \ldots, c_{n-1}$, forming the codeword c, from a constellation $\mathcal{Q}$ with M elements. The rate R of this code is defined to be $R = \frac{k}{n}$ so that $\mathcal{C}$ has $M_1 = 2^{Rn}$ codewords. Again we refer to the elements of c as coordinates of a (row) vector with square length $\|c\|^2$. The average power (also termed the average square length) is given by $P_{\text{av}} = E\,\|c\|^2$. We denote the minimum euclidean distance over all codewords as d_E. The minimum Hamming distance is d_H. If $N = n$ the OFDM symbol can be described by

$$s_c\,(t) = \sum_{k=0}^{N-1} c_k e^{jk\theta}, \quad 0 \leq \theta < 2\pi.$$

Let us introduce some relevant terms with respect to the CF problem. The CF of a code $\mathcal{C}$

is defined by

$$\mathrm{CF}(\mathcal{C}) = \max_{c \in \mathcal{C}} \mathrm{CF}(s_c).$$

The codeword which attains the maximum CF is called the worst-case codeword.

Clearly, it is generally difficult to provide practical design criteria so that the CF of the code is small. For spherical codes in [8] a new geometric criterion was introduced that is presented and improved in the next section. It is shown that, provided that the code supports minimum-distance decoding, the CF can be efficiently calculated. It is further shown how this can be exploited to design new codes with reduced CF.

We call the gain with respect to the CF of the coded scheme over the uncoded scheme the CF gain. In order to trade-off error-correcting properties over the CF in [129] elementary cost-functions were introduced.

Before we discuss methods of examining coded systems let us discuss what can be achieved in principle with coded OFDM systems in terms of capacity. Assuming that the path delays can be fixed to $\tau_l = \frac{lT_s}{N}$ with corresponding power-delay profile r_l which is assumed in the following the capacity of the system for a channel realization is given by

$$C_h = \frac{N - L_1}{N} \sum_{k=0}^{N-1} \log\left(1 + \frac{P_{\mathrm{av}} \left|\beta_k\right|^2}{\sigma^2 N}\right).$$

If a peak-power restriction is invoked the capacity is reduced. However, an analytical treatment of this problem is difficult and only solved for the AWGN channel [130]. Many authors simply model the distortion as additional Gaussian noise and then use the same capacity formula with altered noise power [131, 132].

5.1.1 Crest-factor

We focus now on the computation of the CF of the code $\mathcal{C}$. It was shown in [8] that this problem is intimately related to minimum-distance decoding and we now briefly discuss their approach. Let $\zeta = \frac{f_c}{\Delta f}$ be the ratio of subcarrier bandwidth and carrier frequency the authors established that (Theorem 3.2) whenever $\zeta \geq \zeta_1$

$$|\mathrm{PAPR}\left(\mathcal{C},\zeta\right) - G_L\left(\mathcal{C},\zeta_1\right)| \leq \left[\frac{2\pi N\left(N+\zeta_1\right)}{L} + \frac{2\pi N^2}{\sqrt{3}\zeta_1}\right] \tag{5.1}$$

where

$$\mathrm{PAPR}\left(\mathcal{C},\zeta\right) := \max_{c\in\mathcal{C}} \max_{0\leq\theta<2\pi} \frac{\left[\Re\left(s_c\left(\theta\right)e^{j\zeta\theta}\right)\right]^2}{P_{\mathrm{av}}}$$

is the PAPR of the (normalized) transmit signal and

$$G_L\left(\mathcal{C},\zeta\right) := \max_{0\leq l\leq LN} \max_{c\in\mathcal{C}} \frac{\left[\Re\left(s_c\left(\theta_{l,L}\right)e^{j\zeta\theta_{l,L}}\right)\right]^2}{P_{\mathrm{av}}} \tag{5.2}$$

the discrete-time version of $\mathrm{PAPR}\left(\mathcal{C},\zeta\right)$. Observing that the term

$$s_c\left(\theta\right)e^{j\zeta\theta} = \sum_{k=0}^{N-1} c_k e^{j(\zeta+k)\theta}$$

can be written as the inner product of c and the vector

$$\omega_1\left(\theta,\zeta\right) := \left(e^{j\zeta\theta}, e^{j(\zeta+1)\theta}, \ldots, e^{j(\zeta+N-1)\theta}\right),$$

we have

$$s_c\left(\theta\right)e^{j\zeta\theta} = \left(c, \omega_1\left(\theta,\zeta\right)\right)$$

where $(\cdot,\cdot)$ denotes the inner product of two vectors. Since the real part of the inner product of two vectors that lie on a sphere with radius $\sqrt{N}$ can be expressed as

$$\begin{aligned} 2\Re\left(c,\omega_1\right) &= \|c\|^2 + \|\omega_1\|^2 - \|c-\omega_1\|^2 \\ &= 2N - \|c-\omega_1\|^2, \end{aligned}$$

the term (5.2) can be efficiently computed using a minimum-distance decoding procedure and the error can be estimated by (5.1).

The main disadvantage of this approach is given by the fact that the term $G_L\left(\mathcal{C},\zeta_1\right)$ has to be evaluated on a polynomial in N increasing number of sampling points and that, in addition, the bound is extremely weak so that the PAPR is drastically overestimated. Furthermore, we want to emphasize that good error estimates are essential in this context as the extremal codeword depends on the parameters L and K. The real value and the error estimate are compared in [47].

Therefore, a better approach with lower computational effort is desirable and can be obtained by translating the approach into the baseband as described by the following theorem.

Theorem 27 *Let $\mathcal{C}$ be a code of N symbols over an equal-energy constellation and let $c = (c_0, \ldots, c_{N-1})$ denote its codewords. Let $\omega(\theta, \alpha)$ be defined as*

$$\omega(\theta, \alpha) := \left(e^{j\alpha}, e^{j(\theta+\alpha)}, \ldots, e^{j((N-1)\theta+\alpha)}\right)$$

and

$$\omega^*(\theta, \alpha) = \min_{c \in \mathcal{C}} \|c - \omega(\theta, \alpha)\|^2 .$$

Suppose that $L > 1, K > 2, L, K \in \mathbb{N}$ then the largest CF is enclosed by the inequalities

$$\begin{aligned} & N - \frac{1}{2} \min_{0 \le l_1 < LN} \min_{0 \le l_2 < K} \omega^*(\theta_{l,L}, \alpha_{l_2,K}) \\ & \le \sqrt{N}\mathrm{CF}(\mathcal{C}) \le \\ & C_L C_K \left(N - \frac{1}{2} \min_{0 \le l_1 < LN} \min_{0 \le l_2 < K} \omega^*(\theta_{l,L}, \alpha_{l_2,K})\right) \end{aligned}$$

Before we prove the theorem let us provide the following lemma.

Lemma 28 *Let $c \in \mathcal{C}$. Setting*

$$d_* := \min_{0 \le l_1 < LN} \min_{0 \le l_2 < K} \|c - \omega(\theta_{l_1,L}, \alpha_{l_2,K})\| .$$

then the inequality

$$\left(N - \frac{d_*^2}{2}\right) \le \max_{0 \le \theta < 2\pi} |s_c(\theta)| \le C_L C_K \left(N - \frac{d_*^2}{2}\right)$$

holds.

Proof. Writing

$$\begin{aligned} \max_{0 \le \theta < 2\pi} |s_c(\theta)| &= \max_{0 \le \theta < 2\pi} \max_{0 \le \alpha < 2\pi} \Re\left(s_c(\theta) e^{j\alpha}\right) \\ &= N - \frac{1}{2} \min_{0 \le \theta < 2\pi} \min_{0 \le \alpha < 2\pi} \|c - \omega(\theta, \alpha)\|^2 . \end{aligned}$$

Observing that $\Re\left(s_c\left(\theta\right)e^{j\alpha}\right)$ is a degree 1 trigonometric polynomial in α the oversampling factor can be defined as $\frac{K}{2}$. In this light taking a grid $\left(\theta_{l_1,L},\alpha_{l_2,K}\right), 0\le l_1 < LN, 0 \le l_2 < K$, of the plane $[0,2\pi]\times[0,2\pi]$ yields

$$\begin{aligned}\max_{0\le\theta<2\pi}\left|s_c\left(\theta\right)\right| &\ge \max_{0\le l_1<LN}\max_{0\le l_2<K}\Re\left(s_c\left(\theta_{l_1,L}\right)e^{j\alpha_{l_2,K}}\right)\\ &= N-\frac{d_*^2}{2}\end{aligned}$$

and also

$$\max_{0\le\theta<2\pi}\left|s_c\left(\theta\right)\right|\le C_L C_K\left(N-\frac{d_*^2}{2}\right).$$

where C_L and C_K are the constants dependent on the refinement of the grid. ■

We are now in a position to give a proof to our main theorem.

Proof of theorem. The CF of the code is given by

$$\begin{aligned}\mathrm{CF}\left(\mathcal{C}\right) &= \max_{c\in\mathcal{C}}\mathrm{CF}\left(s_c\right)\\ &= \frac{1}{\sqrt{N}}\max_{c\in\mathcal{C}}\max_{0\le\theta<2\pi}\max_{0\le\alpha<2\pi}\Re\left(s_c\left(\theta\right)e^{j\alpha}\right).\end{aligned}$$

Taking a grid $\left(\theta_{l_1,L},\alpha_{l_2,K}\right), 0\le l_1 < LN, 0 \le l_2 < K$, and changing the order of maximum operators yields for some pairs (l_1, l_2)

$$\begin{aligned}\max_{c\in\mathcal{C}}\Re\left(s_c\left(\theta_{l_1,L}\right)e^{j\alpha_{l_2,K}}\right) &= N-\frac{1}{2}\min_{c\in\mathcal{C}}\left\|c-\omega\left(\theta_{l_1,L},\alpha_{l_2,K}\right)\right\|^2\\ &= N-\frac{1}{2}\omega^*\left(\theta_{l_1,L},\alpha_{l_2,K}\right).\end{aligned}$$

Setting $d_*^2=\min_{0\le l_1<LN}\min_{0\le l_2<K}\omega^*\left(\theta_{l_1,L},\alpha_{l_2,K}\right)$ in Lemma 28 the theorem follows. ■

Theorem 27 states that the CF of a code can be calculated with any desired accuracy in linear running time in the length of the code opposed to polynomial running time as in the approach of [8]. Moreover, as pointed out in Chapter 3 this complexity cannot be significantly improved. The theorem has the geometrical interpretation that the codewords of any OFDM code should have a euclidean distance to the set of vectors $\omega\left(\theta_{l_1,L},\alpha_{l_2,K}\right), 0\le l_1 < LN, 0 \le l_2 < K$, as large as possible. However, it seems hardly possible to devise any explicit design scheme from this observation.

The approach can easily be applied to baseband systems where the transmit signal is given by the real part of the complex baseband signal. Note that, since the signal is real the α's to be considered in Theorem are reduced to $\alpha = 0$ and $\alpha = \pi$. In the passband approach the transmit signal itself is investigated. Defining the transmit signal as $s_c^{(p)}(\theta) = \Re\left(s_c(\theta)\, e^{2\pi j \zeta}\right)$ their approach turns out to be a special case of Theorem 27 if the bandwidth is appropriately replaced. However, the number of sampling points is extremely large if ζ is large and the error is greatly overestimated. Furthermore, the theorem was formulated for equal-energy codes but as pointed out in [8] that if the code has a Trellis diagram the algorithm can be extended by applying the Viterbi algorithm.

The major outcome of Theorem 27 is to provide a simple method of computing the CF of a code. Thus, for a given code we can apply the algorithm to the phase-shift problem of [40] in order to reduce the CF in OFDM. This was also proposed in [8].

The CF of each coordinate in all the codewords is multiplied by fixed phase shifts $\phi_k, k = 0, \ldots, N-1$. Denoting the modified code as $\mathcal{C}(\phi_0, \ldots, \phi_{N-1})$ (this is referred to as the phase-shifted version) we wish to find phase shifts $\left\{\phi_0^*, \ldots, \phi_{N-1}^*\right\}$ where

$$\mathrm{CF}\left(\mathcal{C}\left(\phi_0^*, \ldots, \phi_{N-1}^*\right)\right) \leq \mathrm{CF}\left(\mathcal{C}\left(\phi_0, \ldots, \phi_{N-1}\right)\right)$$

for all $\left\{\phi_0^*, \ldots, \phi_{N-1}^*\right\} \in [0, 2\pi)^N$. As a result we can consider $\mathcal{C}\left(\phi_0, \ldots, \phi_{N-1}\right)$ as a multivariable non-linear function and minimize this function using standard tools, for example a gradient method. However, as the function has also various local minima and saddle points there is not much hope of finding the global minimum.

An example is given by the HIPERLAN 2 standard. The coder is a standard 1/2 constraint length 7 convolutional code. We apply our algorithm to the BPSK case using Newman phase shifts [47]. The phase shifts yield a gain of 4.163 [dB]. The envelope is depicted in Fig. 5.1. It is worth mentioning that we assumed that the memory of the coder is cleared after submitting a codeword.

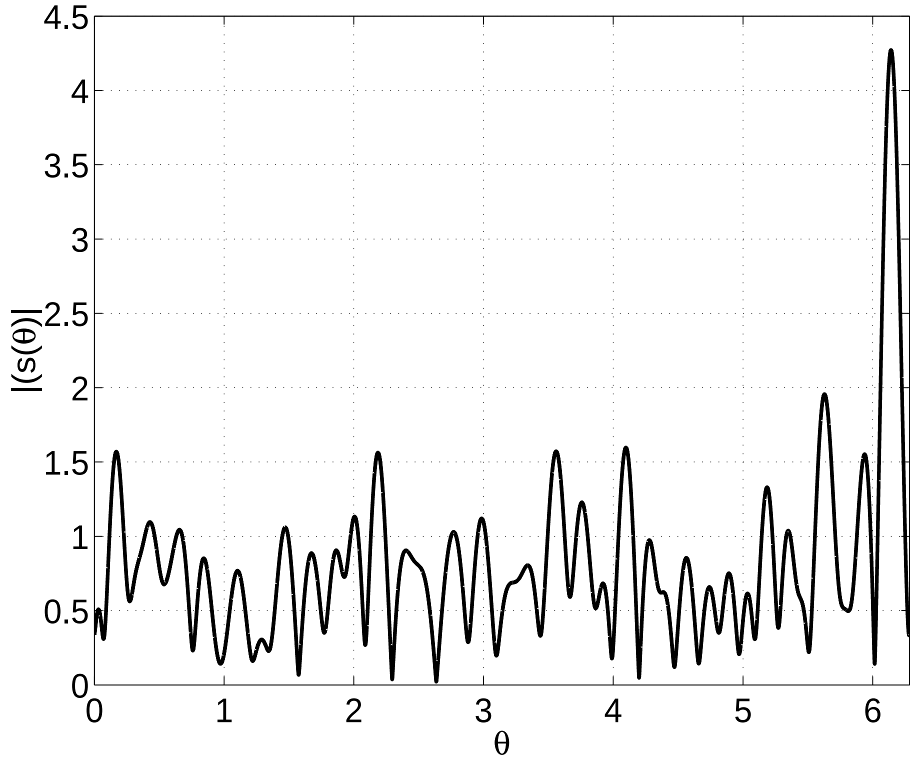

Figure 5.1: Corresponding signal (with unit average power) of worst-case codeword in a HIPERLAN 2 system using phase-shifted BPSK modulation with Newman phases $\phi_k = e^{\frac{j\pi k^2}{N}}$.

5.1.2 Diversity

Using diversity schemes is an important means of overcoming the limitations of wireless multipath channels. Examples are temporal, frequency and antenna diversity schemes. OFDM systems are designed for the broadband multipath channel where frequency diversity is available.

Recall that the received symbols can be written as

$$c'_k = \beta_k c_k + n_k, \quad k = 0, \ldots, N-1$$

where the sequences $n_k, \beta_k, k = 0, \ldots, N-1$, are the noise vector and channel gains, respectively. Assuming perfect channel state information at the receiver and a maximum-likelihood decoder, the receiver computes the decoded sequence according to

$$\widehat{c} = \arg\min_{c\in\mathcal{C}} \sum_{k=0}^{N-1} \left|c'_k - \beta_k c_k\right|^2 .$$

If c and e are two different codewords it was shown in [127] that the average pairwise error probability $\Pr(c \to e)$ over all channel realizations is upperbounded by

$$\Pr(c \to e) \le \prod_{i=0}^{\mathrm{rank}(C)-1} \left(1 + \lambda_i(C) \frac{P_{\mathrm{av}}}{2\sigma^2}\right)^{-1}$$

where

$$C := \sum_{l=0}^{L_m-1} r_l D_l (c-e)^H (c-e) D_l^H$$

with

$$D_l := \mathrm{diag}\left(1, e^{j\frac{l}{N}2\pi}, \ldots, e^{j\frac{(N-1)l}{N}2\pi}\right)$$

and $\lambda_i(C)$ denotes the i-th eigenvalue of C (c^H is the hermitian transpose and $\mathrm{rank}(C)$ is the number of independent rows/columns). The rank of C is the diversity order that we wish to be as large as possible. Clearly, the diversity order can only be as large as the Hamming distance of the underlying code. Since C is rather unwieldy to compute we point out that the matrix C can be rewritten as follows: the outer product $(c-e)^H(c-e)$ can be expressed as

$$(c-e)^H (c-e) = \mathrm{diag}((c-e))\, \mathbb{I}_N\, \mathrm{diag}((c-e)^*)$$

where $\mathbb{I}_N$ is an $N \times N$ matrix filled with ones and $*$ denotes elementwise conjugation. Thus, the matrix can be written as

$$C = \left[(c-e)^H (c-e)\right] \circ W_r W_r^H$$

where $\circ$ is the Hadamard product and W_r is a matrix consisting of entries $(W_r)_{kl} = r_l \exp\left(\frac{2\pi jkl}{N}\right)$ with row index $0 \le k < N$ and column index $0 \le l < L_m$. Furthermore, the diversity order of a frequency coded OFDM system is upperbounded by

$$\mathrm{rank}\left(\left[(c-e)^H (c-e)\right] \circ W_r W_r^H\right) \le \min\{L_m, d_H\}.$$

The upper bound has already been established in [127]. Clearly, it follows that in the case of uncoded systems there is no frequency diversity, i.e. $\operatorname{rank}(C) = 1$. If $L_m = 1$ (irrelevant for OFDM) the relevant coding measure is the minimum distance of the code. On the other hand if $L_m = N$ the relevant coding measure is the Hamming distance of all codeword pairs. In all other cases no simple coding strategy can be given. In particular, a code exhibiting excellent minimum distance that works well in AWGN can have very poor performance in the Rayleigh fading multipath channel. On the other hand, arbitrary fixed phase-shifts of all codewords have no effect on the error correcting capability of the code. The formula can be generalized if $(c - e)$ is itself a matrix $(C - E) \in \mathbb{C}^{N \times N_T}$ rather than a vector for space-frequency codes with N_T transmit antennas [127].

Obviously a necessary condition of achieving a diversity order of d' is a Hamming distance of d'. Moreover, it is also sufficient for collecting the maximum diversity order provided that $d' \geq L_m$ as the following argument shows. We have to investigate the rank of $C = \operatorname{diag}((c - e))\, W_r W_r^H \operatorname{diag}((c - e)^*)$. Since $\operatorname{rank}\left(AA^H\right) = \operatorname{rank}(A)$ it suffices to investigate the rank of $W_r^H \operatorname{diag}((c - e)^*)$. For any pair c, e and modulation scheme the number of zeros in the error vector is smaller than $N - d'$. The resulting matrix has non-zero columns only in the position where non-zero entries in the diagonal matrix occurred. Thus the columns are scaled versions of $e^{\frac{2\pi j k_1 l}{N}}, l = 0, \ldots, L_m - 1$, and $k_i \in [0, N-1], i = 1, \ldots, d'$. Now if the rank of this matrix is lower than $d' \geq L_m$ then the following set of equations has non-trivial solutions

$$\sum_{l=0}^{L_m - 1} c_l z_i^l = 0, \quad z_i = e^{\frac{2\pi j k_i}{N}}, c_l \in \mathbb{R}, \quad i = 1, \ldots, d',$$

which is not possible since the polynomial can only have $d' - 1$ zeros.

On the other hand, clearly the coding gain may become arbitrarily small although the diversity gain is guaranteed. This is because fixing L_m a Hamming distance of at least L_m guarantees two codewords to be distinguished after passing the channel. However, if the positions in which the codewords differ are close the small coding gain can lead to poor performance. On the other hand, in the absence of better (i.e. more tractable) criteria we will stay with the Hamming distance in this thesis.

5.2 Generalized constellations with low crest-factor

In the preceding sections we introduced the design criteria of coded OFDM systems and we will now propose new design schemes that perform well in the discussed sense. We call the outcome of these schemes generalized constellations rather than codes because these constellations do not possess many properties of codes. It is more related to the constellation approach in [84, 85] where we have added a new design criterion.

5.2.1 A new design rule

Generalized constellations are based on properties of de la Vallée-Poussin polynomials. Given n samples $s_c^{(1)}, s_c^{(2)}, \ldots, s_c^{(n)}$ and natural numbers $m, n_1, m > n_1$, we define the degree $m-1$ trigonometric polynomial

$$s_c(\theta) = \frac{1}{n}\sum_{l=0}^{n-1} s_c^{(l)} S_{m,n_1}\left(\theta - \frac{2\pi l}{n}\right)$$

where

$$S_{m,n_1}(\theta) = \frac{\sin\left(\frac{(m+n_1)\theta}{2}\right)\sin\left(\frac{(m-n_1)\theta}{2}\right)}{(m-n_1)\sin^2\left(\frac{\theta}{2}\right)}$$

is the de la Vallée-Poussin kernel (see Chapter 3). Suppose $m + n_1 = n$ then $s_c\left(\frac{2\pi l}{n}\right) = S_{n-n_1,n_1}(0) \cdot s_c^{(l)}$ which is assumed in the sequel. Thus we can write

$$s_c(\theta) = \frac{1}{n}\sum_{l=0}^{n-1} s_c\left(\frac{2\pi l}{n}\right) S_{n-n_1,n_1}\left(\theta - \frac{2\pi l}{n}\right),$$

and

$$\max_{0\le\theta<2\pi} |s_c(\theta)| \le \max_{0\le k<n}\left|s_c\left(\frac{2\pi k}{n}\right)\right| \max_{0\le\theta<\frac{2\pi}{n}} \frac{1}{n}\sum_{k=0}^{n-1}\left|S_{n-n_1,n_1}\left(\theta - \frac{2\pi k}{n}\right)\right|.$$

It was shown in Theorem 8 that

$$\begin{aligned}
&\max_{0\le\theta<\frac{2\pi}{n}} \frac{1}{n}\sum_{k=0}^{n-1}\left|S_{n-n_1,n_1}\left(\theta - \frac{2\pi k}{n}\right)\right| \\
&\le \min\left\{\frac{2}{\pi}\log\left(\frac{n}{n-2n_1}\right) + 2 + \frac{2}{\pi}\log(2) + \frac{2\cot\left(\frac{\pi}{2(n-2n_1)}\right)}{\pi(n-2n_1)}, \sqrt{\frac{n}{n-2n_1}}\right\} \\
&=: C_{n,n_1}^{(S)}.
\end{aligned}$$

Note that if the ratio of n and n_1 is constant the right hand side is uniformly bounded in n. The main idea is to use good codes that have small discrete CF and apply the interpolation. The interpolation can be carried out efficiently in the frequency domain. Obviously, the length of the codewords after the modification has increased, i.e. the spectral efficiency has decreased and clearly there is a trade-off between the offset length and the CF. Also the average power is modified and must be corrected. Note that this approach can be used to design multitone signals with low CF. It is natural to ask whether there are other kernels with even better performance. Firstly, it is certainly possible to use ,,smoother" kernels. However, from the results in Chapter 3 it is clear that these kernels do not necessarily exhibit better performance. One may also be attempted to use asymmetric kernels in order to reduce the offset length. However, we can exclude such kernels given by Theorem 9.

5.2.2 Modulation

The Fourier representation of S_{n-n_1,n_1} is given by

$$S_{n-n_1,n_1}(\theta) = 1 + 2\sum_{k=1}^{n_1} \cos(k\theta) + 2\sum_{k=n_1+1}^{n-n_1-1} \frac{n-n_1-k}{n-2n_1} \cos(k\theta)$$

and the k-th Fourier coefficient $c_k(s_c)$ is given by

$$c_k(s_c) = \frac{1}{n}\sum_{l=0}^{n-1} s_c\left(\frac{l}{n}2\pi\right) e^{\frac{2\pi jkl}{n}} \cdot \frac{1}{2\pi}\int_0^{2\pi} S_{n-n_1,n_1}(\theta)\, e^{jk\theta} d\theta.$$

Thus, all we need is the cyclically extended DFT of the original codeword and the Fourier coefficients of the kernel. This permits an efficient modulation.

5.2.3 Demodulation

The demodulation is more difficult. Since we do not send the original codeword the decoding procedure for the given code can be applied but its performance may be degraded. However, the degradation is small as we will see in an example. In the case of M-ary PSK codes we can do better. The original codeword is received where some of the coordinates are available at different positions and scaled by the kernel. A simple decoding procedure is now given as follows: copies of the same components can be combined and the decoding procedure of the code applied according to [125]. Let $c'_k, \beta_k, k = 0, \ldots, N-1$, be the received symbols and channel gains respectively the receiver aims to minimize the metric

$$\sum_{k=0}^{N-1} |c'_k - \beta_k c_k|^2$$

among all possible sequences. It is easy to see that the problem is equivalent to computing the metric

$$\sum_{k=0}^{N-1} |\beta_k c'_k - c_k|^2$$

if $|c_k| = 1$. After modification of the original codeword some of the coordinates are copies scaled by the kernel. Hence we can write

$$\sum_{k=d}^{N-1-d} |\beta^*_k c^*_k - c_k|^2$$

where β^*_k, c^*_k are the sums of the corresponding values of the copies. In the following we apply our scheme to trace codes.

5.2.4 Trace codes

Let us illustrate our approach with an example using m-sequences. M-sequences constitute a class of length $n = 2^m - 1$ codes encoding m bits where m is a natural number. These codes were proposed for OFDM in [133, 134]. The autocorrelation properties imply small Nyquist-rate CF and the minimum Hamming distance is fairly large. Thus the code is suitable for the transmission scheme proposed here. We chose $m = 4$, i.e. $n = 15$. Here, $n = 15$ corresponds to the original codeword length and $N = 2\,(n - n1) - 1$. The offset is defined by $d_s := \frac{(N-1)}{2} - \frac{(n-1)}{2}$ where n must be odd. If n is even the codeword lies asymmetric in the window defined by the kernel. Fig. 5.2 compares bit error rates (BER) over SNR $:= \frac{P_{\text{av}}}{\sigma^2}$ of uncoded BPSK, m-sequences and the modified scheme with offset $d_s = 1, d_s = 2$ and $d_s = 3$ in peak power limited Rayleigh fading multipath channel with $L_m = 3$ paths and equal power for all paths. In Fig. 5.3 the simulations are shown for $L_m = 4$. We assumed perfect channel state information and perfect

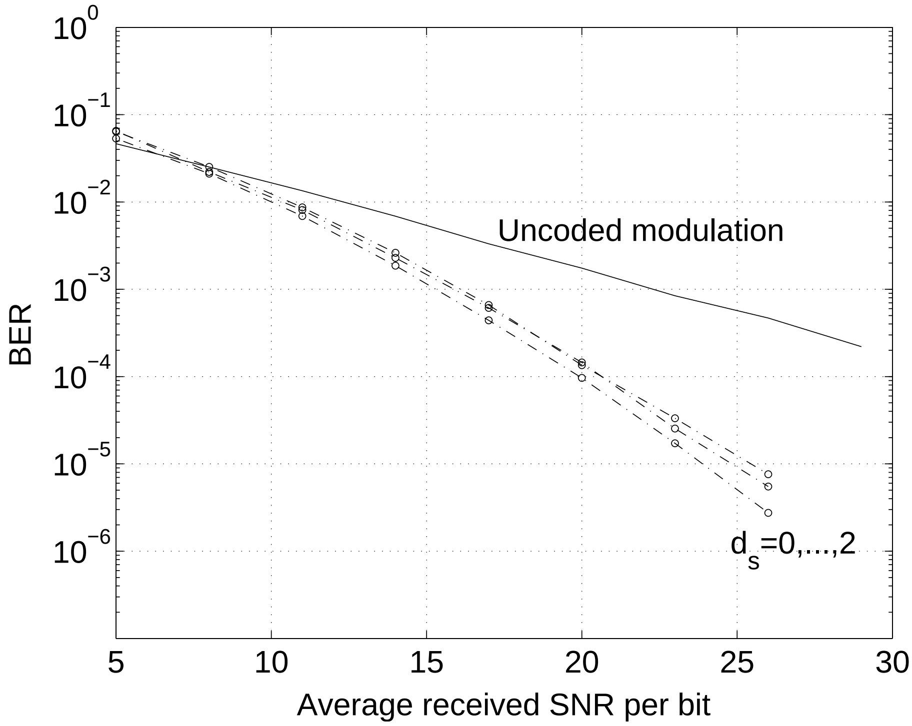

Figure 5.2: Simulation of bit error rates for the M-sequences and the new schemes. The number of paths is $L_m = 3$. At each simulation point a minimum of 20 events was observed. The maximum number is $5 \cdot 10^6$.

synchronization at the receiver. We observe a slightly loss in coding gain with increasing d_s, however the transmitter can now operate with a 1.4 [dB] for $d_s = 2$ (0.9 [dB] for $d_s = 1$) reduced OBO compared to 5.4 [dB] (the theoretical values for the OBO are 12.4 [dB], 10.8 [dB], and 9.9 [dB] times the CF of the coding scheme itself which is too pessimistic here). The reason for this is, that the modified scheme yields a strong gain with respect to the complementary distribution of the CF. It is further observed that the code can exploit the frequency diversity as expected.

The M-sequences belong to a class of codes called trace codes [9] for which good bounds on the Nyquist-rate CF do exist. It is not necessary to quote the relevant results from the theory of finite fields again and we refer to [135]. Combining for example the results in [9] and [136] we obtain an upper bound on the CF for trace codes, i.e.

$$\mathrm{CF}\,(\mathcal{C}) \leq C^{(S)}_{n,n_1} \cdot (2t-2)\,2^{\frac{m}{2}} + 1.$$

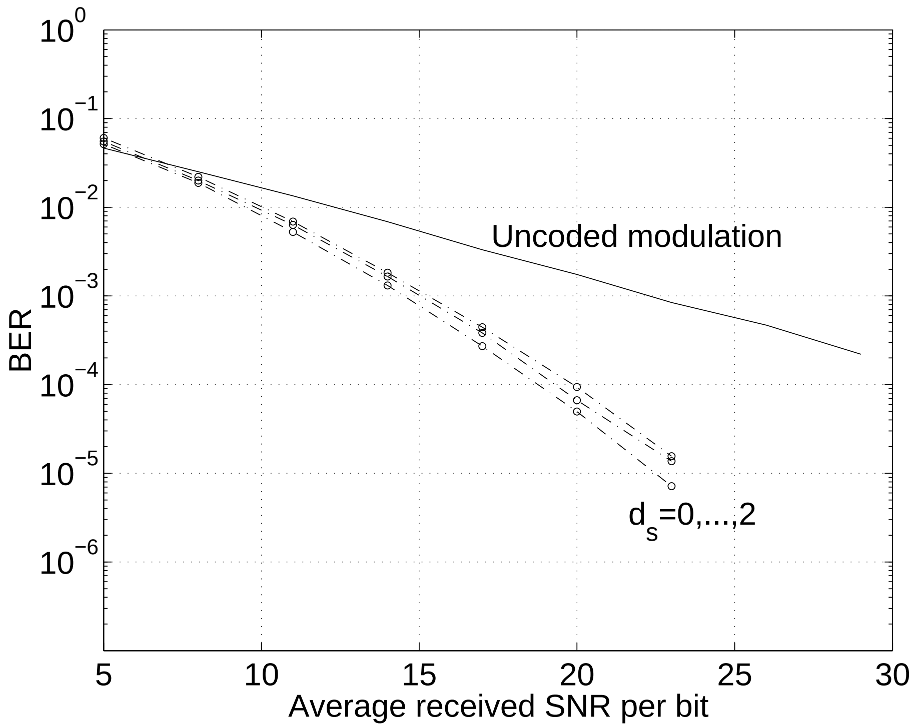

Figure 5.3: Simulation of bit error rates for the M-sequences and the new schemes. The number of paths is $L_m = 4$. At each simulation point a minimum of 20 events was observed. The maximum number is $5 \cdot 10^6$.

where t is the number of errors that can be corrected. Comparing these codes with the Reed-Muller codes it is apparent that they do not yield a rate gain. However, note that the decoding complexity is much less than that for Reed-Muller codes because no look-up tables are needed to identify those cosets with low CF. Thus, the approach can be an attractive alternative in low-cost systems as motivated. In principle, the scheme can be applied to every subcarrier number as long as a minimal polynomial is given. Further improvement of the scheme can be achieved if the all-zero codeword is prevented from transmission. In general, the scheme can be applied to every code with small Nyquist-rate CF as well as modulation with small Nyquist-rate CF. Modulation with small Nyquist-rate CF can be achieved with standard reduction techniques [7].

Unfortunately, the rates of the trace codes, Reed-Muller codes etc. are small and there may be a need to design constellations with higher rates (the techniques in [125, 126] can be used to increase rates). It is of interest what the fundamental limits of the aforementioned scheme are. This is discussed next.

5.3 Consequences for the crest-factor distance trade-off

Let us now discuss the preceding results with respect to the trade-off rate, distance and CF for spherical codes. Certainly, the most interesting question is regarding the existence of codes with good error properties and low CF. Some numerical investigations have been conducted in [137]. We recall some of the results presented in [9] and improve slightly on them.

Define

$$\Omega := \{\omega(\theta,\alpha) : 0 \le \theta < 2\pi, 0 \le \alpha < 2\pi\}$$

and let $A(\Omega, r)$ denote the area covered by the points within distance r on the N-dimensional sphere with radius $\sqrt{N}$, $A(r)$ denote the area of the spherical cap of radius r and $A_{\sqrt{N}}$ denotes the whole area of the surface of the sphere. Fixing the minimum euclidian distance d_E it was shown in [9] that providing

$$A\left(\Omega, \max\left\{d_* - \frac{d_E}{2}, 0\right\}\right) + 2^{NR} A\left(\frac{d_E}{2}\right) \ge A_{\sqrt{N}}, \quad \frac{d_E}{2} \le d_* \le \sqrt{2N}$$

for some d_* yields

$$\mathrm{CF}(\mathcal{C}) \ge \frac{1}{\sqrt{N}}\left(N - \frac{d_*^2}{2}\right).$$

In order to make the theorem practically useful a good lower bound on $A\left(\Omega, \max\left\{d_* - \frac{d_E}{2}, 0\right\}\right)$ is needed. The next lemma improves on Lemma 4 in [9].

Lemma 29 *The area of $A(\Omega, r)$ is lowerbounded by*

$$A(\Omega, r) \ge \begin{cases} (2N-1)A(r) & 0 \le r \le \sqrt{\frac{N}{2} - \frac{1}{4}} \\ N_0 A(r) & \sqrt{\frac{N}{2} - \frac{1}{4}} < r \le \sqrt{N} \\ A(r) & r > \sqrt{N} \end{cases}.$$

where $N_0 = \left\lfloor \frac{\pi}{\arcsin r\sqrt{N}} \right\rfloor$ where $\lfloor\cdot\rfloor$ is the floor-function.

Proof. The last inequality on the right-hand side is trivial since Ω is a non-empty set. The first inequality can be proven as follows. Taking two points $\omega(\theta_1,\alpha)$ and $\omega(\theta_2,\alpha)$ of Ω the squared euclidean distance is given by

$$\begin{aligned}\left\|\omega(\theta_1,\alpha)-\omega(\theta_2,\alpha)\right\|^2 &= 2N-\Re\left(\omega(\theta_1,\alpha),\omega(\theta_2,\alpha)\right)\\ &= 2N-1-1-2\sum_{k=1}^{N-1}\cos\left[k\left(\theta_1-\theta_2\right)\right]\\ &= 2N-1-\frac{\sin\left[\frac{2N-1}{2}\left(\theta_1-\theta_2\right)\right]}{\sin\left[\frac{\theta_1-\theta_2}{2}\right]}\end{aligned}$$

Thus, with $\theta_1-\theta_2=\frac{2\pi k}{2N-1}, 0<k<2N$, yields

$$\left\|\omega(\theta_1,\alpha)-\omega(\theta_2,\alpha)\right\|^2=2N-1.$$

Consequently, assuming $d\le\frac{\sqrt{2N-1}}{2}=\sqrt{\frac{N}{2}-\frac{1}{4}}$ we can choose points $\theta_k, 0\le k<2N$, so that $\left\|\omega(\theta_{k_1},\alpha)-\omega(\theta_{k_2},\alpha)\right\|^2=2N-1, 0\le k_1\ne k_2<2N$ and the spherical caps do not overlap. The second inequality follows from Lemma 4 in [9]. ■

Lemma 29 places a lower bound on the CF of a code provided that $\frac{d_E}{2}\le d_*\le\sqrt{\frac{N}{2}-\frac{1}{4}}$. Clearly if d'_E is the distance such that $2^{NR}A\left(\frac{d'_E}{2}\right)\ge A_{\sqrt{N}}$ then $\mathrm{CF}(\mathcal{C})\ge\frac{1}{\sqrt{N}}\left(N-\frac{d'^2_E}{2}\right)$ with equality at best. Allowing a slightly lower minimum distance leaves space for a lower CF as shown in Fig. 5.4 where the region below the curve is a non-achievable region. The region above the curve is the achievable region for a code. However, note that even if Lemma would hold with the best inequality over the whole possible range $0\le d_E\le\sqrt{2N}$ the range of possible distances differ only in extremely small values. This is due to the linear and exponentially growing number of points. Thus, the statement ,,there is a trade-off between rate, minimum distance and CF" [9] must be considered very carefully at least in light of these results.

The sufficient conditions derived in [9] can also be slightly improved. Using the Varshamov-Gilbert style argument in [9] it can be shown that if

$$A(\Omega,d_*)+2^{NR}A(d_E)\le A_{\sqrt{N}}$$

then there is a code $\mathcal{C}$ with rate R, minimum distance d_E and

$$\mathrm{CF}(\mathcal{C})\le C_LC_K\left(N-\frac{d_*^2}{2}\right).$$

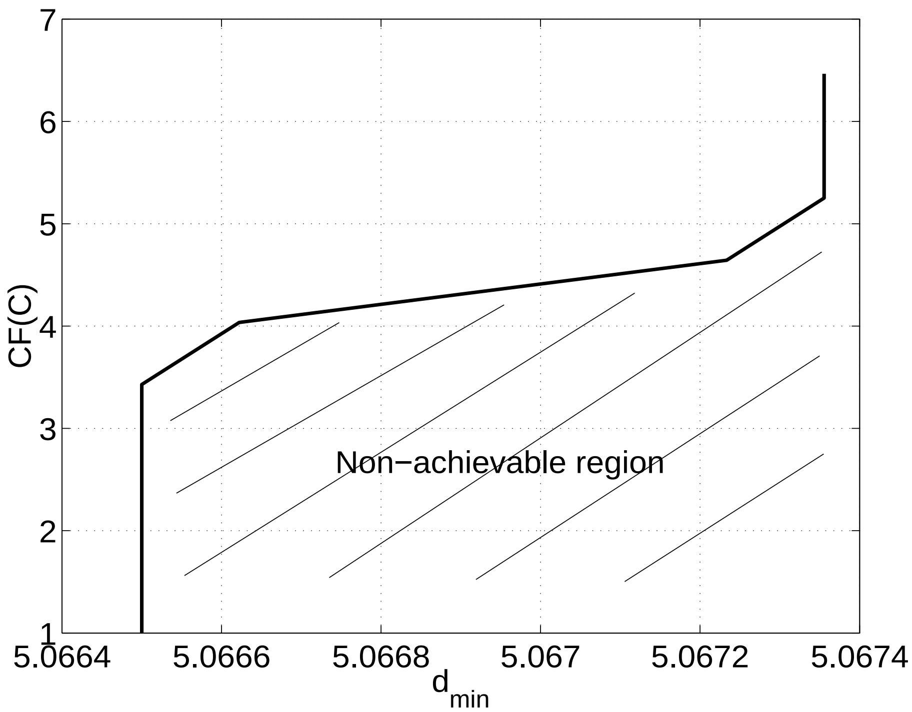

Figure 5.4: Trade-off between CF and distance for rate $R = 0.5$,$N = 48$ codes

A similar statement can be made for the Hamming distance based on a counting argument using (also our improved) bounds on the CF distribution and standard results from literature [82].

Exploring the asymptotic behavior by using the inequality of Shannon it is clear from [9] that by assuming

$$2^R \left(2\Delta \left(1 - \frac{\Delta}{2} \right) \right) < 1 \tag{5.3}$$

with $\Delta > 0$ then there is a code with $d_E = \sqrt{2\Delta N}$ and

$$\mathrm{CF}\left(\mathcal{C}\right) \leq (1 + \varepsilon) \sqrt{\log\left(N\right)}$$

for any $\varepsilon > 0$ and N large enough in a non-probabilistic sense. Note that the condition on Δ has only been introduced in order to fulfil the traditional Varshamov-Gilbert style bound on the existence of codes with distance $d_E = \sqrt{2\Delta N}$. Thus, in this asymptotic regime there is no interaction between error control coding and coding with low CF as long as $\mathrm{CF}\left(\mathcal{C}\right) \geq \sqrt{\log\left(N\right)}$ sustaining the results from the probabilistic analysis. It is also interesting to ask for the converse in the asymptotic case, i.e. whether

$$\mathrm{CF}\left(\mathcal{C}\right) \geq \left(1-\varepsilon\right)\sqrt{\log\left(N\right)}$$

holds if condition (5.3) is satisfied. Indeed, this would be the case if the inequality $A\left(\Omega, r\right) \geq \left(2N-1\right) A\left(r\right)$ would hold for arbitrary r asymptotically and not only for the particular range. In fact $\mathrm{CF}\left(\mathcal{C}\right) \geq \left(1-\varepsilon\right)\sqrt{\log\left(N\right)}$ implies $d^* \to 2N$ for $N \to \infty$ and hence falls out of the range. On the other hand we can determine a Δ^* such that $d^* - \frac{d_E}{2} \leq \sqrt{\frac{N}{2}-\frac{1}{4}}$ holds. A simple calculation yields $\Delta^* \geq 1$. Consequently, if $d_E \geq \sqrt{2N}$ then $\mathrm{CF}\left(\mathcal{C}\right) \geq \left(1-\varepsilon\right)\sqrt{\log\left(N\right)}$ for any $\varepsilon > 0$ which does not seem to have practical implications. We intend to explain this cut-off phenomenon, in part.

As we pointed out it appears to be a challenging task to construct codes with good properties. Let us discuss our scheme in this regard. Obviously we need error-correcting codes that have small discrete CF. These codes do not seem to be known in great abundance. However, if we take, for example, good arbitrary PSK codes with some minimum distance we can multiply every codeword with a unitary matrix and obtain a new code that has the same minimum distance. Both codes are so-called equivalent. In particular, we can pick the DFT as the unitary matrix itself. Hence, we have constructed a code that has small discrete CF and the same minimum distance. Moreover, applying our scheme we again obtain a new code that has even small (real) CF with somewhat reduced minimum distance and power (but the constants do not depend on N). The design scheme is as if the system performed the DFT at the receiver and can therefore dispose of the fluctuations of the signal envelope. Indeed, both systems do behave equally in the AWGN channel but do not in a multipath Rayleigh fading channel and hence the scheme by no means implies good performance in OFDM transmission. For example, taking any constellation it can be shown that the Hamming distance between two different codewords is zero [128]. A further problem is that we cannot supply this scheme with a practical ML decoding algorithm. Combining these ideas with those in [128] and partioning the subcarriers into disjoint subsets and applying fixed phase shifts on these subsets then, if the cardinality of the subsets is not too large, ML decoding can be applied. However, only a certain reduction of the CF can be achieved since a small CF of small subsets of subcarriers does not guarantee small CF of all subcarriers. The approach can be extended to the coded case.

5.4 Open problems

In this chapter we have provided a new class of codes that have both small CF and high diversity. These codes are built on codes that already have small CF on the Nyquist-rate sampling set and a smart interpolation rule. Although their rates are small they may be interesting for certain applications. In order to obtain codes with low CF we used codes that already show exhibit good CF on the Nyquist-rate sampling set. An example for this design was given by using M-sequences that which were an early candidate for OFDM codes until one found that their behavior on the Nyquist-rate sampling set does not imply good behavior for the CF in general. Thus in accordance with [9] we renew the call for codes that have the desired properties.

In a different analysis we investigated the trade-off involved between rate, distance and CF. We improved on the approach in [9] and showed that good codes with asymptotically CF of order $\sqrt{\log(N)}$ exist. We failed to prove the necessity of this result and strongly believe that this is not possible, given the result in Lemma 29 where we showed that the maximum distance between two points on Ω is $2N-1$ and our design rule for constellations with low CF based on properties of certain polynomials that we used to theoretically construct codes with uniformly bounded CF and good distance properties. In addition, we believe that in order to construct asymptotically good codes one needs to apply the scheme. Furthermore, the discussion in the last section leads us to the conclusion that, with respect to the euclidean distance, it is more of a problem to decode rather than to design codes with low CF. On the other hand, we simply shift the dynamics of the signal into the discrete-time domain.

Chapter 6

Conclusions

In this thesis we have presented a general framework for the power control problem in orthogonal frequency division multiplexing (OFDM) transmission. For a number of subproblems relevant in this field we have found new approaches and solutions. We have investigated the relationships between the continuous-time and discrete-time OFDM signal which is interesting for the system designer. Furthermore, we have examined statistical properties of the CF and the effects on performance leading to some interesting and fundamental design limits. Finally, we have presented a new coding scheme based on the previous results and analyzed its performance. It is now clearly possible to design OFDM systems more efficiently but for some problems further work has to be done. We pointed out possible research strands. Some other important aspects of OFDM systems were not considered, for example channel estimation, synchronization and multiple antennas. However, in particular the results for trigonometric polynomials and band-limited signals permit a couple of possible other applications.

Bibliography

[1] P. Ebert S. Weinstein, "Data transmission by frequency-division multiplexing using the discrete fourier transform," *IEEE Transactions on Communication Technology*, vol. Vol. COm-19, pp. No.5:pp.628–634, October 1971.

[2] H. Bölcskei, D. Gesbert, A.J. Paulraj, "On the capacity of OFDM-based spatial multiplexing systems," *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 225–234, February 2002.

[3] M. Riesz, "Eine trigonometrische Interpolationsformel und einige Ungleichungen für Polynome," *Jahresbericht DMV*, vol. 23, pp. 354–368, 1914.

[4] S. M Bernstein, "Sur le maximum absolu d'une somme trigonometrique," *Comptes Rendu de l'Academie des Sciences (Paris)*, , no. 193, pp. 433–436, September 1931.

[5] H. Ehlich and K. Zeller, "Schwankung von Polynomen zwischen den Gitterpunkten," *Math. Zeitschrift*, vol. 86, pp. 41–44, 1964.

[6] K. Jetter, G. Pfander and G. Zimmermann, "The crest-factor for trigonometric polynominals part I: Approximation theoretical estimates," *Rev. Anal. Numr. Thor. Approx.*, vol. 30, pp. 41–56, 2001.

[7] J. Tellado, *Peak to average power reduction for multicarrier modulation*, Ph.D. thesis, Stanford university, 1999.

[8] V. Tarokh and H. Jafarkhani, "On the computation and reduction of the peak-to average power ratio in multicarrier communications," *IEEE Trans. on Comm.*, vol. 48, no. 1, pp. 37–44, January 2000.

[9] K. G. Paterson and V. Tarokh, "On the existence and construction of good codes with low peak-to-average power ratios," *IEEE Trans. on Inf. Theory*, vol. 46, no. 6, pp. 1974–1986, September 2000.

[10] C Tellambura, "Upper bound on peak-factor of n-multiple carriers," *Electron. Letters*, vol. 33, no. 19, pp. 1608–1609, September 1997.

[11] H. Minn, C.Tellambura, and V. K. Bhargava, "On the peak factors of sampled and continuous signals," *IEEE Comm. Letters*, vol. 5, no. 4, pp. 429–431, April 2001.

[12] D. Wulich, "Comments on the peak-factor of sampled and continuous signals," *IEEE Commun. Lett.*, vol. 4, no. 5, pp. 213–214, July 2000.

[13] H. Boche and J. Fischer, "Größenabschätzungen von überabgetasteten Signalen," *FREQUENZ*, vol. 51, pp. 60–66, 1997.

[14] H. Boche and M. Protzmann, "Oversampling and limitation of signals," *IEEE Trans. CAS-I: Fundamental Theory and Applications*, vol. 48, no. 3, pp. 364–365, June 2001.

[15] G. Wunder and H. Boche, "Trigonometric sampling series and their application in mobile communications," in *Kleinheubacher Berichte*, vol. 44, pp. 183–190. T-Nova Technologiezentrum, 2001.

[16] J. R. Partington, *Interpolation, Identification, and Sampling*, Clarendon Press, Oxford, 1997, London Mathematical Society Monographs, New Series.

[17] G. Benke and B. B. Wells, "Estimates for the stability of low-pass filters," *IEEE Trans. on ASSP*, vol. 33, no. 1, pp. 98–105, February 1985.

[18] G. Wunder and H. Boche, "Peak magnitude of oversampled trigonometric polynominals," *Frequenz*, vol. 56, no. 5-6, pp. 102–109, May/June 2002.

[19] G. Wunder and H. Boche, "Peak value estimation of band-limited signals from its samples with application to the peak-to-average power problem in OFDM," in *Proc. IEEE Int. Symp. on Inf.-theory (ISIT)*, Lausanne, July 2002, p. 17.

[20] H. Boche and G. Wunder, "Über eine Verallgemeinerung eines Resultats von Riesz über trigonometrische Polynome auf allgemeine bandbegrenzte Funktionen," *Zeitschrift für angewandte Mathematik und Mechanik*, vol. 82, no. 5, pp. 347–351, May 2002.

[21] G. Wunder and H. Boche, "Peak value estimation of band-limited signals from its samples, noise enhancement and and a local characterisation in the neighborhood of an extremum," *IEEE Trans. on Signal Processing*, vol. 51, no. 3, pp. 771–780, March 2003.

[22] H. Boche and G. Wunder, "Abschätzung der Normen von Operatoren analytischer Funktionen," *Complex Variables*, vol. 48, no. 3, pp. 211–219, 2003.

[23] R. Salem and A. Zygmund, "Some properties of trigonometric series whose terms have random signs," *Acta Mathematica*, , no. 91, pp. 245–301, 1954.

[24] G.Halasz, "On a result of Salem and Zygmund concerning random polynominals," *Studia Scien. Math. Hung.*, , no. 8, pp. 369–377, 1973.

[25] A. Gersho, B. Gopinpath and A. Odlyzko, "Coefficient inaccuracy in transversal filtering," *The Bell System Technical Journal*, vol. 58, no. 10, pp. 2301–2317, December 1979.

[26] D. Mestdagh and P. Spruyt, "A method to reduce the probability of clipping in DMT-based transceivers," *IEEE Transactions on Communications*, vol. 44, no. 10, pp. 1234–1238, October 1996.

[27] S. Müller, R. Bäuml, R. Fischer and J. Huber, "OFDM with reduced peak-to-average power ratio by multiple signal representation," *Annals of Telecommunication*, vol. 52, no. 1-2, pp. 58–67, February 1997.

[28] M. Friese, "On the achievable information rate with peak-power limited OFDM," *IEEE Trans. on Inf. Theory.*, vol. 46, no. 7, pp. 2579–2587, November 2000.

[29] H. Ochiai and H. Imai, "On the distribution of the peak-to-average power ratio in OFDM signals," *IEEE Trans. on Comm.*, vol. 49, no. 2, pp. 282–289, February 2001.

[30] N. Dinur and D. Wulich, "Peak-to average power ratio in high-order OFDM," *IEEE Trans. on Comm.*, vol. 49, no. 6, pp. 1063–1072, June 2001.

[31] A. Bahai, M. Singh, A. Goldsmith, and B. Saltzberg, "A new approach for evaluating clipping distortion in multicarrier systems," *IEEE Sel. Areas on Comm*, vol. 20, no. 5, pp. 1037–1046, June 2002.

[32] H. Boche and G. Wunder, "On the PAPR problem in OFDM systems," in *39th Annual Allerton Conference on Communication, Control, and Computing*, October 2001.

[33] G. Wunder and H. Boche, "New results on the statistical distribution of the crest-factor of OFDM signals," *IEEE Trans. on Inf. Theory*, vol. 49, no. 2, pp. 488–494, February 2003.

[34] G. Wunder and S. Litsyn, "On the statistical distribution of the crest-factor of codes in OFDM transmission," in *Proc. IEEE Inform. Theory Workshop*, Paris, March/April 2003, pp. 191–194.

[35] G. Wunder and H.Boche, "Estimating the impact of the HPA non-linearity on the symbol error rate in OFDM systems," in *ITCom 2002*, Boston, July/August 2002, SPIE, pp. 13–22.

[36] G. Wunder and H.Boche, "The impact of nonlinear devices on the symbol error rate in broadband OFDM transmission," in *VTC Fall 2002*, Vancouver, September 2002.

[37] G. Wunder and H. Boche, "Evaluating the SER in OFDM transmission with nonlinear distortion: An analytic approach," in *7th Int. OFDM Workshop*, Hamburg, September 2002, pp. 193–197.

[38] A. E. Jones, T. A. Wilkinson and S. K. Barton, “Block coding scheme for reduction of peak-to-mean envelope power ratio of multicarrier transmission schemes,” *Electron. Letters*, vol. 30, pp. 2098–2099, 1994.

[39] T. A. Wilkinson and A. E. Jones, “Minimisation of the peak-to-mean envelope power ratio of multicarrier transmission schmes by block coding,” in *IEEE* 45^{th} *Vehicular Technology Conf.*, Chicago, IL, July 1995, pp. 825–829.

[40] A. E. Jones and T. A. Wilkinson, “Combined coding error control and increased robustness to system robustness in OFDM,” in *IEEE VTC Conf.*, Atlanta, April 1996, pp. 904–908.

[41] H. Ochiai and H. Imai, “Forward error correction with reduction of PAPR of QPSK multicarrier signals,” in *IEEE Int. Symposium Inf. Theory*, Ulm, Germany, June/July 1997, p. 120.

[42] H. Ochiai and H. Imai, “Block coding scheme based on complementary sequences for multicarrier signals,” *IEICE Trans. Fundamentals*, pp. 2137–2143, 1997.

[43] R. van Nee, “OFDM codes for peak-to-average power reduction and error correction,” in *Proc. IEEE Globecom*, London, November 1996, pp. 740–744.

[44] J. Jedwab J. Davis, “Peak-to-mean power control and error correction for OFDM transmission using Golay sequences and Reed-Muller codes,” *Electron. Letters*, vol. 33, no. 4, pp. 267–268, February 1997.

[45] J. A. Davis and J. Jedwab, “Peak-to-mean power control in OFDM, Golay complementary sequences, and Reed-Muller codes,” *IEEE Trans. on Inf.-Theory*, vol. 45, no. 7, pp. 2397–2417, November 1999.

[46] K.G. Paterson, “Generelized Reed-Muller codes and power control in OFDM modulation,” *IEEE Trans. Inform. Theory*, vol. 46, no. 1, pp. 104–120, January 2000.

[47] G. Wunder and H. Boche, “A baseband model for computing the PAPR in OFDM systems,” in *4th International ITG Conference on Source and Channel Coding*, Berlin, January 2002, VDE, pp. 273–280, VDE-Verlag GmbH.

[48] G. Wunder and H. Boche, “Generalized constellations with low crest-factor for OFDM transmission,” in *Proc. Int. Symp. Inform. Theory*, Yokohama, Japan, June/July 2003.

[49] K. Gröchnig, *Foundations of time-frequency analysis*, Birkhäuser, Boston, 2001.

[50] W. Kozek and A. Molisch, “Nonorthogonal pulseshapes for multicarrier communications in doubly dispersive channels,” *IEEE Journal Sel. Areas in Commun.*, vol. 16, no. 8, pp. 1579–1589, October 1998.

[51] H.G. Feichtinger and T. Strohmer, Eds., *Gabor Analysis and Algorithms*, Birkhäuser, Boston, 1998.

[52] W. Kozek, Götz Pfander, and Georg Zimmermann, “Perturbation stability of coherent Riesz systems under convolutions and additive noise,” *Appl. Comput. Harmon. Anal.*, vol. 12, pp. 286–308, 2002.

[53] Boelcskei and F. Hlawatsch, *Gabor Analysis and Algorithms, Theory and Applications*, chapter Oversampled modulated filter banks, pp. 295–322, Birkhuser, Boston, 1998.

[54] A.G. Burr, “Performance analysis of COFDM for broadband transmission on fading multipath channel,” *Wireless personal communications*, vol. 10, no. 1, pp. 3–17, June 1999.

[55] H. Ochiai, *Analysis and reduction of peak-to-average power ratio in OFDM system*, Ph.D. thesis, University of Tokio, 2001.

[56] P. Banelli and G. Baruffa, “Mixed BB-IF predistortion of OFDM signals in non-linear channels,” *IEEE Trans. on Broadcast.*, vol. 47, no. 2, pp. 137–145, June 2001.

[57] H. Ochiai, “Power efficiency comparison of OFDM and single-carrier signals,” in *IEEE VTC Fall 2002*, Vancouver, Canada, September 2002.

[58] J. G. Proakis, *Digital Communications*, McGraw-Hill, Inc., 2ed edition, 1983.

[59] J. Jong, K. Yang, W. Stark, G. Haddad, "Power optimization of OFDM systems with dc bias controlled nonlinear amplifiers," in *Proc. IEEE VTC'99*, April 1999, vol. 3, pp. pp.268–272.

[60] E. Costa and S. Pupolin, "M-QAM-OFDM system performance in the presence of a nonlinear amplifier and phase noise," *IEEE Trans. on Comm.*, vol. 50, no. 3, pp. 462–472, March 2002.

[61] O. Edfors et. al., "OFDM channel estimation by singular value decomposition," *IEEE Trans. Comm.*, vol. 46, no. 7, pp. 931–939, July 1998.

[62] K. Seip, "Developments from nonharmonic Fourier series," *Documenta Mathematica, ICM 1998*, vol. II, pp. 713–722, 1998.

[63] J. R. Higgins, *Sampling Theory in Fourier and Signal Analysis*, Clarendon Press, Oxford Science Publications, 1996.

[64] A. Schönhage, *Approximationstheorie*, De Gruyter Lehrbuch, 1971.

[65] P. L. Butzer, W. Splettstößer and R. Stens, "The sampling theorem and linear prediction in signal analysis," *Jber. Deutsch. Math. Vereinigung*, vol. 90, pp. 1–70, 1988.

[66] J. Schiff, *Normal Families*, Springer-Verlag, 1993.

[67] N.I. Achieser, *Vorlesungungen über Approximationstheorie*, Akademieverlag Berlin, 1953.

[68] A.N. Kolmogorov and S.V Fomin, *Introductory real analysis*, Dover Publications, Inc., New York, 1975.

[69] A. Zygmund, *Trigonometric Series*, vol. 1, Cambridge University Press, 1959.

[70] J.B. Garnet, *Bounded analytic functions*, vol. 9, Academic press, pure and applied mathematics, New York, 1981.

[71] H. Boche and H. Schreiber, "Rekonstruktion von Abtastreihen mit Kosinus-roll-off Kernen," *Kleinheubacher Berichte (in German)*, pp. 665–674, 1997.

[72] S. Boyd, "Multitone signals with low crest-factor," *IEEE Transactions on CAS*, vol. 33, no. 10, pp. 1018–1022, October 1986.

[73] R. M. Schroeder, "Synthesis of low-peak-factor signals and binary sequences with low autocorrelation," *IEEE Trans. Inform. Theory*, vol. IT-13, pp. 85–89, 1970.

[74] B. Popovic, "Synthesis of power efficient multitone signals with flat amplitude spectrum," *IEEE Trans. on Comm.*, vol. Vol.39, no. 7, pp. pp. 1031–1033, July 1991.

[75] S. Narahashi and T. Nojima, "New phasing scheme of n-multiple carriers for reducing peak- to-average power ratio," *Electr. Lett.*, vol. 30, pp. 1382–1383, 1994.

[76] J.P. Kahane, "Sur les polynomes a coefficients unimodulaires," *Bull. London Math. Soc.*, pp. 321–342, 1980, in french.

[77] E. van der Ouderaa, J. Schoukens and J. Renneboog, "Peak factor minimization using a time-frequency domain swapping algorithm," *IEEE Trans. Instrum. Meas.*, vol. 37, no. 1, pp. 145–147, March 1988.

[78] M. Friese, "Multitone signals with low crest-factor," *IEEE Trans. on Comm.*, vol. 45, no. 10, pp. 1338–1344, October 1997.

[79] G. Wunder and H. Boche, "Performance bounds and optimal pilot signals in OFDM-MIMO systems," in 10^{th} *Aachen Symposium on Signal Theory*, September 2001, pp. 100–106.

[80] S. Stanczak and H. Boche, "Aperiodic properties of binary Rudin-Shapiro sequences and a lower bound on the merit-factor of sequences with a quadratic phase function," in *ITG-Fachbericht, European Wireless 99*, Munich, Germany, October 1999, pp. 219–224.

[81] S. Wei, D.Goeckel, "A modern extreme value theory approach to calculating the distribution of the PAPR in OFDM systems," in *Proc. IEEE Int. Conf. on Comm.*, May 2002.

[82] C. Tellambura and M.G. Parker, "Relationship between Hamming weight and PMEPR of OFDM," in *IEEE ISIT'02*, Lausanne, Switzerland, June/July 2002, p. 245.

[83] M. R. Leadbetter, G. Lindgren and H. Rootzen, *Extremes and related properties of random sequences and processes*, Springer-Verlag, 1983.

[84] J.D. Forney and L. Wei, "Multidimensional constellations - Part I: Introduction, figures of merit, and generalized cross constellations," *IEEE Journal on Selected Areas in Commun.*, vol. 7, no. 6, pp. 877–892, August 1989.

[85] J.D. Forney and L. Wei, "Multidimensional constellations - Part II: Voronoi constellations," *IEEE Trans. Inform. Theory*, vol. 7, no. 6, pp. 941–958, August 1989.

[86] H. Rohling and V. Engels, "Differential amplitude phase shift keying (DAPSK)-A new modulation method for dtvb," in *Proc. IEE Int. Broadcast. Conv.*, September 1995, pp. 102–108.

[87] D. Slepian, "Group codes for the gaussian channel," *The Bell Syst. Tech. Journal*, vol. 47, pp. 575–602, April 1968.

[88] E. Biglieri, J.K. Karlof, and E. Viterbo, "Representing group codes as permutation codes," *IEEE Trans. on Inf.-Theory*, vol. 45, no. 6, pp. 2204–2207, September 1999.

[89] D. Slepian, "Permutation modulation," in *Proc. IEEE*, March 1965, vol. 53, pp. 228–236.

[90] T Kasami, T. Fujiwara, and S. Lin, "An approximation of the weight distribution of binary linear codes," *IEEE Trans. on Inf. Theory*, vol. 31, no. 6, pp. 769–780, November 1985.

[91] I.Krasikov and S.Litsyn, "On spectra of BCH codes," *IEEE Trans. on Inf. Theory*, vol. 41, no. 3, pp. 786–788, 1995.

[92] O. Keren and S.Litsyn, "More on the distance distribution of BCH codes," *IEEE Trans. on Inf. Theory*, vol. 45, no. 1, pp. 251–255, 1999.

[93] T Kasami, T. Fujiwara, and S. Lin, "An approximation to the weight distribution of binary primitive BCH codes with designed distances 9 and 11," *IEEE Trans. on Inf. Theory*, vol. 32, no. 5, pp. 706–709, September 1986.

[94] R. O'Neill and L.Lopes, "Performance of amplitude limited multitone signals," in *VTC'94*, June 1994, vol. 3, pp. 1675–1679.

[95] M. Renfors J. Rinne, "The behavior of orthogonal frequency division multiplexing signals in an amplitude limiting channel," *Proc. IEEE ICC*, pp. pp.381–385, May 1994.

[96] P. Banelli and S.Cacopardi, "Theoretical analysis and performance of OFDM signals in nonlinear AWGN channels," *IEEE Trans. on Comm.*, vol. 48, no. 3, pp. 430–441, March 2000.

[97] Q. Shi, "OFDM in bandpass nonlinearity," *IEEE Trans. on Cons. Elec.*, vol. 42, no. 3, pp. 253–258, August 1996.

[98] N.M. Blachman, "Detectors, bandpass nonlinearities, and their optimization: Inversion of the Chebyshev transform," *IEEE Trans. Inform. Theory*, vol. 17, no. 4, pp. 398–404, July 1971.

[99] J. Minkoff, "The role of AM-to-PM conversion in memoryless nonlinear systems," *IEEE Trans. Commun.*, vol. 33, no. 2, pp. 139–144, February 1985.

[100] R. Gross and D. Veeneman, "SNR and spectral properties for a clipped DMT ADSL signal," in *ICC*, July 1994, pp. 843–847.

[101] H. Kuchenbecker M. Pauli, "Neue Aspekte zur Reduzierung der durch Nichtlinearitäten hervorgerufenen Außerbandstrahlung eines OFDM-Signals," in *Proc. 2. OFDM-Fachgespräch*, Braunschweig, September 1997.

[102] M. Schilp, W. Sauer-Greff and W. Rupprecht, "Influence of oscillator phase noise and clipping on OFDM for terrestrial broadcasting of digital HDTV," in *ICC'95*, 1995, vol. 3, pp. 1678–1682.

[103] J.E. Mazo, "Asymptotic distortion spectrum of clipped, dc-biased, gaussian noise," *IEEE Tran. on Comm.*, vol. 40, no. 8, pp. 1339–1344, August 1992.

[104] N.M Blachman, "Gaussian noise-part I: The shape of large excursions," *IEEE Trans. Inform. Theory*, vol. 34, no. 6, pp. 1396–1400, November 1988.

[105] D. Wulich, N. Dinur, and A. Glinowiecki, "Level clipped high-order OFDM," *IEEE Trans. on Comm.*, vol. 48, no. 6, pp. 928–930, June 2000.

[106] D. Dardari, V. Trallin, and A. Vaccari, "A theoretical characterisation of nonlinear distortion effects in OFDM systems," *IEEE Trans. on Comm*, vol. 48, no. 10, pp. 1755–1764, October 2000.

[107] N.C. Beaulieu, "An infinite series for the computation of the complementary probability distribution function of a sum of independent random variables and its application to the sum of Raleigh random variables," *IEEE Trans. Commun.*, vol. 38, no. 0, pp. 2056–2057, September 1990.

[108] X. Li, L. Cimini, "Effects of clipping and filtering on the performance of OFDM," in *Proc. IEEE VTC'97*, Phoenix, May 1997, pp. 1634–1638.

[109] H. Rohling T. May, "Reduktion von Nachbarkanalstörungen in OFDM-Funkübertragungssystemen," in *Proc 2. OFDM-Fachgespräch*, Braunschweig, September 1997.

[110] M. Friese, "Multicarrier modulation with low peak-to-average power ratio," *Electr. Lett.*, vol. 32, no. 8, pp. 713–714, April 1996.

[111] S. Müller, J. Huber, "OFDM with reduced peak-to-average power ratio by optimum combination of partial transmit sequences," *Electron. Letters*, vol. 33, no. 5, pp. 368–369, February 1997.

[112] M. Friese, "OFDM signals with low crest-factor," in *IEEE Globecom*, New York,USA, November 1997, vol. 1, pp. 290–294.

[113] J. Huber R. Bäuml, R. Fischer, "Reducing the peak-to-average power ratio of multicarrier modulation by selected mapping," *Electron. Letters*, vol. 32, no. 22, pp. 2056–2057, October 1996.

[114] P. van Eetvelt, G. Wade, and M. Tomlinson, "Peak to average power reduction for OFDM schemes by selective scrambling," *Electron. Lett.*, vol. 32, no. 21, pp. 1963–1964, October 1996.

[115] J. Tellado, L. Hoo and J. Cioffi, "Maximum likelhood detection of nonlinearly distorted multicarrier symbols by iterative decoding," in *IEEE GlOBECOM*, Rio de Janeiro, Brazil, December 1999, pp. 2493–2498.

[116] J. Tellado and J.M. Cioffi, "Peak power reduction for multicarrier transmission," in *Proc. IEEE GlobeCom Commun. Theory MiniConf'98*, Sydney, Australia, November 1998, pp. 219–224.

[117] W. Henkel and B. Wagner, "Another application for Trellis shaping: PAR reduction for DMT (OFDM)," *IEEE Trans Commun.*, vol. 48, no. 9, pp. 1471–1476, September 2000.

[118] S.J. Shepherd, P.W.J. van Eetvelt, C.W. Wyett-Millington, and S.K. Barton, "Simple coding scheme to reduce peak factor in QPSK multicarrier modulation," *Electron. Lett.*, vol. 31, no. 14, pp. 1131–1132, July 1995.

[119] D. Wulich, "Reduction of peak to mean ratio of multicarrier modulation using cyclic coding," *Electron. Letters*, vol. 32, no. 5, pp. 432–433, February 1996.

[120] P. Fan and X. Xia, "Block coded modulation for the reduction of the PAPR in OFDM systems," *IEEE Trans. on Cons. Elec.*, vol. 45, no. 4, pp. 1025–1029, November 1999.

[121] H. Ahn, Y. Shin, and S. Im, "A block coding scheme for PAPR reduction in an ofdm system," in *IEEE Vehicular Technology Conf.*, 2000, pp. 56–60.

[122] A. Grant and R. van Nee, "Efficient maximum likelihood decoding of peak power limiting codes for OFDM," in *Proc 48th IEEE Vehicular Technology Conf.*, 1998, pp. 2081–2084.

[123] A.J. Grant and R.D.J. van Nee, "Efficient maximum likelihood decoding of Q-ary modulated Reed-Muller codes," *IEEE Commun. Lett.*, vol. 2, no. 5, pp. 134–136, May 1998.

[124] K.G. Paterson and A.E. Jones, "Efficient decoding algorithms for generalized Reed-Muller codes," *IEEE Trans. Commun.*, vol. 48, no. 8, pp. 1272–1285, August 2000.

[125] C.V. Chong and V. Tarokh, "A simple encodable/decodable OFDM QPSK code with low peak-to-mean envelope power ratio," *IEEE Trans on Inf.-Theory*, vol. 47, no. 7, pp. 3025–3029, November 2001.

[126] C. Rößling and V. Tarokh, "A construction of OFDM 16-QAM sequences having low peak power," *IEEE Trans. Tnform. Theory*, vol. 47, no. 5, pp. 2091–2094, July 2001.

[127] H. Bölcskei and A. J. Paulraj, "Space-frequency coded broadband OFDM systems," in *WCNC 2000*, Chicago, USA, September 2000, vol. 1, pp. 1–6.

[128] D. L. Goeckel and G. Ananthaswamy, "On the design of multidimensional signal set for OFDM systems," *IEEE Trans. on Comm.*, vol. 50, no. 3, pp. 442–452, March 2002.

[129] D.L. Goeckel, "Coded modulation for peak power constrained wireless OFDM systems," in *28th Allerton Conference on Communications and Control*, 1998.

[130] S. Shamai and I. Bar-David, "The capacity of average and peak-power-limited quadrature Gausssian channels," *IEEE Trans. Inform. Theory*, vol. 41, pp. 1060–1071, July 1995.

[131] H. Ochiai and H. Imai, "Channel capacity of clipped OFDM systems," in *IEEE ISIT'00*, Sorrento, Italy, June 2000, p. 219.

[132] H. Ochiai and H. Imai, "Performance analysis of deliberately clipped OFDM signals," *Trans. Commun.*, vol. 50, no. 1, pp. 89–101, January 2002.

[133] X. Li and J. A. Ritcey, "M-sequences for OFDM peak-to-average power ratio reduction and error correction," *Elec. Lett.*, vol. 33, no. 7, pp. 554–555, March 1997.

[134] J. Jedwab, "Comment: M-sequences for OFDM peak-to-average power reduction," *Electron. Lett.*, vol. 33, pp. 1293–1294, 1997.

[135] R. Lidl and H. Niederreiter, *Finite Fields*, vol. 20 of *Encyclopedia of Mathematics and its Applications*, Cambridge Univ. Press, Cambridge, U.K., 2ed edition, 1997.

[136] F.J. MacWilliams and N.J.A. Sloane, *The theory of error-correcting codes*, North-Holland, Amsterdam, The Netherlands, 2nd edition, 1986.

[137] S. Shepherd, J. Orriss, and S. Barton, "Asymptotic limits in peak envelope power reduction by redundant coding in OFDM," *IEEE Trans Commun.*, vol. 46, no. 1, pp. 5–10, January 1998.