Achievable Rate and Modulation for Bandlimited Channels with Oversampling and 1-Bit Quantization at the Receiver

Sandra Bender

Technische Universität Dresden

Achievable Rate and Modulation for Bandlimited Channels with Oversampling and 1-Bit Quantization at the Receiver

Dipl.-Ing.
Sandra Bender

der Fakultät Elektrotechnik und Informationstechnik der
Technischen Universität Dresden

zur Erlangung des akademischen Grades

Doktoringenieur

(Dr.-Ing.)

genehmigte Dissertation

Vorsitzender: Prof. Dr.-Ing. habil. Uwe Marschner
Gutachter: Prof. Dr.-Ing. Dr. h.c. Gerhard Fettweis
Gutachter: Prof. Dr.-Ing. Martin Bossert

Tag der Einreichung: 14.05.2020
Tag der Verteidigung: 14.10.2020

Abstract

Sustainably realizing applications of the future with high performance demands requires that energy efficiency becomes a central design criterion for the entire system. For example, the power consumption of the analog-to-digital converter (ADC) can become a major factor when transmitting at large bandwidths and carrier frequencies, e.g., for ultra-short range high data rate communication. The consumed energy per conversion step increases with the sampling rate such that high resolution ADCs become unfeasible in the sub-THz regime at the very high sampling rates required. This makes signaling schemes adapted to 1-bit quantizers a promising alternative. We therefore quantify the performance of bandlimited 1-bit quantized wireless communication channels using techniques like oversampling and faster-than-Nyquist (FTN) signaling to compensate for the loss of achievable rate.

As a limiting case, we provide bounds on the mutual information rate of the hard bandlimited 1-bit quantized continuous-time – i.e., infinitely oversampled – additive white Gaussian noise channel in the mid-to-high signal-to-noise ratio (SNR) regime. We derive analytic expressions using runlength encoded input signals. For real signals the maximum value of the lower bound on the spectral efficiency in the high-SNR limit was found to be approximately 1.63 bit/s/Hz.

Since in practical scenarios the oversampling ratio remains finite, we derive bounds on the achievable rate of the bandlimited oversampled discrete-time channel. These bounds match the results of the continuous-time channel remarkably well. We observe spectral efficiencies up to 1.53 bit/s/Hz in the high-SNR limit given hard bandlimitation. When excess bandwidth is tolerable, spectral efficiencies above 2 bit/s/Hz per domain are achievable w.r.t. the 95 %-power containment bandwidth. Applying the obtained bounds to a bandlimited oversampled 1-bit quantized multiple-input multiple-output channel, we show the benefits when using appropriate power allocation schemes.

As a constant envelope modulation scheme, continuous phase modulation is considered in order to relieve linearity requirements on the power amplifier. Noise-free performance limits are investigated for phase shift keying (PSK) and continuous phase frequency shift keying (CPFSK) using higher-order modulation alphabets and intermediate frequencies. Adapted waveforms are designed that can be described as FTN-CPFSK. With the same spectral efficiency in the high-SNR limit as PSK and CPFSK, these waveforms provide a significantly improved bit error rate (BER) performance. The gain in SNR required for achieving a certain BER can be up to 20 dB.

Zusammenfassung

Die nachhaltige Realisierung von zukünftigen Übertragungssystemen mit hohen Leistungsanforderungen erfordert, dass die Energieeffizienz zu einem zentralen Designkriterium für das gesamte System wird. Zum Beispiel kann die Leistungsaufnahme des Analog-Digital-Wandlers (ADC) zu einem wichtigen Faktor bei der Übertragung mit großen Bandbreiten und Trägerfrequenzen werden, z. B. für die Kommunikation mit hohen Datenraten über sehr kurze Entfernungen. Die verbrauchte Energie des ADCs steigt mit der Abtastrate, so dass hochauflösende ADCs im Sub-THz-Bereich bei den erforderlichen sehr hohen Abtastraten schwer einsetzbar sind. Dies macht Signalisierungsschemata, die an 1-Bit-Quantisierer angepasst sind, zu einer vielversprechenden Alternative. Wir quantifizieren daher die Leistungsfähigkeit von bandbegrenzten 1-Bit-quantisierten drahtlosen Kommunikationssystemen, wobei Techniken wie Oversampling und Faster-than-Nyquist (FTN) Signalisierung eingesetzt werden, um den durch Quantisierung verursachten Verlust der erreichbaren Rate auszugleichen.

Wir geben Grenzen für die Transinformationsrate des Extremfalls eines strikt bandbegrenzten 1-Bit quantisierten zeitkontinuierlichen – d.h. unendlich überabgetasteten – Kanals mit additivem weißen Gauß'schen Rauschen bei mittlerem bis hohem Signal-Rausch-Verhältnis (SNR) an. Wir leiten analytische Ausdrücke basierend auf lauflängencodierten Eingangssignalen ab. Für reele Signale ist der maximale Wert der unteren Grenze der spektralen Effizienz im Hoch-SNR-Bereich etwa 1,63 Bit/s/Hz.

Da die Überabtastrate in praktischen Szenarien endlich bleibt, geben wir Grenzen für die erreichbare Rate eines bandbegrenzten, überabgetasteten zeitdiskreten Kanals an. Diese Grenzen stimmen mit den Ergebnissen des zeitkontinuierlichen Kanals bemerkenswert gut überein. Im Hoch-SNR-Bereich sind spektrale Effizienzen bis zu 1,53 Bit/s/Hz bei strikter Bandbegrenzung möglich. Wenn Energieanteile außerhalb des Frequenzbandes tolerierbar sind, können spektrale Effizienzen über 2 Bit/s/Hz pro Domäne – bezogen auf die Bandbreite, die 95 % der Energie enthält – erreichbar sein. Durch die Anwendung der erhaltenen Grenzen auf einen bandbegrenzten überabgetasteten 1-Bit quantisierten Multiple-Input Multiple-Output-Kanal zeigen wir Vorteile durch die Verwendung geeigneter Leistungsverteilungsschemata.

Als Modulationsverfahren mit konstanter Hüllkurve betrachten wir kontinuierliche Phasenmodulation, um die Anforderungen an die Linearität des Leistungsverstärkers zu verringern. Beschränkungen für die erreichbare Datenrate bei rauschfreier Übertragung auf Zwischenfrequenzen mit Modulationsalphabeten höherer Ordnung werden für Phase-shift keying (PSK) and Continuous-phase frequency-shift keying (CPFSK)

untersucht. Weiterhin werden angepasste Signalformen entworfen, die als FTN-CPFSK beschrieben werden können. Mit der gleichen spektralen Effizienz im Hoch-SNR-Bereich wie PSK und CPFSK bieten diese Signalformen eine deutlich verbesserte Bitfehlerrate (BER). Die Verringerung des erforderlichen SNRs zur Erreichung einer bestimmten BER kann bis zu 20 dB betragen.

Acknowledgement

Foremost I would like to express my gratitude to Professor Gerhard Fettweis for giving me the opportunity to conduct this research at the Vodafone Chair Mobile Communication Systems at the Technische Universität Dresden. I am indebted to him not only for his guidance and support with regard to my research and the chance to work in the interdisciplinary collaborative research cluster "Highly Advanced Energy Efficient Computing" but also for entrusting me with managing the industry cooperation project with Vodafone, which provided interesting insights and complementary challenges to my field of responsibilities. I am furthermore grateful to Prof. Martin Bossert for serving as the second reviewer for this thesis.

I am indebted to Dr. Meik Dörpinghaus for sharing his experience and expertise in countless, long scientific discussions and for thoroughly proofreading this thesis and other publications of mine. My thank goes also to Dr. Heinrich Nuszkowski firstly for arousing my interest in this area of communication technology with his lectures many years ago and secondly for the collaboration and guidance when I was the teaching assistant for his mobile communication lecture. I am thankful to Dr. Stefan Krone for making me consider working towards a PhD degree and to Elisa Kaiser and Robert Schmoll for proofreading the outcome. Moreover, I wish to thank Prof. Lukas Landau for introducing me to this topic when supervising my Diploma thesis and for his help and fruitful discussions at the beginning of my PhD.

I would like to thank my colleagues, especially my group members Lukas Landau, Najeeb ul-Hassan, Christoph Jans, Albrecht Wolf, Martin Schlüter, and Peter Neuhaus, but also Philipp Schulz, Yaning Zou, Tom Hößler, Arturo Gonzalez, Nick Schwarzenberg, Eva Bolza-Schünemann, and Lucas Scheuvens for discussions, support, activities, and advise of any kind and for creating a great working atmosphere.

Finally, my thank goes to my family and friends, foremost to Paul for enduring those last months and to my parents and grandparents for their love and support. I therefore wish to dedicate this work to my late grandfathers, Bruno Bender and Herbert Röll.

Dresden, May 2020 Sandra Bender

Contents

1 Introduction

> *In a deep-submicron CMOS process, time-domain resolution of a digital signal edge transition is superior to voltage resolution of analog signals.*
>
> **— R. B. Staszewski [Sta12]**

1.1 Motivation and Scope

The ever increasing demand for performance of networks (wireless and wired) is more and more opposed by the limited resources available, especially w.r.t. power consumption. Thus, sustainably realizing applications of the future with high performance demands requires that energy efficiency becomes a central design criterion for the entire system as, e.g., in [FDC+19]. This clearly has an impact on the system design process of communication and computation systems: instead of separately designing building blocks for best performance under reasonable power consumption, all building blocks and interfaces need to jointly provide sufficient performance to meet an overall target while minimizing the energy consumed. This approach is for example described in [ABG+10] from a system level point of view as trade-off between system capacity, quality of service (QoS), and energy consumption of a mobile communication network.

The example in [FDC+19] aims for small scale, energy efficient high-performance compute nodes, e.g., for the mobile edge cloud. Its system design from software down to hardware over all layers is centered around energy efficiency. This includes the concept of the wireless backplane, which enables a new energy efficient computing architecture. In this case, the requirement on the wireless link is to deliver ultra-short range and energy efficient communication with high data rates. Naturally, for high data rates at short transmission range high carrier frequencies are promising, e.g., the extremely high frequency (EHF) band, given the large bandwidths available. In this case of short range communication with a line-of-sight (LOS) component between printed circuit boards, it has further been measured that such channels can be considered largely frequency flat, despite the large bandwidth [FuHLF13]. The targeted performance is to provide 100 Gbps at 100 to 300 GHz carrier frequency over 2 to 10 cm range.

This results in challenging requirements on the hardware, e.g., the analog-to-digital converter (ADC). In digital communications, it is typically assumed that the ADC at the receiver provides a sufficiently fine grained quantization of the magnitude of the received signal. However, for very high data rate short link communication the power consumption of the ADC becomes a major factor. The consumed energy per conversion step increases with the sampling rate [Mur18] such that high resolution ADCs become unfeasible in the sub-THz regime at the very high sampling rates required. This and the small voltage headroom of downscaled complementary metal–oxide–semiconductor (CMOS) technology [Sta12] make signaling schemes adapted to 1-bit quantizers an attractive alternative. The resulting advantage is that the 1-bit ADC is just a simple comparator and can be implemented with very low power consumption [SMS09] since, for example, an automatic gain control (AGC) may not be required [SPM09]. While this approach can reduce the power consumption, the resulting loss in data rate needs to be mitigated. By using oversampling, the loss in amplitude information can be partly recovered such that quantization resolution of the signal magnitude is traded off by resolution in time domain. Optimal communication over the resulting channel including the ADC requires a modulation and signaling scheme adapted to this specific channel. Thus, in this case hardware restrictions imply the choice of the signaling scheme resulting in a bottom-up approach in systems design.

When it comes to the achievable rate of such a communication scheme, early works target noise-free scenarios [Gil93; Sha94] while more recent studies focus on the low signal-to-noise ratio (SNR) domain [KL10] or the mismatched capacity [Zha12]. Moreover, simulation-based bounds are evaluated, e.g., in [LDF18; LDF17] for different signaling strategies, often w.r.t. the 90 %-power containment bandwidth as a bandwidth measure. However, an analytical evaluation of the channel capacity of the bandlimited 1-bit quantized oversampled additive white Gaussian noise (AWGN) channel in the mid-to-high SNR domain is still open. This capacity depends on the oversampling factor since Nyquist-sampling – like any other sampling rate – does not provide a sufficient statistic due to 1-bit quantization.

Thus, the scope of this work is to quantify the performance of bandlimited 1-bit quantized wireless communication channels targeting energy efficiency. Hereby, a central performance metric applied is the achievable rate or mutual information rate. Since the amplitude information is discarded in 1-bit quantized systems, techniques like oversampling and faster-than-Nyquist (FTN) signaling in combination with run-length modulation are applied to compensate for the loss of achievable rate. The overall objective is to derive bounds on the achievable rate that can be evaluated without a complex simulation framework, preferably either by analytical or numerical computations. Thus, the influence of varying system parameters on the performance can be evaluated with comparably small effort and insights on suitable signal designs can be obtained. Given the outlined application scenario of short range multigigabit/s

communication, we focus on the system performance in the mid-to-high SNR domain and over AWGN channels for our derivations. An additional concern is relieving the power consumption at the transmitter, especially caused by the power amplifier (PA), using constant envelope modulation. Here, the focus is on performance limits of the achievable rate in a noise-free scenario as well as the bit error rate (BER) performance.

1.2 Outline and Main Contributions

In the following, Chapter 2 is concerned with the energy consumption in wireless links, where especially PAs, ADCs, and channel decoders have been identified as power bottlenecks. Challenges as well as recent and expected advances w.r.t. these building blocks are reviewed. Additionally, previous results regarding the performance – especially the achievable rate – of 1-bit quantized oversampled communication systems are reviewed and summarized illustrating the need for analytical and numerical frameworks.

Chapter 3 provides bounds on the mutual information rate of the hard bandlimited 1-bit quantized continuous-time AWGN channel in the mid-to-high SNR regime. The continuous-time scenario is the limiting case of increasing the oversampling rate to infinity, however, it imposes challenges on the analysis. We derive analytic expressions for the obtained bounds, such that they can be used as benchmarks for 1-bit quantized communication systems and support design choices. Assumptions and approximations that need to be made in order to obtain analytical results are justified – if not by rigorous proof – by numerical computations and simulations.

Chapter 4 focuses on the achievable rate in a discrete-time scenario. Bounds on the achievable rate are derived that are simple to evaluate and match the results of Chapter 3 remarkably well. Given the popularity of low-resolution quantizers in multiple-input multiple-output (MIMO) scenarios, the results are applied to a MIMO-channel.

In Chapter 5 continuous phase modulation (CPM) modulation is considered as a constant envelope modulation scheme. Noise-free performance limits are investigated for phase shift keying (PSK) and continuous phase frequency shift keying (CPFSK) using higher-order modulation alphabets. Furthermore, adapted waveforms are designed that can be described as FTN-CPM. With the same spectral efficiency (SE) in the high-SNR limit as PSK and CPFSK, these waveforms provide a significantly improved BER performance. Finally, Chapter 6 provides the conclusion and an outlook.

1.3 Notation

Throughout this thesis the following notations are applied: vectors are set bold, random variables sans serif. Thus, $\mathbf{X}^{(K)}$ is a random vector of length K. Omitting the superscript denotes the corresponding random process $\mathbf{X}$ for $K \to \infty$. Sets are denoted in calligraphic font, in particular, the alphabet from which a random variable X can be chosen is denoted by $\mathcal{X}$.

Sequences of a variable x are denoted by x_l^k, where the sequence is considered from time instant l to k. If l corresponds to 1 (first symbol), it is omitted. This notation can also apply to sequences of vectors $\boldsymbol{x}_l^k$, random variables x_l^k, or random vectors $\mathbf{x}_l^k$ while x_k refers to the value of x at time instant k.

For information measures, $(\cdot)'$ denotes the corresponding rate. Bounds on any quantity x are referred to using the indices x_{UB} and x_{LB} for upper and lower bound, respectively. Furthermore, $(a)^+$ denominates the maximum of a and zero.

State of the Art

2

This work considers approaches to wirelessly communicate high data rates as energy efficiently as possible, where the focus is on short distances. One central measure in this regard is applying low-resolution 1-bit quantizers at the receiver. This chapter is therefore divided in two main parts: review of the energy consumption of a wireless transceiver chain and known results w.r.t. receivers applying coarse quantization.

2.1 Energy Consumption Considerations

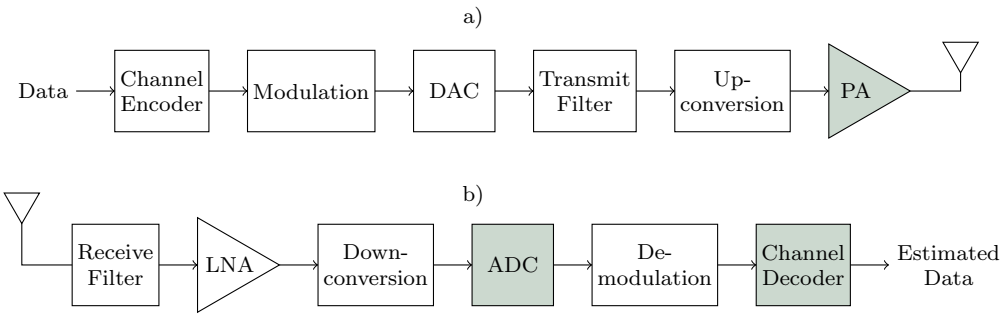

Fig. 2.1. Wireless transceiver chain: a) transmitter and b) receiver. Critical elements w.r.t. power consumption are highlighted in gray.

Fig. 2.1 depicts a typical transceiver chain in wireless communications. In [Lan17, Chapter 2.1], a wireless communication system has been analyzed with regard to the most power consuming elements. Although progress can be assumed in terms of the actual numbers since, it is safe to assume that the power bottlenecks that have been identified remained such. Thus, three elements can be considered as major energy consumers and are highlighted in gray in Fig. 2.1: the power amplifier (PA), the analog-to-digital converter (ADC), and the channel decoder. In the following, we will thus focus on these three and their contribution to the power consumption of a wireless communication system, especially with focus on transmission in the extremely high frequency (EHF) band, i.e., carrier frequencies between 30 GHz to 300 GHz. Furthermore, link budget considerations for wireless high-speed short range communications shall be reviewed in order to define the SNR-range, which is of interest for the operation of such a system.

2.1.1 Power Amplifier

Common efficiency measures for PAs are the drain efficiency, given by the ratio of output, i.e., transmitted power P_T divided by the power consumption P_{PA} of the PA

$$\eta_{\mathrm{PA}} = \frac{P_T}{P_{\mathrm{PA}}} \tag{2.1}$$

and the power added efficiency (PAE) given by

$$PAE = \frac{P_T - P_S}{P_{\mathrm{PA}}} = \frac{P_T}{P_{\mathrm{PA}}}\left(1 - \frac{1}{G_{\mathrm{PA}}}\right) \tag{2.2}$$

where P_S is the power of the signal to be amplified and G_{PA} is the gain of the amplifier [JHAS15]. With this, the power P_D dissipated by the PA can be written as [McC13]

$$\frac{P_D}{P_T} \approx \frac{1}{PAE} - 1 \tag{2.3}$$

given that PAs are designed to provide a high gain G_{PA}. From (2.3) one can compute that for a PAE of 50 % the dissipated power is equal to the output power while in order for the dissipated power to fall below 10 % of the output power, the PAE must be at least 91 %. These parameters highly depend the class of the PA as well as its desired operation point. Surveys [JHAS15; RKD12] show that with increasing operational frequency, the efficiency of the PA decreases drastically – in [RKD12] a drain efficiency drop from 90 % to around 40 % and below is observed for operating frequencies above 10 GHz for the highly efficient class E amplifiers. However, in wireless communications amplifiers of the classes A, B and AB are widely used [JHS12] given their good linearity characteristics. In [JHAS15] drain efficiencies between approximately 20 % and 30 % are reported for amplifiers operating in the ultra high frequency (UHF) band, 300 MHz to 3 GHz, and the super high frequency (SHF) band, 3 GHz to 30 GHz. Table 2.1 summarizes the parameters of recently published PAs in the SHF and EHF band. It confirms that obtaining high gains and efficiencies becomes increasingly difficult with increasing frequency.

This is especially problematic for operation in the EHF band. Circuit measurements are usually carried out using an unmodulated so-called continuous wave signal. However, in practice modulation techniques like orthogonal frequency-division multiplexing (OFDM) can have a high peak-to-average power ratio (PAPR) and require highly linear amplifiers. While a so-called input backoff, i.e., reducing P_T sufficiently below the peak output power of the amplifier, can reduce the distortion by clipping the signal, it also reduces the efficiency of the amplifier [JHS12]. This can be seen in Table 2.1, e.g., for [MLL17; PJJ+16; IR17] where the PAE has also been given for linear operation or under transmission of modulated data. There are techniques to improve the tradeoff between linearity and efficiency, such as outphasing, envelope elimination

Reference	Technology	f [GHz]	P_{Sat} [dBm]	Gain [dB]	P_{1dB} [dBm]	PAE [%]
[MLL17]	0.25 um GaN/SiC HEMT	15	34.1	28.2	32.2	38.7 (peak), 30 (linear)
[MLL18]	22 nm FD CMOS SOI	28	17.5	27.0	13.0	29.1 (peak), 14.1 (P1dB)
[PJJ+16]	28 nm CMOS	28	19.8	13.6	18.6	43.3 (peak), 28.5 (linear)
[NAU19]	130 nm SiGe	28	18.4	16.2		33 (peak)
		60	17.2	11.8		21 (peak)
[RHS+19]	22 nm CMOS FDSOI	29.45	16.3	11.9	13.9	23 (peak)
[IR17]	28nm bulk CMOS	32	19.8	22.0	16.0	21 (peak), 5.8 (64QAM)
[LNG+16]	90 nm CMOS	60	13.2	16.0		19.5 (peak)
		94	10.6	10.4		8.9 (peak)
[TNB19]	45 nm CMOS SOI	61	14.0	9.0		20.4 (peak)
[FRCE19]	0.13 um SiGe	61.5	14.6	18.0	12.0	11 (peak)
[GURP16]	250nm InP HBT	71-95	19.8-21.1	10.0-11.5	16.7	40 (peak), 10.5 (P1dB)
		96-120	19.3-20.2	10.0-13.0	17.7	21.2 (peak), 13.5 (P1dB)
[BHZ17]	130 nm SiGe	120	15.5	19.0		6.4 (peak)
[AFAS16]	130 nm SiGe	160	11.0	24.8	9.5	5 (peak)
[SALP16]	130 nm SiGe BiCMOS	215	9.6	25.0	4.0	0.5 (peak)

Tab. 2.1. Parameters of state-of-the-art power amplifiers

and restoration, envelope tracking or Doherty techniques [JHS12]. However, this shows that modulation techniques with low PAPR are promising for communication in the EHF band when it comes to energy efficiency. This includes continuous phase modulation (CPM) and modulation optimized for 1-bit quantization, e.g., runlength based modulation schemes, which are discussed in this work.

2.1.2 Data Converter

Data converters are often classified in two main types: (i) amplitude domain converters where the input signal amplitude is transformed in a binary number, (ii) time domain converters (TDCs) or frequency converters, in which the duration between zero-crossings or the frequency of the input signal is converted into the digital domain. While the former are well studied and used successfully for decades, TDCs have recently received increasing attention due to issues arising with continuing device

scaling. The resulting voltage supply scaling reduces the signal headroom for amplitude processing [Sta12] including difficult trade-offs w.r.t. noise, linearity, bandwidth, power consumption etc. and yielding performance limitations of analog circuits operating in the amplitude-domain (voltage and current) [Man18], whereas in time-domain circuits reduced switching times present a promising alternative.

Analog-to-Digital Converters

Regarding the power consumption of ADCs, [SMS09] presents a theory for lower bounding the power consumption of Nyquist sampling ADCs. By assuming that the noise power of sampling noise is designed to be equal to the quantization noise, it results as minimum power for sampling the input signal

$$P_{\mathrm{samp,min}} = 48 k_B T f_s 2^{2\mathrm{ENOB}} \tag{2.4}$$

where k_B is the Boltzmann constant, T is the temperature, and f_s is the sampling frequency. In the case of matched noise mentioned above, the effective number of bits (ENOB) is given by $\mathrm{ENOB} = b - 0.5$, where b is the quantizer resolution in bits. Furthermore, lower bounds on the power consumption of comparator elements P_{C} and amplifiers P_{A} are derived, which result to be proportional to $P_{\mathrm{samp,min}}$ when using the same criterion for sampling and quantization noise as above. This means that, neglecting other effects, the power consumption of an ADC increases at least by factor 2 for doubling f_s and factor 4 for adding 1 bit resolution.

A figure of merit (FoM) to relate the power consumption P_{ADC} of the ADC, f_s and the ENOB was introduced by Walden [Wal99] as

$$FoM_{\mathrm{W,ADC}} = \frac{2^{\mathrm{ENOB}} f_{\mathrm{samp,nyq}}}{P_{\mathrm{ADC}}} \tag{2.5}$$

where $f_{\mathrm{samp,nyq}}$ relates to f_s in the case of Nyquist sampling and to the maximum possible signal bandwidth in the case of oversampling architectures. Despite not capturing the relations between power consumption, resolution, and sampling frequency observed above as criticized, e.g., in [Mur08] where it is recommended for comparing low-resolution ADCs, Walden's FoM remains widely used. Often the inverse of (2.5) is used, e.g., in the annually updated ADC survey [Mur18], where it represents the energy spent per conversion step. In Fig. 2.2, said energy consumption in fJ per conversion step is depicted for state-of-the-art ADCs according to the data in [Mur18]. It can be observed that for sampling rates above 1 GHz it becomes increasingly difficult to build energy efficient ADCs with high resolution. Thus, this constitutes the so-called ADC-bottleneck and motivates the use of low-resolution or time-based converters with appropriate signaling strategies for wireless communication in the EHF band.

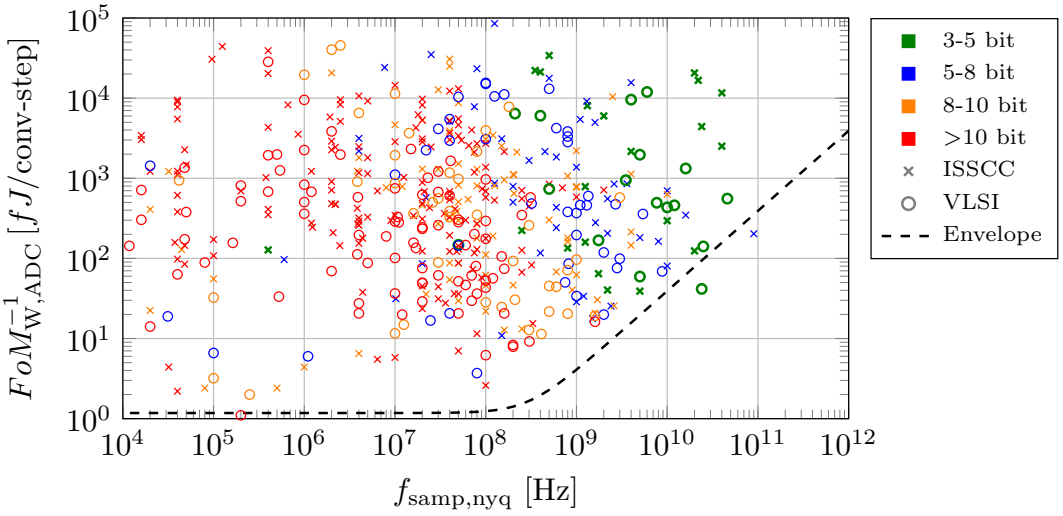

Fig. 2.2. FoM of (2.5) sorted by ENOB based on a survey of publications at the International Solid-State Circuits Conference (ISSCC) and the Symposium on VLSI Circuits (VLSI) 1997-2018 [Mur18]

Time-to-Digital Converters

It can be observed in Fig. 2.2 that the effective resolution, ENOB, of the ADCs is at least 3 bit. However, this does not mean that this number of bits is encoded in amplitude information. Using digital-to-time converters (DTCs) and TDCs, a number of bits can be encoded into the time differences between given events, e.g., the crossing of a threshold or zero-crossing (ZC).

Basic building blocks for TDCs and DTCs are inverters and D-type edge-triggered flip-flops [Man18]. While inverters are used to build voltage controlled delay units that later can be used for architectures based on ring-oscillators and voltage controlled delay lines, e.g., like in [UKCE18], the flip-flop can be used to realize a comparator function [Man18; Nar09]. Basic time-domain based data converter architectures can be found in [RA10; Nar09].

With regard to performance, it is observed that based on the data in [Mur18] "*the performance of the published architectures so far spans the medium signal-to-noise-and-distortion ratio (SNDR) range [and] medium bandwidth range*" and "*the energy efficiency isn't the most competitive, though recent data points seem to show appreciable improvement in conversion efficiency*" [Man18]. It is however argued further that beyond this first impression, time based data converters are generally more compact area wise and especially successful in embedded and mobile applications as well as low-cost, low-power sensing scenarios, e.g., for the Internet of Things (IoT). Moreover, it is expected that TDCs/DTCs will scale well – with respect to chip area but especially w.r.t. conversion efficiency – and that they will see improvement in jitter performance when transitioning from planar MOS to FinFET based architectures [Man18; Nar09].

2.1.3 Decoder

A decoder converts the received data stream into a sequence of symbols or bits, which could have been sent. Ideally, the decoder is able to return the sequence that has actually been sent – otherwise an error occurs in transmission. If this decision is based on the conditional probability density function (pdf) $p(\mathbf{y}|\mathbf{x})$ or likelihood function, where $\mathbf{y}$ and $\mathbf{x}$ are a the sequences of data at the channel output and channel input, respectively, then

$$\hat{\mathbf{x}} = \arg\max_{\mathbf{x}} p(\mathbf{y}|\mathbf{x}) \tag{2.6}$$

is called the maximum likelihood (ML) decoder [Nus15]. For general channels, the following issues can be associated with (2.6)

1. The likelihood function may not be known or not available in closed form.
2. Depending on the length of the sequences $\mathbf{y}$ and $\mathbf{x}$ and the size of the alphabet, the search space can become huge rapidly.

In order to address these problems, a multitude of decoder architectures has been developed like sequence- and symbol-based algorithms, e.g., the Viterbi algorithm [Vit67], the Bahl-Cocke-Jelinek-Raviv-algorithm (BCJR) [BCJR74], and iterative algorithms. The energy consumption estimation in [Lan17, Chapter 2.1] has been made based on an iterative decoder for low density parity check (LDPC) codes, which are for example used in 5G New Radio [RK18], yielding an estimate of 1116 mW for a communication system operating at 180 GHz with 20 GHz bandwidth.

Recently, interest in decoder architectures based on neural networks (NNs) is growing since an NN could even address the problem of an unkown likelihood function by learning it. NN-based decoders are considered to be efficient during run-time but to consume more energy during training phase. In conclusion, any modulation scheme that seeks to relieve the expense of energy in the transceiver, must consider the price of decoder complexity.

2.1.4 Link Budget

As outlined previously, this work focuses on the scenario of ultra-short range communication of high data rates with low energy consumption per transmitted bit. Specifically, we consider line-of-sight (LOS) transmission over a range of several centimeters. In [FDC+19], based on measured and estimated hardware parameters, a link budget calculation is made for transmitting at 180 GHz carrier frequency over a distance of approximately 2 to 10 cm for wireless communication within the chassis of a small-scale high performance compute node. Since all transceivers have fixed position, antenna arrays with a passive beam-switching network are used. Based on this architecture and assuming that the PA can provide 6 dBm output power, worst case (i.e., for the longest

link) receive signal-to-noise ratio (SNR) estimates of 13.6 dB are obtained. Thus, we are not interested in achievable data rates in the low-SNR regime of 0 dB or below but are rather looking for characterizing the achievable rate for SNRs values above of 5 dB or even 10 dB.

2.2 Relevant Signaling Concepts for 1-Bit Quantization

This section introduces methods and concepts, which are relevant in the context of 1-bit quantized wireless communication and which are applied in this work.

2.2.1 Faster-Than-Nyquist Signaling

Since its first introduction by Mazo in 1975 [Maz75] faster-than-Nyquist (FTN) signaling has received wide attention [ARO13; ET13; ZGQ+19]. The concept is directly related to the channel impulse response $h(t)$, which usually is orthogonal w.r.t. to shifts of multiples of the symbol duration T_{symb}. These pulses are also called Nyquist pulses. Due to the orthogonality, at the optimal sampling instances, neighboring symbols are invisible to each other and there is no intersymbol interference (ISI). If T_{Nyq} is the Nyquist interval, i.e., the time interval w.r.t. which $h(t)$ is orthogonal, and the signaling rate $f_{\mathrm{sig}} = {}^{1}/_{T_{\mathrm{symb}}}$ is increased beyond ${}^{1}/_{T_{\mathrm{Nyq}}}$, we speak of FTN signaling. Then, the FTN-factor is

$$M_{\mathrm{FTN}} = \frac{T_{\mathrm{Nyq}}}{T_{\mathrm{symb}}} = f_{\mathrm{sig}} T_{\mathrm{Nyq}}. \tag{2.7}$$

Originally, [Maz75] showed for binary transmission using sinc-pulses that the minimum distance $d_{\min}$ does not decrease below 2 (its value in the case of orthogonality) until $M_{\mathrm{FTN}} \approx 1.247$, meaning a 25 % increase of data rate. The limit, at which $d_{\min}$ falls below its orthogonal value, is known as Mazo limit. For other Nyquist-pulses like, e.g., raised cosine (RC) pulses, with some excess bandwidth beyond ${}^{1}/_{T_{\mathrm{Nyq}}}$ and, thus, less ISI due to a faster decay of $h(t)$, data rate can be increased by 30 % to 100 % at the same cost w.r.t. bandwidth and energy per bit. However, this is at the expense of a more complex receiver architecture, which is needed to resolve the ISI [ARO13; ET13]. The concept of the Mazo limit can be applied to non-binary transmission, non-orthogonal pulses, and non-linear modulation [ARO13]. It is known that the FTN capacity is

$$C_{\mathrm{FTN}} = \int_0^\infty \log_2\left(1 + \frac{2P}{N_0}|H(f)|^2\right) df \,\mathrm{bits/s} \tag{2.8}$$

where P is the signal power and $H(f)$ is its spectral distribution, normalized to unit integral over the real line. C_{FTN} exceeds the orthogonal pulse linear modulation

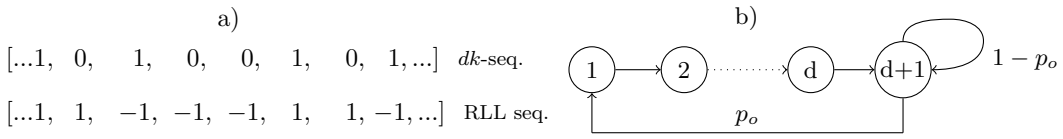

Fig. 2.3. Illustration of a) the conversion of a dk-sequence to a RLL sequence and b) the state diagram of a d-sequence ($k \to \infty$)

capacity for all except sinc-pulses [RA09]. The potential of augmenting the amount of data transmitted in a given time and with a given bandwidth, makes FTN signaling appealing for communication with 1-bit quantization in order to increase the achievable rate.

2.2.2 Runlength Limited Sequences

The concept of runlength limited (RLL) sequences has been widely studied in the field of magnetic and optical recording [Imm90]. Hereby, runlength refers to the number of consecutive alike symbols. We denote the runlength of the kth run of an RLL sequence by a_k. It is limited by two constraints: a minimum and a maximum runlength, respectively. RLL sequences are closely related to so called dk-sequences, which are binary sequences where every *one* has to be followed by at least d_{rll} and at most k_{rll} *zeros*. RLL sequences can be generated from dk-sequences by non-return-to-zero inverse (NRZI) coding, where every *one* translates into a zero-crossing. This is illustrated in Fig. 2.3a).

The minimum and maximum runlength then becomes $a_{\min} = d_{\mathrm{rll}} + 1$ and $k_{\mathrm{rll}} + 1$, respectively. The minimum runlength $a_{\min}$ controls the maximum transition frequency. Therefore, it has a major impact on the spectrum of the signal and, thus, the ISI when the signal is conveyed over a bandlimited channel. The maximum runlength on the other hand is relevant for facilitating clock recovery. In our work we are mainly interested in quantifying the impact of bandlimitation and, thus, ISI such that we consider $k_{\mathrm{rll}} \to \infty$. Fig. 2.3b) depicts the state diagram, which can be used to describe the resulting d-sequence, where the probability of returning to state one is given by p_o, which results in the occurrence of symbol *one*.

The entropy maximizing distribution for a positive discrete random variable with given expectation is the geometric distribution [Kap93, Section 1.9.32], which means that for max-entropic RLL sequences the runlength a_k is geometrically distributed [Imm90]. The capacity per sample of a d-sequence is given by [Imm90] as

$$C_{\mathrm{rll}}(d_{\mathrm{rll}}, \infty) = \log_2 \lambda_{\mathrm{rll}} \tag{2.9}$$

Tab. 2.2. Capacity per sample, average runlength, 90 % power containment bandwidth and corresponding spectral efficiency of RLL sequences weighted with a triangular pulse of duration $2T_s$

$a_{\min}$	1	2	3	4	5	6
$C_{\mathrm{rll}}(a_{\min})$ [bit]	1	0.6942	0.5515	0.4650	0.4057	0.3620
$\mathbb{E}[\mathsf{a}_k]$	2	3.618	5.148	6.63	8.08	9.506
$B_{90\%}T_s$	0.849	0.507	0.375	0.311	0.276	0.255
$\frac{C_{\mathrm{rll}}(a_{\min})}{B_{90\%}T_s}$ [bit/s/Hz]	1.178	1.369	1.472	1.495	1.469	1.417

with λ_{rll} being the largest real root of $z^{d_{\mathrm{rll}}+1} - z^{d_{\mathrm{rll}}} - 1 = 0$. For max-entropic sequences, which achieve C_{rll}, the spectrum of the d-sequence is given as [ZW88; BLDF16]

$$S_{\mathrm{rll}}(\omega) = \frac{1}{\mathbb{E}[\mathsf{a}_k]\sin^2\left(\frac{\omega}{2}\right)} \frac{1 - \left|\frac{e^{j(d_{\mathrm{rll}}+1)\omega}}{\lambda_{\mathrm{rll}}{}^{d_{\mathrm{rll}}}(\lambda_{\mathrm{rll}} - e^{j\omega})}\right|^2}{\left|1 + \frac{e^{j(d_{\mathrm{rll}}+1)\omega}}{\lambda_{\mathrm{rll}}{}^{d_{\mathrm{rll}}}(\lambda_{\mathrm{rll}} - e^{j\omega})}\right|^2}. \tag{2.10}$$

Intuitively, $S_{\mathrm{rll}}(\omega)$ becomes narrower with increasing $a_{\min}$ since the maximum transition frequency decreases. Thus, the runlength coding reduces the effective two-sided bandwidth $2W$ of the transmit signal such that it is favorable to choose channel filters with a bandwidth $2W < \frac{1}{T_s}$. It is known that from a system design point of view this is equivalent to increasing the signaling rate while maintaining the filter bandwidth [ARO13], i.e., FTN signaling. The spectral efficiency of such an FTN-RLL system depends on how fast $C_{\mathrm{rll}}(d_{\mathrm{rll}}, \infty)$ decreases compared to the reduction of the occupied bandwidth.

Table 2.2 illustrates the decrease of the capacity per sample and the corresponding increase of the average runlength $\mathbb{E}[\mathsf{a}_k]$ with $a_{\min}$. Furthermore, the normalized 90 % power containment bandwidth $B_{90\%}T_s$ is given assuming that the RLL sequence is transmitted using a triangular pulse of duration $2T_s$. As a result, we see an optimum of the maximum spectral efficiency w.r.t. to the 90 % power containment bandwidth $\frac{C_{\mathrm{rll}}(a_{\min})}{B_{90\%}T_s}$ at $a_{\min} = 4$. This indicates that RLL coding in combination with FTN signaling is a suitable way of boosting the spectral efficiency of a communication system that is restricted to a resolution of one bit at the ADC. We will see this confirmed in Section 2.3 as well as in the following chapters. Note, however, that the toy example in Table 2.2 does neither consider noise nor ISI due to suppression of out-of-band emissions. This will be considered in more detail in Section 4.1.

2.2.3 Intersymbol Interference

From the previous sections it becomes clear that the treatment and modeling of ISI is very relevant if one would like to apply these concepts to a communication system. Sequence estimation techniques are required at the receiver for detection and

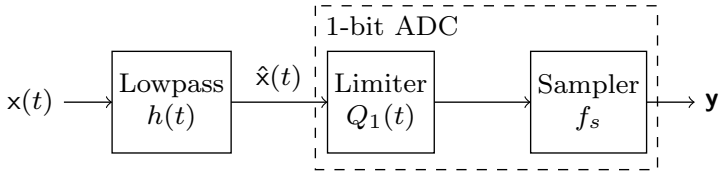

Fig. 2.4. Simplified baseband system model of a noise free bandlimited 1-bit quantized communication channel

it has been shown also in the context of 1-bit quantized communication systems that appropriate sequence design is beneficial in terms of achievable rate, cf. Section 2.3.2 or [LDF18; Lan17]. Furthermore, Section 2.3.2 presents a technique how sequence estimation, namely the BCJR, can be used to evaluate a lower bound on the achievable rate. However, as discussed below depending on the parameter choices this quickly becomes computational intensive.

Parts of this work are concerned with deriving low-complexity frameworks to evaluate achievable rates or spectral efficiencies of bandlimited 1-bit quantized wireless communication channels. In order to achieve this, we choose to model ISI as noise. Fig. 2.4 depicts a simplified baseband system model of a noise free bandlimited 1-bit quantized communication channel. Here, $h(t)$ is the overall channel impulse response, which is typically a lowpass (LP) filter. The distortion caused by LP filtering is then

$$\tilde{\mathsf{x}}(t) = \hat{\mathsf{x}}(t) - \mathsf{x}(t). \tag{2.11}$$

The more energy of $\mathsf{x}(t)$ is removed by filtering, the stronger the ISI $\tilde{\mathsf{x}}(t)$ becomes. The amount of power removed by LP filtering can be used to quantify the amount of ISI as

$$\mathbb{E}\left[\,|\tilde{\mathsf{x}}(t)|^2\,\right] = \mathbb{E}\left[\,|\hat{\mathsf{x}}(t) - \mathsf{x}(t)|^2\,\right] = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_{\hat{\mathsf{x}}}(\omega) - S_{\mathsf{x}}(\omega)d\omega \tag{2.12}$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} (|H(\omega)|^2 - 1)S_{\mathsf{x}}(\omega)d\omega \tag{2.13}$$

where $S_{\mathsf{x}}(\omega)$ and $S_{\hat{\mathsf{x}}}(\omega)$ are the power spectral density (PSD) of $\mathsf{x}(t)$ and $\hat{\mathsf{x}}(t)$, respectively, and $H(\omega)$ is the Fourier transform of $h(t)$.

2.3 Achievable Rate with 1-Bit Quantization

This section elaborates on previous results regarding the achievable data rate when considering systems with 1-bit quantization. This includes theoretical bounds and simulation-based approaches to evaluate the achievable rate. Furthermore, known results and definitions w.r.t. to the trade-off between time and amplitude resolution are discussed.

2.3.1 Time vs. Amplitude Domain

The trade-off between resolution in amplitude and resolution in time as illustrated from a hardware perspective in Section 2.1.2, is equally of interest from a communication and information theoretic point of view. In order to analyze this trade-off, early works [Gil93; Sha94] consider noise free scenarios and are based on system models similar to the one depicted in Fig. 2.4. Thereby a bandlimited process $\hat{\mathrm{x}}(t)$, which may be generated by filtering an input process $\mathrm{x}(t)$ using an ideal LP filter $h(t)$ with one-sided bandwidth W, is fed to a limiter or 1-bit quantizer. This limiter realizes what we will further refer to as 1-bit quantization function

$$Q_1(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ -1 & \text{if } x < 0 \end{cases}. \tag{2.14}$$

The output of the limiter is fed into a sampling unit, which maps the signal onto a discrete time grid. Given a finite signal bandwidth $2W$ and according to the sampling frequency f_s, the following cases can be distinguished

- **Nyquist sampling** ($f_s = 2W$): According to the Nyquist–Shannon sampling theorem [Sha49], an unquantized signal is completely described by its values at the sampling instances. The sampled signal provides a sufficient statistic and no information is lost due to sampling. Only the information loss due to quantization has to be considered.

- **Sub-Nyquist sampling** ($f_s < 2W$): Although information is lost due to sampling, this is still an interesting case especially in the context of rate-distortion theory and compressed sensing. If, for example, distortion is introduced due to quantization, the sampling rate can be reduced below the Nyquist rate without a penalty w.r.t. the rate-distortion trade-off [KGE15]. In the context of this work, this case is not of interest and merely mentioned for the sake of completeness.

- **Oversampling** ($f_s > 2W$): In the case of unquantized signals, oversampling would provide only redundant information, however, due to coarse quantization, the additional samples can help to compensate for the amplitude information lost through the limiter. The oversampling ratio (OSR) w.r.t. the signal bandwidth is defined as
$$M_{\mathrm{osr},f} = \frac{f_s}{2W}. \tag{2.15}$$

Another option to define oversampling is w.r.t. the symbol duration, i.e.,

$$M_{\mathrm{osr},t} = \frac{T_{\mathrm{symb}}}{T_s} \tag{2.16}$$

where T_s is the sampling period. We will refer to this case as *temporally* oversampled. Oversampling and temporal oversampling can be equivalent. Consider a transmission using Nyquist signaling and ideal LP filters with symbol duration $T_{\mathrm{symb}} = \frac{1}{2W}$ and sampling period T_s. Then

$$M_{\mathrm{osr},t} = \frac{T_{\mathrm{symb}}}{T_s} = \frac{f_s}{2W} = M_{\mathrm{osr},f}. \tag{2.17}$$

This changes when considering FTN signaling as introduced in Section 2.2.1. For an ideal LP with one-sided bandwidth W, we have $T_{\mathrm{Nyq}} = \frac{1}{2W}$. Thus, from (2.7) it can be seen that there is a connection between FTN signaling and oversampling, which yields for the relation of signaling and sampling rate

$$\frac{f_s}{f_{\mathrm{sig}}} = \frac{M_{\mathrm{osr},f} 2W}{M_{\mathrm{FTN}} 2W} = \frac{M_{\mathrm{osr},f}}{M_{\mathrm{FTN}}}. \tag{2.18}$$

If now $M_{\mathrm{osr},f} = M_{\mathrm{FTN}} > 1$, then $f_s = f_{\mathrm{sig}}$ and the system is oversampled w.r.t. its bandwidth but not w.r.t. the symbol duration. Both types of oversampling are used in this work and it will be indicated in the corresponding chapter, whether oversampling or temporal oversampling is applied.

Performance Results for Low-Resolution Quantizers and Nyquist Sampling

In the case of Nyquist signaling and Nyquist sampling, the maximum achievable rate is clearly limited by the quantizer resolution yielding

$$R_{\max} = \frac{b}{T_{\mathrm{Nyq}}}. \tag{2.19}$$

The performance loss by decreasing the quantizer resolution under finite SNR and different communication channels has been subject to various studies. In [SDM09], the impact of low-precision ADCs (1-3 bits) was analyzed with focus on Nyquist sampling for communication over an additive white Gaussian noise (AWGN) channel. It results that the optimal input distribution does not need to have more than $2^b + 1$ mass points. Furthermore, numerical computations show that up to moderate SNRs of 20 dB, quantization with 2 or 3 bits only yields 10-20 % reduction in spectral efficiency compared to unquantized observations and that fractionally spaced samples can provide small performance gains. In [ZSK12], it was shown that in the case of ISI channels, uniform quantizers are in general suboptimal. Moreover, a framework for the design of low-precision quantizers was developed that maximize a lower bound on the information rate for finite SNR. The ergodic capacity of Rayleigh fading channels with 1-bit quantization at the receiver was analyzed in [KF10].

Theoretical Results on Performance Limits with Oversampling

First results focus on noise free scenarios. In [Gil93], the achievable rate of a two-fold oversampled bandlimited process is given when using a Gaussian source as well as a discrete source with a four-letter alphabet and equally likely symbols. It is found that the achievable rate is 1.072 bits per Nyquist interval (bpni) for the Gaussian source and 1.049 bpni for the discrete source using evenly spaced samples. Based on one example, it is further shown that non-evenly spaced samples have the potential to increase the achievable rate. For the specific example, 1.089 and 1.063 bpni are obtained for the Gaussian and the discrete source, respectively.

Lower bounds on the achievable rate when oversampling a bandlimited process are first given in [Sha94] based on Zakai bandlimited processes. The bounds are

$$I_{\mathrm{LB},1}(\hat{\mathrm{x}};\mathbf{y}) = \log_2(M_{\mathrm{osr}}+1)\,\text{bpni} \tag{2.20}$$

$$I_{\mathrm{LB},2}(\hat{\mathrm{x}};\mathbf{y}) = \begin{cases} M_{\mathrm{osr}}\log_2 2\,\text{bpni} & M_{\mathrm{osr}}=1,2 \\ \log_2 M_{\mathrm{osr}} + (M_{\mathrm{osr}}-1)\log_2\left(\frac{M_{\mathrm{osr}}}{M_{\mathrm{osr}}-1}\right)\,\text{bpni} & M_{\mathrm{osr}}>2 \end{cases} \tag{2.21}$$

where the index $(\cdot)_{\mathrm{LB}}$ denotes the lower bound. The expression in (2.20) is at the same time an upper bound on the achievable rate when quantizing any input process in amplitude using a quantizer with M_{osr} thresholds, which illustrates well the potential of trading amplitude resolution for time resolution. Interesting questions arising from this idea include performance under noise as well as the design of suitable signaling schemes.

Considering the low SNR domain, in [KL10] a lower bound on the capacity per unit-cost, i.e., the slope of the capacity-vs-input-power curve at zero, is obtained for two-fold oversampling. It shows an increase of at least 17 % compared to Nyquist sampling. Considering more general SNR ranges, in [KF12] capacity expressions for non-bandlimited 1-bit quantized oversampled channels are obtained by numerical computation. Moreover, closed-form expression for the high-SNR-limit and an high-OSR-approximation are given. Promising results w.r.t. the achievable rate are obtained by computation using the cutting-plane algorithm, e.g., 4 bits per channel use (bpcu) for $M_{\mathrm{osr}}=16$ in the high-SNR. However, the study does not take into account that in practice independent noise samples at the receiver (as assumed) would lead to decreasing SNR with increasing OSR since the filter bandwidth would need to be increased. It further finds that transmission under coarse quantization can benefit from random dithering yielding higher achievable rates in the mid- than the high-SNR domain. Dithering has also found to be beneficial [SDF19a; SDF19b] w.r.t. to the performance limits of channel estimation for 1-bit quantized, oversampled channels. Based on derivation of the Fisher information and the Cramér-Rao lower bound, it has

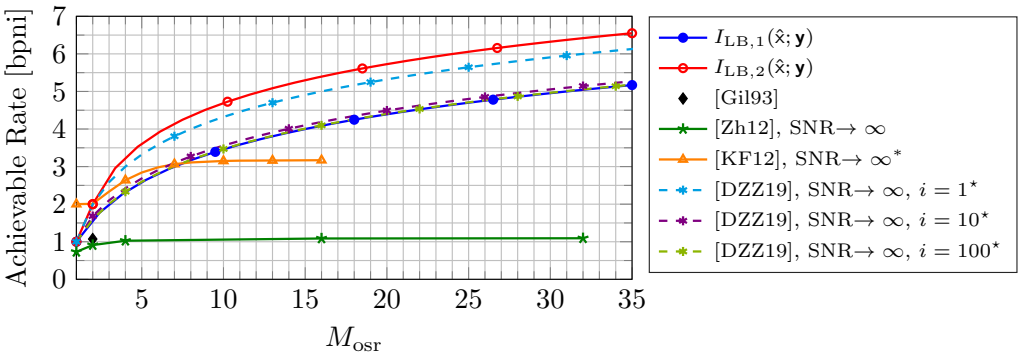

Fig. 2.5. Analytical and numerical results and bounds on the achievable rate of oversampled processes under 1-bit quantization in the high-SNR limit (*estimated from figure, not bandlimited, ⋆only approximately bandlimited due to truncation)

been shown that phase and timing dithering can improve the respective estimation performance.

In [Zha12], a framework was derived for analyzing channels with non-linear distortion using the performance measure of generalized mutual information (GMI) with a Gaussian codebook ensemble and nearest-neighbor decoding. It is shown that that oversampling increases the GMI for all SNRs. Furthermore, w.r.t. 1-bit quantized channels, the GMI can exceed 1 bpcu by approximately 0.1 bpcu in the high-SNR limit. These rather small improvements can be explained by the fact that the GMI represents a mismatched capacity [GLT00], where the tightness of the lower bound depends strongly on the choice of the input distribution [Zha12]. In [DZZ19], a transceiver is designed based on truncation of the pulses of infinite duration in [Sha94]. By employing an integrate-and-dump filter at the receiver, which is in general suboptimal, a discrete memoryless channel (DMC) model can be obtained facilitating the analysis. The study provides analytical results in the high-SNR limit as well as numerical results on the achievable rate showing that reliable communication at rates exceeding one bit per Nyquist interval is possible in the mid-SNR domain and that performance comparable to (2.20) can be achieved. Where applicable, the results reviewed in this section have been included in Fig. 2.5. Note that the results in [DZZ19] are not strictly bandlimited due to truncation of the pulses to integer multiples of the Nyquist interval iT_{Nyq}. For $i \to \infty$, the impact of truncation vanishes and the achievable rate in the high-SNR limit converges to (2.20).

2.3.2 Simulation-based Performance Evaluation of Oversampled 1-bit Quantized Communication

Performance evaluation of any communication scheme requires the choice of a performance metric. Very common metrics chosen are the bit error rate (BER) performance

and the achievable rate or mutual information rate, which can both be evaluated by Monte-Carlo simulation. Hereby, the main difference between the latter two is that achievable rate requires the existence of a coding theorem. While BER-simulations are comparably straight forward to implement, for obtaining the information rate some thought is required.

Obtaining the Mutual Information Rate by Simulation

Consider two random processes $\mathbf{A}$ and $\mathbf{D}$. If $\mathbf{A}$ and $\mathbf{D}$ are discrete, the mutual information I between two random vectors $\mathbf{A}^{(K)}$ and $\mathbf{D}^{(K)}$ of length K is defined as

$$I(\mathbf{A}^{(K)};\mathbf{D}^{(K)}) = \sum_{(\boldsymbol{a},\boldsymbol{d})\in\mathcal{A}^K\times\mathcal{D}^K} p_{\mathbf{A},\mathbf{D}}(\boldsymbol{a},\boldsymbol{d}) \log_2 \frac{p_{\mathbf{A},\mathbf{D}}(\boldsymbol{a},\boldsymbol{d})}{p_{\mathbf{A}}(\boldsymbol{a}) p_{\mathbf{D}}(\boldsymbol{d})} \tag{2.22}$$

where $p_{\mathbf{A},\mathbf{D}}(\boldsymbol{a},\boldsymbol{d})$ is their joint and $p_{\mathbf{A}}(\boldsymbol{a})$ and $p_{\mathbf{D}}(\boldsymbol{d})$ are their individual probability distributions, respectively. Furthermore, $\mathcal{A}^K$ and $\mathcal{D}^K$ are the corresponding alphabets. The mutual information rate I', i.e., the amount of data transmitted per unit time is then given as

$$I'(\mathbf{A};\mathbf{D}) = \lim_{K\to\infty} \frac{1}{K} I(\mathbf{A}^{(K)};\mathbf{D}^{(K)}) = H'(\mathbf{D}) - H'(\mathbf{D}|\mathbf{A}) \tag{2.23}$$

where H' is the entropy rate. When oversampling w.r.t. the signaling rate is used, every symbol A_k at the input generates a vector of observations $\boldsymbol{D}_k$ at the channel output. For stationary processes, applying the Shannon-McMillan-Breimann theorem to (2.23), we have

$$I'(\mathbf{A};\mathbf{D}) = \lim_{K\to\infty} \left\{ -\frac{1}{K} \log_2 p(\boldsymbol{D}^{(K)}) \right\} - \lim_{K\to\infty} \left\{ -\frac{1}{K} \log_2 p(\boldsymbol{D}^{(K)}|\boldsymbol{A}^{(K)}) \right\}. \tag{2.24}$$

For independent and identically distributed (i.i.d.) input symbols, it is known that independent channel realizations yield a lower bound on the mutual information rate

$$I'(\mathbf{A};\mathbf{D}) = \lim_{K\to\infty} \frac{1}{K} I(\mathbf{A}^{(K)};\mathbf{D}^{(K)}) \geq I(\mathsf{A}_k;\mathsf{D}_k) \tag{2.25}$$

where k is the time index of the current symbol. Thus, a memoryless channel results and (2.24) can be obtained by Monte-Carlo simulation on symbol basis. If, however, the channel does have memory, an algorithm to obtain the mutual information rate between $\mathbf{A}$ and $\mathbf{D}$ is described in [ALV+06].

Under the assumption that the channel is ergodic and can be represented by a finite state machine (FSM), [ALV+06] shows that $I'(\mathbf{A};\mathbf{D})$ can be obtained via an algorithm based on generating two very long sequences A^K and $\boldsymbol{D}^K$ and evaluating (2.24) numerically. If the channel law is not exactly known or memory-wise too computationally expensive, an auxiliary channel law satisfying $\tilde{p}(\cdot) > 0$ whenever

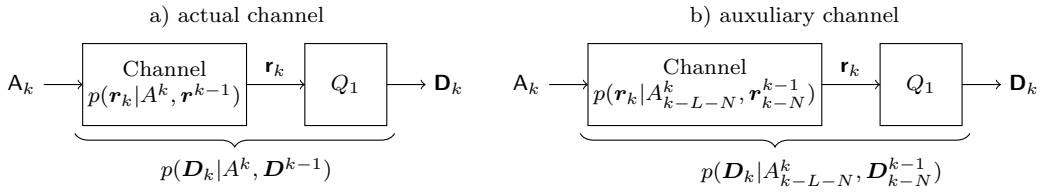

Fig. 2.6. Actual and auxiliary channel for simulation-based evaluation of the achievable rate

$p(\cdot) > 0$ will yield a lower bound on $I'(\mathbf{A}; \mathbf{D})$ [ALV+06]. Thus, limiting the channel memory w.r.t. input and output by L and N symbols, respectively, using

$$\tilde{p}(\boldsymbol{D}_k|A^k, \boldsymbol{D}^{k-1}) = p(\boldsymbol{D}_k|A_{k-L-N}^k, \boldsymbol{D}_{k-N}^{k-1}) \tag{2.26}$$

yields an *auxiliary channel lower bound*. This is illustrated in Fig. 2.6. Note that the accuracy of the results depends on the length K of the sampled sequences. For the computation of $\tilde{p}(\boldsymbol{D}^K)$ and $\tilde{p}(\boldsymbol{D}^K|A^K)$, the forward recursion of the BCJR [BCJR74] can be used based on (2.26). Hereby the computational complexity strongly depends on the number of states in the FSM and, thus, the memory $L + N$. Based on the auxiliary channel, [ALV+06] also gives an upper bound on the mutual information rate. In [LF14a], this technique is applied specifically to oversampled 1-bit quantized communication channels confirming the benefit of oversampling in terms of the achievable rate in this context.

For the evaluation of (2.24), assuming 1-bit quantization and oversampling, the transition probabilities in the BCJR need to be obtained via numerical integration over a multivariate distribution. The probability that the kth oversampled, quantized received vector is the ith element in $\mathcal{D}^{M_{\text{osr}}}$ when the jth element in $\mathcal{A}$ was sent, is

$$\Pr(\mathbf{D}_k = \boldsymbol{D}_i|\mathsf{A}_k = A_j) = \int_{\mathbb{D}_i^{M_{\text{osr}}}} \cdots \int p_{\mathbf{r}|\mathsf{A}}(\boldsymbol{r}|A_j) d\boldsymbol{r} \tag{2.27}$$

where $\mathbf{r}$ is the unquantized received vector and $\mathbb{D}_i^{M_{\text{osr}}}$ is the subspace of $\mathbb{R}^{M_{\text{osr}}}$ which is mapped to $\boldsymbol{d}_i$. While (2.27) can be computed offline, i.e., prior to the execution of the BCJR, for every combination of input and output symbol, it needs to be repeated for every slight variation of the channel. Its dimension is M_{osr} for real channels and $2M_{\text{osr}}$ for complex channels.

Simulation-based Performance Results

The BER performance has been investigated for example in [HLF14b] based on amplitude shift keying (ASK) transmission over an ISI channel using a BCJR decoder. It is shown that the BER can be improved by oversampling and a specific design of the

input distribution. The optimization of the input distribution was refined in [SLF15]. However, the choice of the input distribution will have an impact on the achievable data rate, which is not visible directly when considering BERs. Thus, the mutual information rate is a widely used metric, e.g., in [HLF14a] where with (2.25) the benefit of FTN signaling for i.i.d. 4- and 16-quadrature amplitude modulation (QAM) is shown for a 1-bit quantized channel. Furthermore, [LDF18] summarizes results on the achievable rate for different signaling strategies, namely optimized sequences of ASK symbols [LF14b], as well as RLL sequences and sequences optimized using the Blahut-Arimoto algorithm [LDF15]. Due to sequence design and FTN signaling, the achievable rate cannot be fairly compared in bpcu but either in bits per Nyquist interval or as spectral efficiencies. For the latter, a reference bandwidth is required, e.g., the 3dB-bandwidth or a power containment bandwidth. In [LDF18] the 90 % power containment bandwidth and 95 % power containment bandwidth are chosen, while in [DZZ19] (cf. Section 2.3.1), the 95 % power containment bandwidth is chosen for spectral efficiency evaluation. Both show spectral efficiencies comparable to (2.20). In [LDF17], the signaling strategies in [LDF15] are applied to bandlimited channels, which reduces the achievable rate but still achieves more than 1 bits per Nyquist interval.

Furthermore, channel filter optimization was carried out based on compliance with a spectral mask and using FTN signaling [BLDF16] yielding 10 to 20 % performance improvements in terms of achievable rate compared to Gaussian shaped filtering. In [LDdF18], CPM is considered for data transmission over 1-bit quantized oversampled channels. Also for this case oversampling is beneficial as the additional samples provide information on phase transitions, which convey the information. The resulting spectral efficiency for the 90 % power containment bandwidth is slightly lower compared to the previous studies while gaining the advantages of the constant envelope modulation, i.e., the option of using non-linear PAs, cf. Section 2.1.1.

For high SNR, the spectral efficiencies of the different approaches are compared in Fig. 2.7 based on the effective oversampling ratio

$$M_{\mathrm{osr,eff}} = \frac{M_{\mathrm{osr,t}}}{B_x T_{\mathrm{symb}}} = \frac{M_{\mathrm{osr,t}} M_{\mathrm{FTN}}}{B_x T_{\mathrm{Nyq}}} \tag{2.28}$$

where B_x is the bandwidth measure applied. Note that the results are based on different bandwidth measures; if not stated otherwise $B_x = B_{90\%}$ is the 90 % power containment bandwidth. Fig. 2.7 especially underlines the suitability of FTN signaling for increasing the spectral efficiency under 1-bit quantization.

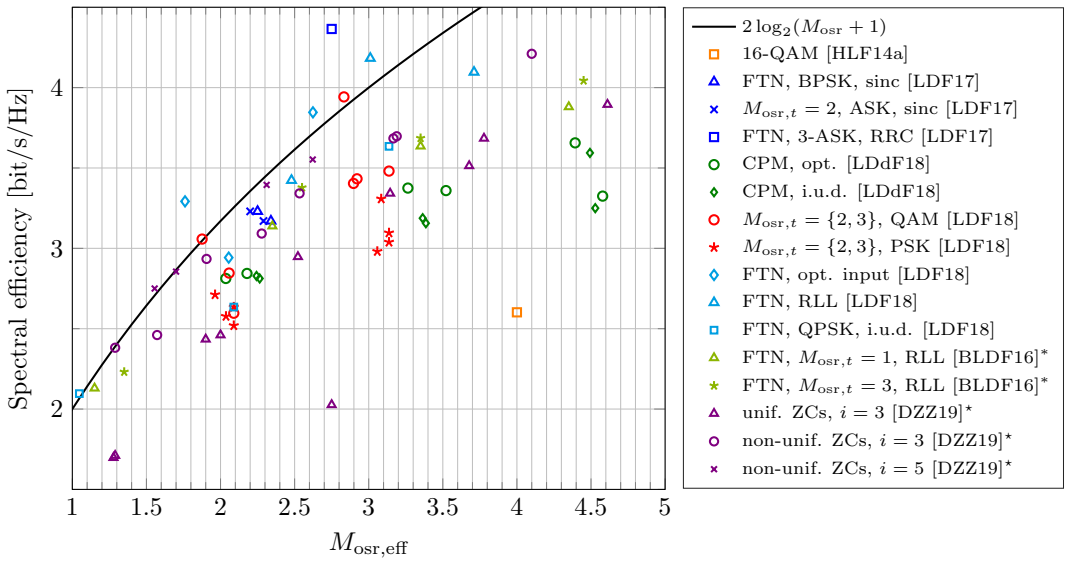

Fig. 2.7. Simulation-based results on the achievable rate of oversampled 1-bit quantized communication strategies in the high-SNR regime ($^*B_x = B_{3\mathrm{dB}}$, $^\star B_x = B_{95\%}$, pulse length of iT_{Nyq})

2.3.3 1-Bit Quantization for Multiple-Input Multiple-Output Systems

So-called multiple-input multiple-output (MIMO) channels are very attractive especially for communication in the range of millimeter wave (mmWave) carrier frequencies. This is due to the small size of antennas at these frequencies allowing for reasonably sized antenna arrays and the larger bandwidths available in mmWave regime, which both boost the achievable data rate. However, this makes mmWave MIMO channels even more affected by the ADC-bottleneck discussed in Section 2.1.2 since not only one but many fast ADCs are required for the many radio frequency (RF) chains. This is reflected in a multitude of research regarding communication over MIMO channels using few-bit and especially 1-bit quantizers. Research questions range from beam-forming strategies and pre-coding, over channel estimation, to detection performance and achievable data rates [ZDL+18]. Fig. 2.8 depicts a typical system model with N_t transmit and N_r receive antennas, which can be summarized as

$$\mathbf{y}_k = Q_1\left(\mathbf{H}\mathbf{C}\mathbf{x}_k + \hat{\mathbf{n}}_k\right). \tag{2.29}$$

Here, $\mathbf{y}$ is the $N_r \times 1$ sampled receive vector, $\mathbf{H}$ is the channel state matrix, $\mathbf{C}$ is the precoding operation, $\mathbf{x}$ is the $N_t \times 1$ transmit vector, and $\hat{\mathbf{n}}$ is the $N_r \times 1$ sampled and filtered AWGN vector. Variations of this model include the number of RF chains per quantizer [RPSN18] or a small number of high-resolution quantizers scattered in-between the 1-bit quantizers [PS17]. Other works consider multi-user MIMO channels, e.g., [JDC+17; ÜY18; MCLH17]. Mostly, the analysis is focused on Nyquist-signaling

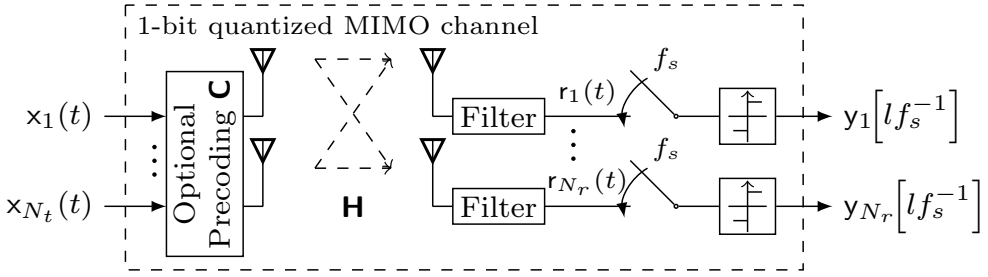

Fig. 2.8. Typical system model for investigations on 1-bit quantized multiple-input multiple-output channels

and sampling, however, temporal oversampling is considered for example in [GBLV17; ÜY18; SLd18]. The performance impact of spatial oversampling is investigated in [HL17].

With respect to the achievable rate, simulation-based performance evaluations are plenty, e.g, [JDC+17; RPL14; HL17], where the latter is considering spatial oversampling. Analytical results can be found for example in [MN07; MN12] and [MH15]. A common approach for analyzing the achievable rate, e.g., in [MN12; DJM+17] is based on Bussgang's decomposition [Bus52]. This means that the output of the nonlinear quantizer is decomposed into a desired signal component and an uncorrelated distortion, which represents a linearization using an effective channel $\mathbf{H}'$. In [MH15], the focus is on the high-SNR capacity and among other results, lower bounds on the capacity of the 1-bit quantized MIMO channel are given based on channel inversion (CI) and the additive quantization noise model (AQNM). For quantitative comparison of selected results the reader is referred to Section 4.2.

3 On the Achievable Rate of the 1-Bit Quantized Continuous-Time Channel

In this chapter, we study the achievable rate of the 1-bit quantized continuous-time additive white Gaussian noise (AWGN) channel. Continuous-time hereby means infinite resolution in time at the receiver and can be interpreted as the limiting case of increasing the oversampling rate to infinity. Like for the AWGN channel capacity, which has been given by Shannon [Sha48], without time quantization we have no quantization in the information carrying dimension. However, the non-linearity imposed by 1-bit quantization makes the analysis more involved.

We therefore base the analysis on a specially constructed input signal. Since with 1-bit quantization the information is effectively transmitted in the zero-crossings (ZCs) of the transmitted signal, we encode the data into the distances of these ZCs. This approach has similarities to the channel considered in [AV96]. The slope of the signal at the ZC is then limited by the available bandwidth, which is why we will consider a transmission pulse shaping the transition and, thus, to some extend the spectrum. We will see that besides well-known magnitude errors in the transmitted symbols this channel is prone to synchronization errors (insertions and deletions). Furthermore, the bandlimitation requires a model of the intersymbol interference (ISI), which we will base on Section 2.2.3.

Analytic expressions for upper and lower bounds on the mutual information rate are obtained based on our specific signaling scheme, which can be used as benchmarks for 1-bit quantized communication systems and gives insights on design choices. In order to treat the problem analytically, assumptions and approximations are needed, which are justified – if not by rigorous proof – by numerical computations and simulations. For a special case of a sine-shaped transition waveform, an early model has been published in [BDF17a]. Parts of this chapter with the results for the special case of a sine-shaped transition waveform have been submitted for publication to EURASIP Journal for Wireless Communication and Networks [BDFa].

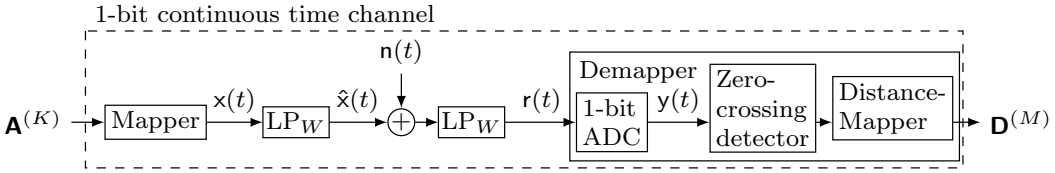

Fig. 3.1. System model for the bandlimited 1-bit quantized continuous-time channel

3.1 System Model

The analysis is based on the system model depicted in Fig. 3.1. Given a 1-bit quantizer, we must recover all information from the sequence of time instants of ZCs[1]. Thus, the channel input and output vectors $\mathbf{A}^{(K)} = [\mathsf{A}_1, ..., \mathsf{A}_K]^T$ and $\mathbf{D}^{(M)} = [\mathsf{D}_1, ..., \mathsf{D}_M]^T$ contain the temporal distances of two consecutive ZCs of the signal $\mathsf{x}(t)$ and the received signal $\mathsf{r}(t)$, respectively. As the resolution in time is infinite, A_k and D_m are continuous random variables. Note that K is not necessarily equal to M since the noise can add or remove ZCs.

3.1.1 Signal Structure and Input Distribution

The mapping between the random vector $\mathbf{A}^{(K)}$ and the continuous-time signal $\mathsf{x}(t)$ is depicted in Fig. 3.2. The signal $\mathsf{x}(t)$ alternates between two levels $\pm\sqrt{\hat{P}}$ with $\hat{P}$ being the peak power of $\mathsf{x}(t)$. The time when the kth transition between the levels $\pm\sqrt{\hat{P}}$ begins, is defined as

$$\mathsf{T}_k = \sum_{i=1}^{k} \mathsf{A}_i + t_0. \tag{3.1}$$

where without loss of generality we set $t_0 = 0$. In order to reduce the bandwidth of the signal, the transitions between the signal levels are controlled by a pulse shape

$$g(t) = \left(1 + \zeta\left(t - \frac{\beta}{2}\right)\right) \cdot \mathbb{1}_{[0,\beta]}(t) + 2 \cdot \mathbb{1}_{[\beta,\infty)}(t) \tag{3.2}$$

where $\zeta(t)$ is an odd function between $(-\beta/2, -1)$ and $(\beta/2, 1)$ and zero otherwise. Furthermore, $\mathbb{1}_{[u,v]}(x)$ is the indicator function being one in the interval $[u, v]$ and zero otherwise. The resulting time instant of the kth ZC is

$$\mathsf{T}'_k = \mathsf{T}_k + \frac{\beta}{2}. \tag{3.3}$$

[1] One additional bit is carried by the sign of the first sample. However, its effect on the mutual information rate between channel input and output can be neglected as it converges to zero for infinite blocklength.

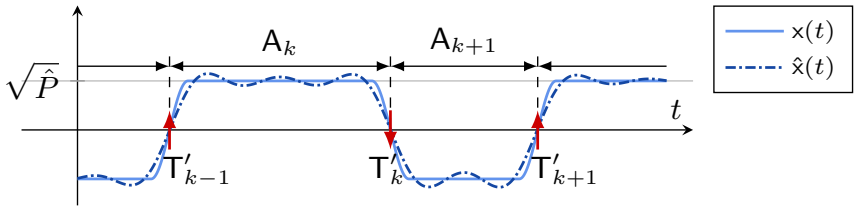

Fig. 3.2. Mapping input sequence $\mathbf{A}^{(K)}$ to $\mathsf{x}(t)$ and transmit signal $\hat{\mathsf{x}}(t)$ for a sine-shaped transition

This can be seen in Fig. 3.2, where the solid curve depicts $\mathsf{x}(t)$. The transition from the level $\sqrt{\hat{P}}$ begins at time $\mathsf{T}_k = \mathsf{T}'_k - \beta/2$ and follows $\zeta(t)$ – in this example sine-shaped – until time $\mathsf{T}_k + \beta$, where the level is maintained until time T_{k+1}. Thus, $\zeta(t)$ describes the waveform and β the duration of the transition.

The input symbols A_k are the temporal distances between the kth and the $(k-1)$th ZC of $\mathsf{x}(t)$. The signal $\mathsf{x}(t)$ is given by the superposition of pulses $g(t)$ with alternating signs as

$$\mathsf{x}(t) = \left(\sum_{k=1}^{K} \sqrt{\hat{P}} (-1)^k g(t - \mathsf{T}_k) \right) + \sqrt{\hat{P}}. \tag{3.4}$$

The A_k are considered to be independent and identically exponentially distributed with minimum length β, i.e.,

$$p_{\mathsf{A}_k}(a) = \lambda e^{-\lambda(a-\beta)} \mathbb{1}_{[\beta,\infty)}(a). \tag{3.5}$$

The distribution in (3.5) is chosen since it is the entropy maximizing distribution for positive continuous random variables with mean T_{avg}. The mean symbol duration of A_k is then given by

$$T_{\text{avg}} = \frac{1}{\lambda} + \beta \tag{3.6}$$

and the variance of the input symbols is

$$\sigma^2_{\mathsf{A}} = {}^1\!/\!_{\lambda^2}. \tag{3.7}$$

The distribution of the T_k or any other sum L of a subset of the A_k is then the Erlang-distribution, which is known to be the distribution of a sum of exponentially distributed random variables and which is a special case of the Gamma-distribution. Thus,

$$p_{\mathsf{T}_k}(t) = \frac{\lambda^k e^{-\lambda(t-k\beta)} (t - k\beta)^{k-1}}{(k-1)!}, t \geq k\beta. \tag{3.8}$$

By choosing β to be the minimum value of the A_k, we can guarantee that $\mathsf{x}(t)$ reaches the level $\sqrt{\hat{P}}$ between two transitions. While this is not necessarily capacity-achieving, it facilitates the derivation of a lower bound on the mutual information rate.

3.1.2 Channel Model

The signal $\mathsf{x}(t)$ is lowpass (LP)-filtered yielding the transmit signal $\hat{\mathsf{x}}(t)$, which is transmitted over an AWGN channel. The LP-filters at transmitter and receiver are considered to be ideal LPes with one-sided bandwidth W and amplitude one. The available bandwidth W is related to the transition time β, where the latter is chosen such that

$$\beta = \frac{1}{2W}. \tag{3.9}$$

Then, in the limiting case of $\lambda \to \infty$, $\mathsf{x}(t)$ has one-sided signal bandwidth W. However, for $\lambda < \infty$ the signal $\mathsf{x}(t)$ is not strictly bandlimited, such that a distortion

$$\tilde{\mathsf{x}}(t) = \hat{\mathsf{x}}(t) - \mathsf{x}(t) \tag{3.10}$$

is introduced by LP-filtering. Further implications of (3.9) will be discussed in Sections 3.1.3 and Appendix A.2. The relation between the channel bandwidth and the randomness at the channel input is captured via the normalized bandwidth

$$\kappa = {}^{W}\!/_{\lambda}. \tag{3.11}$$

The filtered transmit signal $\hat{\mathsf{x}}(t)$ is depicted in Fig. 3.2 as the dash-dotted curve. It can be seen that the main effect of filtering is a reduced slope at the ZCs as well overshoot and ringing w.r.t. the levels $\pm\sqrt{\hat{P}}$. The filtered signal $\hat{\mathsf{x}}(t)$ can be obtained by superposition, analogous to (3.4), of the filtered transmit pulses

$$\hat{g}(t) = \frac{1}{2\pi}\int_{-2\pi W}^{2\pi W} G(\omega)e^{j\omega t}d\omega \tag{3.12}$$

where

$$G(\omega) = \int_{0}^{\infty} g(t)e^{-j\omega t}dt \tag{3.13}$$

is the Fourier transformation of the waveform $g(t)$ in (3.2). The distortion $\tilde{\mathsf{x}}(t)$ has the variance

$$\sigma_{\tilde{\mathsf{x}}}^2 = \mathbb{E}\left[\,|\hat{\mathsf{x}}(t) - \mathsf{x}(t)|^2\,\right] = \frac{1}{\pi}\int_{2\pi W}^{\infty} S_{\mathsf{x}}(\omega)d\omega \tag{3.14}$$

where $S_{\mathsf{x}}(\omega)$ is the power spectral density (PSD) of $\mathsf{x}(t)$. The resulting transmit power of the system is $P_{\tilde{\mathsf{x}}} = P - \sigma_{\tilde{\mathsf{x}}}^2$, where P is the average power of $\mathsf{x}(t)$. It is given by

$$P = \frac{\hat{P}}{T_{\text{avg}}} \left(\int_0^{\beta} \zeta^2 \left(t - \frac{\beta}{2} \right) dt + \frac{1}{\lambda} \right) = \hat{P} \frac{m + 2\kappa}{1 + 2\kappa} \tag{3.15}$$

where m captures the impact of the transition waveform by integration over $\zeta^2(t)$. At the receiver, LP-filtering ensures bandlimitation of the noise and the demapper realizes the conversion between the noisy received signal $\mathsf{r}(t)$ and the sequence $\mathbf{D}^{(M)}$ of ZC distances. The received signal after quantization and LP-filtering is given by

$$\mathsf{y}(t) = Q_1(\mathsf{r}(t)) = Q_1(\hat{\mathsf{x}}(t) + \hat{\mathsf{n}}(t)) \tag{3.16}$$

where $Q_1(\cdot)$ denotes the 1-bit quantization function, cf. (2.14). Here, $\hat{\mathsf{n}}(t)$ is the filtered additive white Gaussian noise $\mathsf{n}(t)$ with mean zero and PSD $N_0/2$. Its variance, i.e., the total in-band noise power, is

$$\sigma_{\hat{\mathsf{n}}}^2 = N_0 W \tag{3.17}$$

and its PSD is given by

$$S_{\hat{\mathsf{n}}}(f) = \begin{cases} N_0/2 & \text{for } |f| \leq W \\ 0 & \text{otherwise} \end{cases}. \tag{3.18}$$

The total distortion $\mathsf{z}(t) = \mathsf{r}(t) - \mathsf{x}(t)$ between the designed signal $\mathsf{x}(t)$ and the received signal $\mathsf{r}(t)$ results to

$$\mathsf{z}(t) = \hat{\mathsf{n}}(t) + \tilde{\mathsf{x}}(t). \tag{3.19}$$

Note that despite the deterministic nature of the filtering, it is not clear yet if and how the resulting ISI can be equalized at the receiver given that 1-bit quantization does not provide a sufficient statistic. For the purpose of lower-bounding the mutual information rate, we thus treat the ISI as noise, cf. Section 2.2.3. The implication of this approach are discussed in more detail in the corresponding sections, especially in Section 3.1.5. On the other hand, an upper bound on the achievable rate can be obtained by not considering the filter distortion.

Furthermore, we only obtain an upper and a lower bound on $\sigma_{\tilde{\mathsf{x}}}^2$, cf. Section 3.1.4, and therefore we cannot evaluate the exact transmit power $P_{\tilde{\mathsf{x}}}$. Hence, the signal-to-noise ratio (SNR) is defined w.r.t. $\mathsf{x}(t)$ as

$$\rho = \frac{P}{N_0 W}. \tag{3.20}$$

3.1.3 Error Events

When transmitting the signal $\mathsf{x}(t)$ over the channel described in Section 3.1.2, which includes LP-distortion and AWGN, three types of errors can occur:

- shifts of ZCs, which lead to magnitude errors of the received symbols
- insertion of ZC pairs, which cause insertions of received symbols
- deletion of ZC pairs, which lead to deletions of received symbols.

To the best of our knowledge, only capacity bounds for binary discrete-time channels are available for channels with insertions and deletions, e.g., [Gal61; Zig69; FDE11; DMP07].

The transition time β of the input sequence is matched to the channel bandwidth, cf. (3.9). Hence, we can assume that the filtered noise process at time instances spaced by a temporal distance larger than β is uncorrelated and, therefore, the possibility of a noise event inverting an entire symbol can be neglected. This argument leads to the conclusion that, given (3.9), deletions do not need to be included in the model. This is supported by the simulation results given in Appendix A.2.

Thus, the remaining error events are shifts and insertions of ZC. Insertions are synchronization errors, which prevent that the receiver correctly identifies where a transmit symbol begins. Information stability and Shannon's coding theorem for channels with synchronization errors has been proven by Dobrushin for discrete and finite random variables [Dob67], although to him "*it appears that these restrictions are not essential*". In the case of continuous random processes **A** and **D** this problem is still open and remains for future work.

Given a temporal separation of the two error events (shifts and insertions of ZCs), we are able to separately evaluate their impact in order to analyze the achievable rate. This separation is given as long as there are no insertions in a transition interval (TI) $[\mathsf{T}_k, \mathsf{T}_k + \beta]$ and a ZC cannot be shifted out of the TI. The dynamics of the noise within the TI are limited since the noise is bandlimited with bandwidth $2W$, which is matched to the length β of one TI, cf. (3.9). Therefore, multiple ZCs per TI occur only with very small probability in the mid-to-high SNR regime. We conduct a numerical evaluation of curve-crossing problems for Gaussian random processes, which support this argument for an SNR above 5 dB, see Appendix A.3. In this analysis, $\tilde{\mathsf{x}}(t)$ is assumed to follow a Gaussian distribution, cf. Section 3.1.5.

On the other hand, a ZC cannot be shifted out of the TI if the shift introduced by the noise is smaller than $\frac{\beta}{2}$ with high probability. We show in Appendix A.4 that this is valid for $\rho \gtrsim 10\,\text{dB}$. These assumptions affect the validity region of the bounds derived.

3.1.4 Some Signal Properties Induced by Filtering

In order to model the impact of the ISI that is caused by LP-distortion, some parameters are required to quantify the impact of the ISI w.r.t. the relevant error events mentioned above. In (2.13) the power of the LP-distortion $\tilde{\mathsf{x}}(t)$ is given. In this chapter due to the ideal LP-filter this reduces to (3.14), which requires an expression for the spectrum $S_{\mathsf{x}}(\omega)$ of $\mathsf{x}(t)$. Based on the definition of $\mathsf{x}(t)$ in (3.4), we are able to derive bounds on the spectrum of $\mathsf{x}(t)$ for $|\omega| > 0$ in Appendix A.1. Given any transition waveform $\zeta(t)$, this yields

$$S_{\mathsf{X},\mathrm{LB}}(\omega) = \frac{\hat{P}\,|G(\omega)|^2}{T_{\mathrm{avg}}(1+2c(\omega))} \leq S_{\mathsf{X}}(\omega) \leq \frac{\hat{P}\,|G(\omega)|^2}{T_{\mathrm{avg}}}(1+2c(\omega)) = S_{\mathsf{X},\mathrm{UB}}(\omega). \quad (3.21)$$

with

$$c(\omega) = \frac{1}{\sqrt{1+\omega^2\lambda^{-2}}-1}. \quad (3.22)$$

Properties of $\tilde{\mathsf{x}}(t)$ that are of interest are its variance $\sigma^2_{\tilde{\mathsf{x}}}$, cf. (3.14), and later on $s''_{\tilde{\mathsf{x}}\tilde{\mathsf{x}}}(0)$, which is the second derivative of the auto-correlation function (ACF) of $\tilde{\mathsf{x}}(t)$ at $\tau = 0$, see Sections 3.1.5 and 3.2.3. With (3.21), (3.14), and

$$\Gamma_0 = \int_{2\pi W}^{\infty} |G(\omega)|^2\, d\omega \quad (3.23)$$

we can bound $\sigma^2_{\tilde{\mathsf{x}}}$ by

$$\sigma^2_{\tilde{\mathsf{x}},\mathrm{LB}} = \frac{\hat{P}}{(1+2c_1)\pi T_{\mathrm{avg}}}\Gamma_0 \leq \sigma^2_{\tilde{\mathsf{x}}} \leq \frac{\hat{P}(1+2c_1)}{\pi T_{\mathrm{avg}}}\Gamma_0 = \sigma^2_{\tilde{\mathsf{x}},\mathrm{UB}}. \quad (3.24)$$

where c_1 is obtained by one further bounding step: $c(\omega)$ is monotonically decreasing w.r.t. $|\omega|$ and, hence, for all $|\omega| \geq 2\pi W$ it holds $c(\omega) \leq c(2\pi W) = c_1$.

The ACF of the LP-distortion $\tilde{\mathsf{x}}(t)$ is given by

$$s_{\tilde{\mathsf{x}}\tilde{\mathsf{x}}}(\tau) = \frac{1}{\pi}\int_{2\pi W}^{\infty} S_{\mathsf{X}}(\omega)\cos(\omega\tau)d\omega \quad (3.25)$$

such that its second derivative is

$$s''_{\tilde{\mathsf{x}}\tilde{\mathsf{x}}}(\tau) = \frac{1}{\pi}\int_{2\pi W}^{\infty} S_{\mathsf{X}}(\omega)\frac{\partial^2}{\partial\tau^2}\cos(\omega\tau)d\omega. \quad (3.26)$$

Here, the exchangeability of differentiation and integration has been shown via Lebesgue's dominated convergence theorem [Rud87, Theorem 1.34] where the dominating function is $f(\omega) = \omega^2 S_{\mathsf{X},\mathrm{UB}}(\omega)$. Furthermore, an upper bound on $S_{\mathsf{X}}(\omega)$ results

in a lower bound on $s''_{\tilde{\mathsf{x}}\tilde{\mathsf{x}}}(0)$ because $\frac{\partial^2}{\partial\tau^2}\cos(\omega\tau)\big|_{\tau=0} = -\omega^2$ in (3.26) and since $S_{\mathsf{X}}(\omega)$ is positive for all ω. With (3.21) and

$$\Gamma_2 = \int_{2\pi W}^{\infty} \omega^2 \left|G(\omega)\right|^2 d\omega \tag{3.27}$$

we thus can obtain the lower bound

$$s''_{\tilde{\mathsf{x}}\tilde{\mathsf{x}}}(0) \geq -\frac{(1+2c_1)\hat{P}}{\pi T_{\text{avg}}}\Gamma_2. \tag{3.28}$$

Furthermore, the description of the filtered pulse $\hat{g}(t)$ can be tedious since for $t > \beta$ the pulse $\hat{g}(t)$ exhibits the typical ringing, which can be difficult to characterize compactly. The value

$$u = (\hat{g}(\beta) - 1)\sqrt{\hat{P}} \tag{3.29}$$

represents the lowest signal level of $\hat{g}(t)$ for $t > \beta$ and thus can serve as a lower bound on $\hat{g}(t)$ in this region where required. A simplified description for the transition can be obtained by using the slope of $\hat{g}(t)$ at $t = \beta/2$, which corresponds to the slope of the filtered version $\hat{\zeta}(t)$ of $\zeta(t)$ at $t = 0$. Thus,

$$\hat{\zeta}_0 = \frac{d\hat{\zeta}(t)}{dt}\big|_{t=0} \tag{3.30}$$

and with $\hat{\zeta}(t) = \hat{g}(t + \beta/2) - 1$, $0 \leq t \leq \beta$ we can define an approximated version of $\hat{g}(t)$ as

$$\hat{g}_{\text{appr}}(t) = \begin{cases} 0, & t < 0 \\ \hat{\zeta}_0\left(t - \frac{\beta}{2}\right) + 1, & 0 \leq t \leq \beta \\ 1 + \frac{u}{\sqrt{\hat{P}}}, & t > \beta \end{cases}. \tag{3.31}$$

3.1.5 Model of the Intersymbol Interference

As mentioned above, for the purpose of lower-bounding the mutual information rate, we model the ISI as an additional noise component. For this, we approximate its distribution and obtain the corresponding parameters in this section. The original sequence $\mathsf{x}(t)$ is designed such that there is no ISI. Due to LP-filtering, $\hat{g}(t)$ shows the typical ringing such that depending on the temporal distances between the pulses – given by the data symbols A_k – interference occurs. Our approach to approximate the ISI distribution is depicted in Fig. 3.3. Starting from $\hat{g}_{\text{appr}}(t)$, we already have a characterization of the impact of filtering on the pulse starting at T_k, which we will refer to as the kth pulse. It remains to characterize the ISI generated by all neighboring pulses. Due to the separability of the error events, cf. Section 3.1.3, we divide the time interval belonging to the kth pulse in a TI and a hold period (HP).

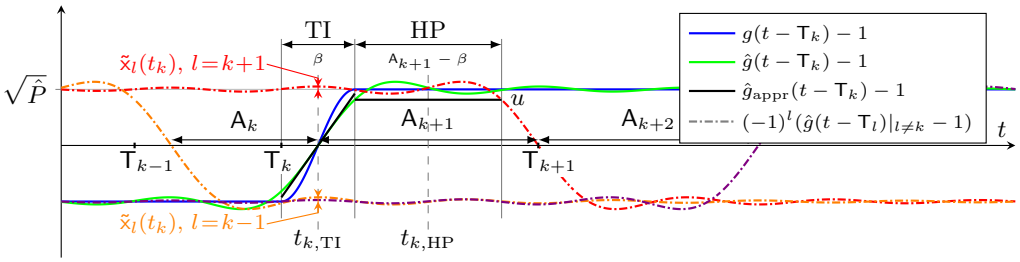

Fig. 3.3. ISI model for the kth pulse for transition interval (TI) and hold period (HP): Original pulse $g(t-\mathsf{T}_k)-1$, filtered pulse $\hat{g}(t-\mathsf{T}_k)-1$ and its approximation $\hat{g}_{\mathrm{appr}}(t)-1$, and interfering pulses $(-1)^l(\hat{g}(t-\mathsf{T}_l)|_{l\neq k}-1)$; exemplarily for a sine-shaped transition waveform

In the TI, the filtering does not change the position of the ZC, i.e., the designed waveform $g(t-\mathsf{T}_k)$ and the transmit waveform $\hat{g}(t-\mathsf{T}_k)$ both generate a ZC at $\mathsf{T}_k+\beta/2$. However, the ringing of the lth neighboring pulse $(-1)^l\hat{g}(t-\mathsf{T}_l)$, $l\neq k$ causes an interference of $\tilde{\mathsf{x}}_l(t_{k,\mathrm{TI}}=\mathsf{T}'_k)$ w.r.t. the kth pulse. In the HP, there is a symmetry w.r.t. $t_{k,\mathrm{HP}}=\mathsf{T}_k+\frac{\mathsf{A}_{k+1}+\beta}{2}$, which we exploit to approximate the distribution of the ISI. Overall, the interfering signal $\tilde{\mathsf{x}}(t_k)$ at time $t_k\in[\mathsf{T}_k,\mathsf{T}_{k+1}]$ can be represented as the sum of ISI-contributions of all other pulses as

$$\tilde{\mathsf{x}}(t_k)=\sum_{l=1}^{k-1}\tilde{\mathsf{x}}_l(t_k)+\sum_{l=k+1}^{K}\tilde{\mathsf{x}}_l(t_k)=\tilde{\mathsf{x}}_{\mathrm{lhs}}(t_k)+\tilde{\mathsf{x}}_{\mathrm{rhs}}(t_k). \tag{3.32}$$

The $\tilde{\mathsf{x}}_l(t_k)$ can be obtained via a deterministic mapping as

$$\tilde{\mathsf{x}}_l(t_k)=\begin{cases}(-1)^l\tilde{g}(\sum_{i=l+1}^{k}\mathsf{A}_i+\tilde{t}_k) & l<k\\ (-1)^{l+1}\tilde{g}(\sum_{i=k+1}^{l}\mathsf{A}_i-\tilde{t}_k+\beta) & l>k\end{cases} \tag{3.33}$$

where $\tilde{g}(t)=\hat{g}(t)-g(t)$, $\tilde{t}_k=t_k-\mathsf{T}_k$ and the sums

$$\mathsf{L}_{n+1}^{m}=\sum_{i=n+1}^{m}\mathsf{A}_i,\ m>n \tag{3.34}$$

follow the Erlang-distribution in (3.8). The two terms $\tilde{\mathsf{x}}_{\mathrm{lhs}}(t_k)$ and $\tilde{\mathsf{x}}_{\mathrm{rhs}}(t_k)$ are independent as the A_k are independent and identically distributed (i.i.d.). Therefore, we have

$$p(\tilde{\mathsf{x}}(t_k))=p(\tilde{\mathsf{x}}_{\mathrm{lhs}}(t_k))*p(\tilde{\mathsf{x}}_{\mathrm{rhs}}(t_k)). \tag{3.35}$$

Unfortunately, the analytical derivation of $p(\tilde{\mathsf{x}}(t_k))$ becomes infeasible because $\tilde{g}(t)$ cannot be inverted. Moreover, the $\tilde{\mathsf{x}}_l(t_k)$ for $l>k$ and $l<k$, respectively, are not independent. Thus, the distributions of $p(\tilde{\mathsf{x}}_{\mathrm{lhs}}(t_k))$ and $p(\tilde{\mathsf{x}}_{\mathrm{lhs}}(t_k))$ cannot be obtained by convolution of the densities $p(\tilde{\mathsf{x}}_l(t_k))$. Hence, we obtain an empirical distribution for $\tilde{\mathsf{x}}_{\mathrm{lhs}}(t_k)$ and $\tilde{\mathsf{x}}_{\mathrm{rhs}}(t_k)$ by analyzing 10^4 sequences $\mathbf{A}^{(K)}$ with 2000 interfering pulses each.

The results are depicted in Fig. 3.4. Due to the symmetry, it is sufficient to analyze the scenario of an up-crossing symbol[2] as shown in Fig. 3.3. As mentioned above, we consider two time instants of special interest in order to analyze the behavior of the ISI:

(a) The time instant $t_k = t_{k,\mathrm{TI}} = \mathsf{T}_k + \frac{\beta}{2}$ of the ZC of the kth pulse in the TI. Any distortion here affects the position of the ZC and, thus, the magnitude of the received symbol.

(b) The middle of the HP $t_k = t_{k,\mathrm{HP}} = \mathsf{T}_k + \frac{\mathsf{A}_{k+1}+\beta}{2}$, where in our model the ISI can obtain its largest values as discussed below. This is relevant for assuring synchronization in the receiver as large noise amplitudes here could insert additional ZCs.

The numerical evaluation is carried out for two exemplaric transition waveforms, namely a sine-shaped and a linear transition. It will become clear in Section 3.3 that the characteristics of the signal are mainly dominated by the bandwidth W and that the waveform plays a subordinated role. It is therefore no coincidence that the simulation results for both waveforms are almost identical.

Case (a) is depicted in Fig. 3.4a) and c). Here, we have with (3.33)

$$\tilde{\mathsf{x}}_l(t_{k,\mathrm{TI}}) = \begin{cases} (-1)^l \tilde{g}(\mathsf{L}_{l+1}^{k} + \frac{\beta}{2}) & l < k \\ (-1)^{l+1} \tilde{g}(\mathsf{L}_{k+1}^{l} + \frac{\beta}{2}) & l > k \end{cases}. \tag{3.36}$$

It can be seen from (3.36) that interfering pulses are weighted with inverted signs if they are separated by the same number of symbols ν from $t_{k,\mathrm{TI}}$, i.e., if they have the same probability distribution of L_{n+1}^{m}. Thus, the convolution of the probability density functions (pdfs) of this pair of $(\tilde{\mathsf{x}}_{l=k+\nu}(t_k), \tilde{\mathsf{x}}_{l=k-\nu}(t_k))$ becomes an auto-correlation, such that if all $\tilde{\mathsf{x}}_l(t_k)$ were independent we would expect an even function with mean zero as result of (3.35). As can be seen in Fig. 3.4a) and c) this still holds for low to moderate κ. With the bounds on the variance of the ISI obtained below, cf. (3.41) and (3.24), we observe that the distribution of the ISI can be approximated by a Gaussian distribution as a worst-case scenario up to ratios $\kappa = {}^{W}/_{\lambda} \lesssim 3$. For larger κ the simulated distribution starts to exceed the Gaussian distribution based on the upper bound on the variance of the ISI.

Since we consider that $\hat{g}(t - \mathsf{T}_k)$ is approximated by $\hat{g}_{\mathrm{appr}}(t - \mathsf{T}_k)$, it can be seen in Fig. 3.3, why case (b) describes the worst case scenario in the HP: The kth pulse $\hat{g}(t - \mathsf{T}_k)$ (green) is lower-bounded by its lowest value u, cf. (3.29). In the interval $\mathsf{T}_k + \beta < t < t_{k,\mathrm{HP}}$, the strongest interference comes from the $(k+1)$th (red) pulse. Since the envelope of $\tilde{g}(t)$ is monotonically decreasing for $t \geq \beta$, the interference

[2] As upcrossing we denote zero-crossings with positive transition slope, i.e., from $-\sqrt{\hat{P}}$ to $\sqrt{\hat{P}}$. Correspondingly, downcrossings have a negative slope.

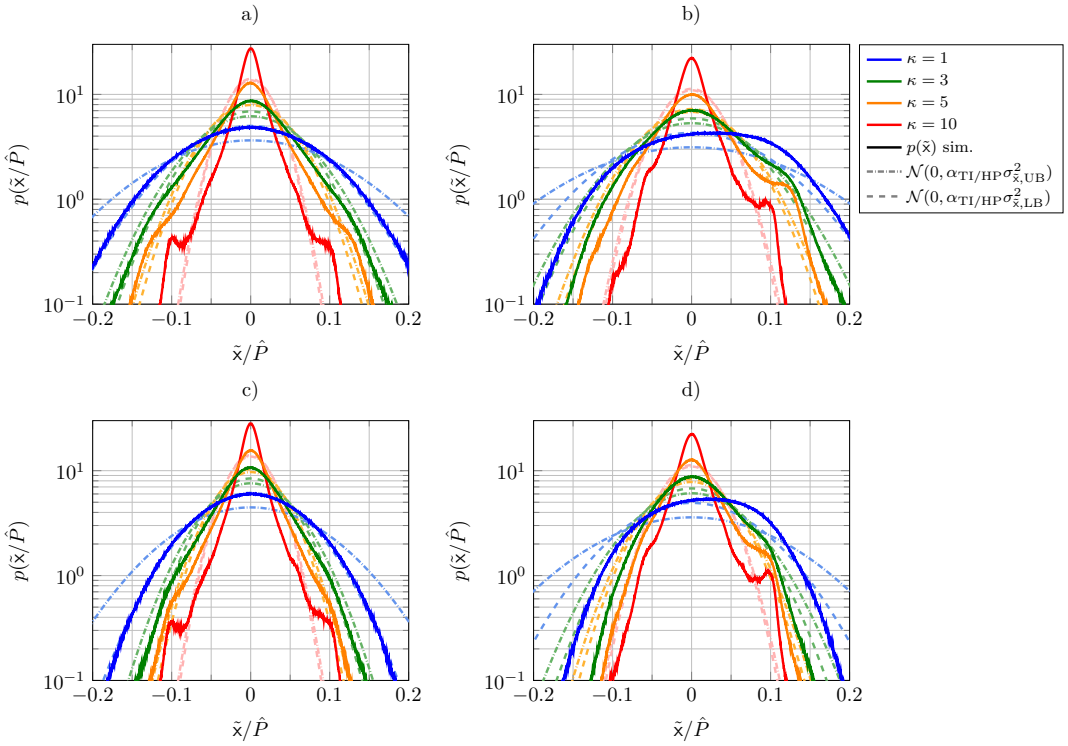

Fig. 3.4. Numerically obtained distribution of $\tilde{\mathsf{x}}(t_k)$: a) sine-shaped transition, $t_k = t_{k,\mathrm{TI}}$, b) sine-shaped transition, $t_k = t_{k,\mathrm{HP}}$, c) linear transition, $t_k = t_{k,\mathrm{TI}}$, and d) linear transition, $t_k = t_{k,\mathrm{HP}}$.

of $(k+1)$th pulse can be highest at $t_{k,\mathrm{HP}}$. For $t_{k,\mathrm{HP}} < t < \mathsf{T}_{k+1}$, the scenario can be analyzed with the red pulse approximated by u and the green one as interferer. Thus, in the middle of the HP (3.33) becomes

$$\tilde{\mathsf{x}}_l(t_{k,\mathrm{HP}}) = \begin{cases} (-1)^l \tilde{g}(\mathsf{L}^k_{l+1} + \frac{A_{k+1}}{2} + \frac{\beta}{2}) & l < k \\ (-1)^{l+1} \tilde{g}(\mathsf{L}^l_{k+2} + \frac{A_{k+1}}{2} + \frac{\beta}{2}) & l > k \end{cases}. \tag{3.37}$$

Now, differently to (3.36), the interferers with the same probability distribution of L^m_{n+1} are weighted with the same sign. For $\mathsf{A}_k + \frac{\mathsf{A}_{k+1}}{2}$ and $\frac{\mathsf{A}_{k+1}}{2} + \mathsf{A}_{k+2}$, these are the pulses $k-1$ (orange) and $k+2$ (violet). Therefore, a function with a mean deviating from zero towards positive values results from the convolution in (3.35), cf. Fig. 3.4b). This is actually beneficial, since in this region additional ZCs that can be inserted by large noise amplitudes are the relevant error event, cf. Section 3.1.3. Given that we look at an up-crossing, i.e., $\mathsf{x}(\mathsf{T}_k + \beta) > 0$, the tail of the distribution towards negative $\tilde{\mathsf{x}}$ is important. It can be seen that a Gaussian distribution with mean zero and upper-bounded variance, see below (3.41) and (3.24), has a heavier tail towards negative values compared to the actual distribution. This holds for $\kappa \lesssim 10$. Therefore, the Gaussian approximation can be used to give an upper bound on the probability of additional ZCs in that region of κ.

The variance of $\tilde{\mathrm{x}}(t)$ depends for both, TI and HP, on the amount of energy $\sigma^2_{\tilde{\mathrm{x}}}$ removed by LP-filtering. We have derived bounds on $\sigma^2_{\tilde{\mathrm{x}}}$ in Section 3.1.4, however, $\sigma^2_{\tilde{\mathrm{x}}}$ captures besides the ISI also the distortion of the current pulse, which is already included in the approximation $\hat{g}_{\mathrm{appr}}(t)$. The amount of the energy of $\tilde{g}(t)$ that contributes to ISI is the one, for which $t \geq t_{\mathrm{min}}$, where t_{min} is the minimum temporal distance between an interfering pulse and $t_{k,\mathrm{TI}}$ or $t_{k,\mathrm{HP}}$, respectively. With $\mathsf{A}_k \geq \beta$, (3.34), (3.36), and (3.37), we find that for for $l > k$ it holds for the TI that

$$t_{\mathrm{min}} = \min(\mathsf{L}^l_{k+1}) + \frac{\beta}{2} = \min(l-k)\beta + \frac{\beta}{2} = \beta + \frac{\beta}{2} \tag{3.38}$$

while for the HP

$$t_{\mathrm{min}} = \min(\mathsf{L}^l_{k+2}) + \frac{\min(\mathsf{A}_{k+1})}{2} + \frac{\beta}{2} = \beta. \tag{3.39}$$

Thus, we consider the fraction α of $\sigma^2_{\tilde{\mathrm{x}}}$ that contributes to the ISI to be

$$\alpha = \frac{\int_{t_{\mathrm{min}}}^{\infty} \tilde{g}^2(t)dt}{\int_{\frac{\beta}{2}}^{\infty} \tilde{g}^2(t)dt},\ t_{\mathrm{min}} = \begin{cases} \beta & \mathrm{HP} \\ \frac{3\beta}{2} & \mathrm{TI} \end{cases}. \tag{3.40}$$

Based on $\zeta(t)$, (3.40) can be evaluated by numerical integration for TI and HP yielding

$$\sigma^2_{\mathrm{ISI}} = \begin{cases} \alpha_{\mathrm{HP}}\sigma^2_{\tilde{\mathrm{x}}} & \text{in the HP} \\ \alpha_{\mathrm{TI}}\sigma^2_{\tilde{\mathrm{x}}} & \text{in the TI} \end{cases} \tag{3.41}$$

and

$$s''_{\mathrm{ISI}}(0) = \alpha_{\mathrm{HP}} s''_{\tilde{\mathrm{x}}\tilde{\mathrm{x}}}(0). \tag{3.42}$$

3.2 Bounding the Achievable Rate

The capacity of a communication channel is defined as the highest rate at which data can be transmitted over the channel with an arbitrary small error probability. It is given by

$$C = \sup\ I'(\mathbf{A}; \mathbf{D}) \tag{3.43}$$

where the supremum is taken over all distributions of the input signal, for which $\hat{\mathrm{x}}(t)$ is constrained to the average power $P - \sigma^2_{\tilde{\mathrm{x}}}$ and the bandwidth W. In (3.43) the mutual information rate is given by

$$I'(\mathbf{A}; \mathbf{D}) = \lim_{K\to\infty} \frac{1}{KT_{\mathrm{avg}}} I\left(\mathbf{A}^{(K)}; \mathbf{D}^{(M)}\right) \tag{3.44}$$

with $I(\mathbf{A}^{(K)};\mathbf{D}^{(M)})$ being the mutual information. Here, the mutual information rate is defined based on a normalization w.r.t. the expected transmission time KT_{avg}. In this section we derive a lower bound on the capacity by using the class of input signals introduced in Section 3.1.1. Later on, we will consider the supremum of $I'(\mathbf{A};\mathbf{D})$ over the rate parameter λ of the exponential input distribution, cf. (3.5). Additionally, we derive an upper bound on the achievable rate of this specific signaling scheme. While this bound will not serve as an upper bound on the capacity, it is useful to quantify the impact of the applied bounding steps. However, due to the data processing inequality, the AWGN capacity serves as an upper bound on the capacity of the considered communication system.

3.2.1 Bounds on the Achievable Rate

We apply the concept of a genie-aided receiver as, e.g, given in [FDE11]. The genie-aided receiver is supplied with information on the inserted ZCs via an auxiliary process $\mathbf{V}$. The criterion for the choice of $\mathbf{V}$ is that it provides sufficient information to the genie-aided receiver such that the additional ZCs can be removed. Let $\hat{\mathbf{D}}$ contain the temporal distances of the ZCs at the receiver without the additional ZCs. With a suitable choice of $\mathbf{V}$, the process $\hat{\mathbf{D}}$ can be determined based on $\mathbf{D}$ and $\mathbf{V}$ such that the mutual information rate in case the receiver has side information about the inserted ZCs is given by

$$I'(\mathbf{A};\hat{\mathbf{D}}) = I'(\mathbf{A};\mathbf{D},\mathbf{V}). \tag{3.45}$$

Using the chain rule of mutual information, we have

$$I'(\mathbf{A};\mathbf{D}) = I'(\mathbf{A};\mathbf{D},\mathbf{V}) - I'(\mathbf{A};\mathbf{V}|\mathbf{D}). \tag{3.46}$$

Here, $I'(\mathbf{A};\mathbf{D})$ is the mutual information rate without the side information on additional ZCs at the receiver. The impact of the shifted ZCs is contained in $I'(\mathbf{A};\mathbf{D},\mathbf{V})$ and the effect of the inserted ZCs is described by $I'(\mathbf{A};\mathbf{V}|\mathbf{D})$.

Given that mutual information is always non-negative, i.e., $I'(\mathbf{A};\mathbf{V}|\mathbf{D}) \geq 0$, the mutual information rate $I'(\mathbf{A};\mathbf{D})$ can be upper-bounded by

$$I'(\mathbf{A};\mathbf{D}) \leq I'(\mathbf{A};\mathbf{D},\mathbf{V}) = I'(\mathbf{A};\hat{\mathbf{D}}) \tag{3.47}$$

independently of the nature of the auxiliary process $\mathbf{V}$. In order to characterize the auxiliary process $\mathbf{V}$, we consider the transmission of a single input symbol A_k. Its

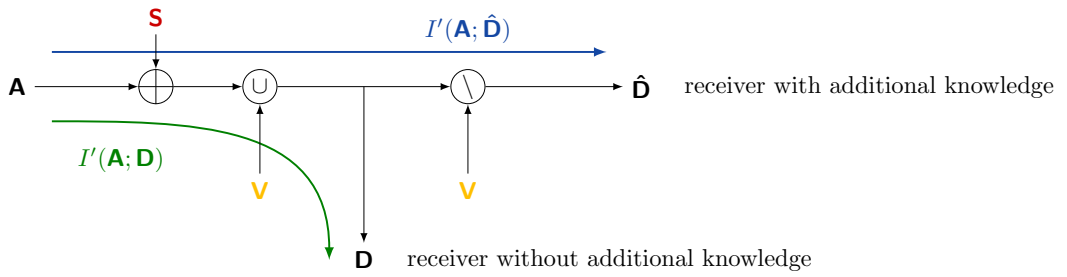

Fig. 3.5. Illustration of the error processes and the mutual information rates with and without genie-aided receiver

adjacent ZCs T'_{k-1} and T'_k will be shifted to $\hat{\mathsf{T}}_{k-1}$ and $\hat{\mathsf{T}}_k$ by the noise process, such that

$$\hat{\mathsf{T}}_k = \mathsf{T}'_k + \mathsf{S}_k \tag{3.48}$$

where S_k denotes the error introduced by the shift. On the other hand, inserted ZCs will lead to a vector of received symbols $\mathbf{D}_k$ belonging to A_k. This can be reversed, if the receiver knows which of the received ZCs correspond to the ones originally transmitted. Then the receiver can sum up the distances D_m contained in $\mathbf{D}_k$ that are separated by the additional ZCs, which yields the corresponding symbol $\hat{\mathsf{D}}_k$. Intuitively, such an algorithm would start with the first received symbol, which leads to the definition of $\mathbf{V}$ applied here: instead of providing the receiver with the exact positions in time of the additional ZCs, it is sufficient to know how many received symbols have to be summed up to obtain $\hat{\mathsf{D}}_k$. Thus, the auxiliary sequence $\mathbf{V}^{(K)}$ consists of positive integer numbers $\mathsf{V}_k \in \mathbb{N}$, which represent the number of output symbols that correspond to each input symbol A_k. Hence, we have a discrete auxiliary process $\mathbf{V}$ that can be used to lower-bound the mutual information rate in (3.46) as

$$\begin{aligned} I'(\mathbf{A};\mathbf{D}) &= I'(\mathbf{A};\mathbf{D},\mathbf{V}) - H'(\mathbf{V}|\mathbf{D}) + H'(\mathbf{V}|\mathbf{D},\mathbf{A}) && (3.49)\\ &\geq I'(\mathbf{A};\mathbf{D},\mathbf{V}) - H'(\mathbf{V}|\mathbf{D}) && (3.50)\\ &\geq I'(\mathbf{A};\mathbf{D},\mathbf{V}) - H'(\mathbf{V}). && (3.51) \end{aligned}$$

Here, the inequalities hold since the entropy rate of a discrete random process is non-negative (3.50) and due to the fact that conditioning cannot increase entropy (3.51). The error processes and mutual information rates are illustrated in Fig. 3.5. Subsequently, bounds on $I'(\mathbf{A};\mathbf{D},\mathbf{V})$ and $H'(\mathbf{V})$ will be derived.

3.2.2 Achievable Rate of the Genie-Aided Receiver

In this section we evaluate the mutual information rate of the genie-aided receiver $I'(\mathbf{A};\hat{\mathbf{D}}) = I'(\mathbf{A};\mathbf{D},\mathbf{V})$ between the sequences of temporal spacings of ZCs $\mathbf{A}^{(K)}$ and $\hat{\mathbf{D}}^{(K)}$. Since the additional ZCs have been removed by the genie-aided receiver, both vectors $\mathbf{A}^{(K)}$ and $\hat{\mathbf{D}}^{(K)}$ have the same length. The remaining errors are the shifts S_k of the zero-crossing time instants T'_k to $\hat{\mathsf{T}}_k$. Thus, on symbol level the channel output becomes with (3.48) and (3.1)

$$\hat{\mathsf{D}}_k = \hat{\mathsf{T}}_k - \hat{\mathsf{T}}_{k-1} = \mathsf{A}_k + \mathsf{S}_k - \mathsf{S}_{k-1} = \mathsf{A}_k + \Delta_k. \tag{3.52}$$

In order to derive bounds on the mutual information rate of this channel, we require knowledge on the probability distribution of S_k.

Approximation of the Distribution of the Shifting Errors

The bandlimited signal $\hat{\mathsf{x}}(t)$ can be completely described by a sampled representation with sampling rate $1/\beta$, which fulfills the Nyquist condition, cf. (3.9). Note that the concept of sampling is used here only to evaluate the overall distortion $\mathsf{z}(t)$, cf. (3.19), at the time instants T'_k of the transmitted ZCs. The receiver is still able to resolve the time instants of the ZCs with infinite resolution.

In order to obtain the distribution of the shifting error S_k, we map the pdf of the additive noise $\mathsf{z}(\mathsf{T}'_k)$ at the time T'_k into the pdf of the ZC-shift S_k on the time axis. The mapping depends on the filtered transmit pulse $\hat{g}(t)$, cf. (3.12), and is given by

$$\mathsf{z}(\mathsf{T}'_k) = -\sqrt{\hat{P}}\hat{\zeta}(\mathsf{S}_k) = \sqrt{\hat{P}}\left(1 - \hat{g}\left(\mathsf{S}_k + \frac{\beta}{2}\right)\right) \approx -\sqrt{\hat{P}}\hat{\zeta}_0\mathsf{S}_k. \tag{3.53}$$

Here we assume that $\mathsf{S}_k \ll \beta$ since we are considering the behavior in the mid-to-high SNR regime. We therefore use the linearization of $\hat{g}(t)$ around $t = \frac{\beta}{2}$ in (3.30). In Appendix A.4 we show that this holds for $\rho \gtrsim 10\,\text{dB}$.

Deriving the pdf of S_k requires the pdf of the additive noise $\mathsf{z}(\mathsf{T}'_k)$, which consists of two parts: the LP-filtered Gaussian noise $\hat{\mathsf{n}}(\mathsf{T}'_k)$ and the ISI. The latter originates from the oscillation of the neighboring pulses $\hat{g}(t)$ due to the LP-filtering, cf. Section 3.1.5. Since due to the 1-bit quantization it is not clear if and how we can equalize the ISI, which affects the position of the ZCs, we model the ISI as additional noise for lower-bounding the mutual information rate. W.r.t. an upper bound on the mutual information rate, the ISI is not considered. Then, only $\hat{\mathsf{n}}(\mathsf{T}'_k)$ contributes to $\mathsf{z}(\mathsf{T}'_k)$. We have shown in Section 3.1.5 that we can approximate the distribution of $\tilde{\mathsf{x}}(t)$ by a Gaussian distribution for ratios $\kappa = {}^{W}/_{\lambda}$ in the order of one, which will prove to be the relevant ones in the considered scenario.

We thus have

$$\mathsf{z}(t) \sim \mathcal{N}(0, \sigma_{\mathsf{z}}^2) \tag{3.54}$$

with

$$\sigma_{\mathsf{z}}^2 = \begin{cases} \sigma_{\hat{\mathsf{n}}}^2 + \sigma_{\text{ISI}}^2 & \text{for the lower bound} \\ \sigma_{\hat{\mathsf{n}}}^2 & \text{for the upper bound} \end{cases} \tag{3.55}$$

where σ_{ISI}^2 is the variance of the ISI, cf. (3.41). This yields

$$p_{\mathsf{S}}(s) = \left| -\sqrt{\hat{P}} \frac{\partial \hat{\zeta}(s)}{\partial s} p_{\mathsf{z}} \left(\sqrt{\hat{P}} \hat{\zeta}(s) \right) \right| \tag{3.56}$$

$$\approx \sqrt{\frac{\hat{P}\hat{\zeta}_0{}^2}{2\pi\sigma_{\mathsf{z}}^2}} \exp\left\{ -\frac{\hat{P}\hat{\zeta}_0{}^2}{2\sigma_{\mathsf{z}}^2} s^2 \right\}. \tag{3.57}$$

Thus, the ZC errors S_k are approximately Gaussian distributed in the mid-to-high SNR domain for κ in the order of one, i.e., $\mathsf{S}_k \sim \mathcal{N}\left(0, \sigma_{\mathsf{S}}^2\right)$ with

$$\sigma_{\mathsf{S}}^2 = \frac{\sigma_{\mathsf{z}}^2}{\hat{P}\hat{\zeta}_0{}^2}. \tag{3.58}$$

Upper Bound on the Achievable Rate of the Genie-Aided Receiver

With $\mathsf{S}_k \sim \mathcal{N}\left(0, \sigma_{\mathsf{S}}^2\right)$ and (3.55), it results for $\Delta_k = \mathsf{S}_k - \mathsf{S}_{k-1}$ in (3.52)

$$\Delta_k \sim \mathcal{N}(0, 2\sigma_{\mathsf{S}}^2) \tag{3.59}$$

since the S_k are approximately independent as the minimum distance of two ZCs is β, which in turn is matched to the bandwidth, cf. (3.9). However, the Δ_k are correlated since they always depend on the current and the previous S_k. Thus, the ACF of $\boldsymbol{\Delta}$ is

$$\phi_{\Delta\Delta}(m) = \mathbb{E}\left[\Delta_l \Delta_{l+m}\right] = \begin{cases} 2\sigma_{\mathsf{S}}^2, & m = 0 \\ -\sigma_{\mathsf{S}}^2, & |m| = 1 \\ 0, & \text{otherwise} \end{cases} \tag{3.60}$$

such that the covariance matrix $\mathbf{R}_{\Delta}^{(K)}$ of ZC shifts $\boldsymbol{\Delta}^{(K)} = [\Delta_1, ..., \Delta_K]^T$ becomes

$$\mathbf{R}_{\Delta}^{(K)} = \mathbb{E}\left[(\boldsymbol{\Delta}^{(K)})^T \boldsymbol{\Delta}^{(K)} \right] = \sigma_{\mathsf{S}}^2 \begin{pmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & -1 & 2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -1 \\ 0 & \dots & 0 & -1 & 2 \end{pmatrix}. \tag{3.61}$$

This shows that the channel with the genie-aided receiver is a colored additive Gaussian noise channel with input $\mathbf{A}$, output $\hat{\mathbf{D}}$, and noise $\boldsymbol{\Delta}$. Here, $\boldsymbol{\Delta}$ is independent of $\mathbf{A}$ since the ISI is not considered, cf. (3.55). It is known that the capacity of the colored additive Gaussian noise channel is achieved by a Gaussian input distribution [CT06, Chapter 9, Eq. (9.97)], which gives us an upper bound on the mutual information rate of the channel with the genie-aided receiver and the exponential input distribution. Thus, we have

$$I'(\mathbf{A};\hat{\mathbf{D}}) \leq \frac{1}{2}\int_{-\frac{1}{2}}^{\frac{1}{2}} \log\left(1+\frac{(\nu - S_{\Delta}(f))^{+}}{S_{\Delta}(f)}\right) df \tag{3.62}$$

where the water level ν is given by

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} (\nu - S_{\Delta}(f))^{+} df = \sigma_{\mathsf{A}}^2 \tag{3.63}$$

with σ_{A}^2 given in (3.7). Furthermore, $S_{\Delta}(f)$ denotes the PSD of $\boldsymbol{\Delta}$, which is given by the z-transform of (3.60) as

$$S_{\Delta}(f) = 2\sigma_{\mathsf{S}}^2(1-\cos(2\pi f)), \qquad |f| < 0.5. \tag{3.64}$$

Although $S_{\Delta}(f)$ is equal to zero for $f = 0$, the integral in (3.62) exists. This can be shown by using that $\nu \geq (\nu - S_{\Delta}(f))^{+}\ \forall f$ and solving

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} \log\left(1+\frac{{}^{\nu}/_{(2\sigma_{\mathsf{S}}^2)}}{1-\cos(2\pi f)}\right) df = \operatorname{arcosh}\left(\frac{\nu}{2\sigma_{\mathsf{S}}^2}+1\right). \tag{3.65}$$

Lower Bound on the Achievable Rate of the Genie-Aided Receiver

The mutual information between $\mathbf{A}^{(K)}$ and $\hat{\mathbf{D}}^{(K)}$ of the genie-aided receiver is

$$I(\mathbf{A}^{(K)};\hat{\mathbf{D}}^{(K)}) = h(\mathbf{A}^{(K)}) - h(\mathbf{A}^{(K)}|\hat{\mathbf{D}}^{(K)}) \tag{3.66}$$

$$= h(\mathbf{A}^{(K)}) - h(\mathbf{A}^{(K)} - \hat{\mathbf{A}}_{\text{LMMSE}}^{(K)}|\hat{\mathbf{D}}^{(K)}) \tag{3.67}$$

where $h(\cdot)$ is the differential entropy. Furthermore, $\hat{\mathbf{A}}_{\text{LMMSE}}^{(K)}$ denotes the estimate of $\mathbf{A}^{(K)}$ based on $\hat{\mathbf{D}}^{(K)}$ with a linear minimum mean-squared error (LMMSE) estimator. Equality in (3.67) holds due to the fact that the addition of a constant does not change the differential entropy. Since $\hat{\mathbf{A}}_{\text{LMMSE}}^{(K)}$ is a deterministic function of $\hat{\mathbf{D}}^{(K)}$, it can be treated as a constant while conditioning on $\hat{\mathbf{D}}^{(K)}$.

As a next step, we upper-bound the term $h(\mathbf{A}^{(K)} - \hat{\mathbf{A}}_{\text{LMMSE}}^{(K)}|\hat{\mathbf{D}}^{(K)})$ on the right-hand side of (3.67). It represents the randomness of the estimation error of the LMMSE estimator when estimating $\mathbf{A}^{(K)}$ based on $\hat{\mathbf{D}}^{(K)}$. An upper bound on this error can be given by the differential entropy of a Gaussian random variable with the same

covariance matrix [CT06, Theorem 8.6.5]. Then, with (3.52), the estimation error covariance matrix of the LMMSE estimator becomes

$$\begin{aligned}\mathbf{Q}_{\text{err}}^{(K)} &= \mathbb{E}\left[(\mathbf{A}^{(K)} - \hat{\mathbf{A}}_{\text{LMMSE}}^{(K)})(\mathbf{A}^{(K)} - \hat{\mathbf{A}}_{\text{LMMSE}}^{(K)})^T\right] \\ &= \mathbf{Q}_{\mathsf{A}} - (\mathbf{Q}_{\mathsf{A}} + \mathbf{Q}_{\mathsf{A}\Delta})(\mathbf{Q}_{\mathsf{A}} + \mathbf{Q}_{\Delta} + \mathbf{Q}_{\mathsf{A}\Delta} + \mathbf{Q}_{\mathsf{A}\Delta}^T)^{-1}(\mathbf{Q}_{\mathsf{A}} + \mathbf{Q}_{\mathsf{A}\Delta})^T. \end{aligned} \tag{3.68}$$

Here, all covariance matrices $\mathbf{Q}$ have size $K \times K$ and $\mathbf{Q}_{\mathsf{A}} = \sigma_{\mathsf{A}}^2 \mathbf{I}^{(K)}$. Moreover, $\mathbf{I}^{(K)}$ is the identity matrix of dimension $K \times K$, and σ_{A}^2 is given in (3.7). The cross-covariance matrix of $\mathbf{A}^{(K)}$ and $\mathbf{\Delta}^{(K)}$ is $\mathbf{Q}_{\mathsf{A}\Delta} = \mathbb{E}[\mathbf{A}^{(K)}(\mathbf{\Delta}^{(K)})^T]$ and the covariance matrix of $\mathbf{\Delta}^{(K)}$ is $\mathbf{Q}_{\Delta} = \mathbb{E}[\mathbf{\Delta}^{(K)}(\mathbf{\Delta}^{(K)})^T]$.

If we ignore the correlation between $\mathbf{A}^{(K)}$ and $\mathbf{\Delta}^{(K)}$, this corresponds to $\mathbf{Q}_{\mathsf{A}\Delta} = \mathbf{0}$. Furthermore, neglecting the correlation between the ISI samples $\tilde{\mathsf{x}}(\mathsf{T}'_k)$, $k = 1, ..., K$ is equivalent to $\mathbf{Q}_{\Delta} = \mathbf{R}_{\Delta}^{(K)}$, with $\mathbf{R}_{\Delta}^{(K)}$ in (3.61). It is shown in Appendix A.5 that these assumptions, i.e., $\mathbf{Q}_{\mathsf{A}\Delta} = \mathbf{0}$ and $\mathbf{Q}_{\Delta} = \mathbf{R}_{\Delta}^{(K)}$, result in an upper bound on $h(\mathbf{A}^{(K)} - \hat{\mathbf{A}}_{\text{LMMSE}}^{(K)} | \hat{\mathbf{D}}^{(K)})$ yielding

$$h(\mathbf{A}^{(K)} - \hat{\mathbf{A}}_{\text{LMMSE}}^{(K)} | \mathbf{D}^{(K)}) \leq \frac{1}{2} \log \det \left(2\pi e \mathbf{Q}_{\text{err}}^{(K)}\right) \tag{3.69}$$

$$\leq \frac{1}{2} \log \det \left(2\pi e \mathbf{R}_{\text{err}}^{(K)}\right) \tag{3.70}$$

with

$$\mathbf{R}_{\text{err}}^{(K)} = \sigma_{\mathsf{A}}^2 \mathbf{I}^{(K)} - \sigma_{\mathsf{A}}^4 (\sigma_{\mathsf{A}}^2 \mathbf{I}^{(K)} + \mathbf{R}_{\Delta}^{(K)})^{-1}. \tag{3.71}$$

Thus, the lower bound on the mutual information in (3.67) becomes

$$\begin{aligned} I(\mathbf{A}^{(K)}; \hat{\mathbf{D}}^{(K)}) &\geq h(\mathbf{A}^{(K)}) - \frac{1}{2} \log \det \left(2\pi e \mathbf{R}_{\text{err}}^{(K)}\right) \\ &= K h(\mathsf{A}_k) + \frac{1}{2} \log \det \left((2\pi e)^{-1} \left(\sigma_{\mathsf{A}}^{-2} \mathbf{I}^{(K)} + (\mathbf{R}_{\Delta}^{(K)})^{-1}\right)\right). \end{aligned} \tag{3.72}$$

Here, $h(\mathbf{A}^{(K)}) = K h(\mathsf{A}_k)$ follows from the independence of the input symbols A_k. For the second term of (3.72) we have used (3.71) and the matrix inversion lemma. With (3.72) the lower bound on the mutual information rate in (3.45) is given by

$$I'(\mathbf{A}; \hat{\mathbf{D}}) \geq \lim_{K \to \infty} \frac{1}{K T_{\text{avg}}} \left\{ K h(\mathsf{A}_k) + \frac{1}{2} \log \det \left[\frac{1}{2\pi e} \left(\sigma_{\mathsf{A}}^{-2} \mathbf{I}^{(K)} + (\mathbf{R}_{\Delta}^{(K)})^{-1}\right)\right] \right\} \tag{3.73}$$

$$= \frac{1}{T_{\text{avg}}} \left\{ h(\mathsf{A}_k) + \frac{1}{2} \int_{-\frac{1}{2}}^{\frac{1}{2}} \log \left(\frac{\sigma_{\mathsf{A}}^{-2}}{2\pi e} \left(1 + \frac{\sigma_{\mathsf{A}}^2}{S_{\Delta}(f)}\right)\right) df \right\} \tag{3.74}$$

where we have used Szegö's theorem on the asymptotic eigenvalue distribution of Hermitian Toeplitz matrices [Gra06], [GS58, pp. 64-65] for (3.74). Again, $S_{\Delta}(f)$ is

the PSD of $\mathbf{\Delta}$, cf. (3.64), which corresponds to the sequence of covariance matrices $\mathbf{R}_{\Delta}^{(K)}$. As shown in (3.65), the integral exists. The entropy at the channel input is

$$h(\mathsf{A}_k) = 1 - \log(\lambda) \tag{3.75}$$

since the A_k are exponentially distributed. If we now insert the expressions for T_{avg}, σ_{A}^2, β, and $h(\mathsf{A}_k)$, i.e., (3.6), (3.7), (3.9), and (3.75), we can write with (3.65) for the lower bound in (3.74)

$$I'(\mathbf{A}; \hat{\mathbf{D}}) \geq \frac{W\left[\log\left(\frac{e}{2\pi}\right) + \operatorname{arcosh}\left(\frac{1}{2\sigma_{\mathsf{S}}^2\lambda^2} + 1\right)\right]}{1 + 2W\lambda^{-1}}. \tag{3.76}$$

3.2.3 Bounds on the Entropy of the Auxiliary Process

In Section 3.2.1 we defined the elements V_k of the auxiliary process $\mathbf{V}$ as random variables that describe the number of received symbols corresponding to the transmitted symbols A_k. Based on this, we now have to find an upper bound for $H'(\mathbf{V})$, cf. (3.51). The V_k depend on the number N_k of inserted ZCs as $\mathsf{V}_k = \mathsf{N}_k + 1$. Since we assume separability of shift and insertion errors, we do not have to consider the TIs, which contain only the shifted ZCs. Thus, it remains the HP of duration $T_{\mathrm{sat}} = \mathbb{E}[\mathsf{A}_k] - \beta = \lambda^{-1}$ where $\mathsf{x}(t)$ maintains the level $\pm\sqrt{\hat{P}}$. Without the LP-filters, this would yield a level-crossing problem w.r.t. the level $\pm\sqrt{\hat{P}}$. However, for the actual transmit signal $\hat{\mathsf{x}}(t)$ we have a curve crossing problem due to the typical ringing, cf. Fig. 3.2. Thus we apply a further bounding step such that we can obtain a closed form expression for an upper bound on N_k: We consider a level crossing problem w.r.t. the lowest value u of the kth pulse in the HP. Here, u is given by (3.29).

Level-crossing problems have been widely studied, especially for Gaussian processes, for example in [Kac43; Ric44; CL67]. Our upper bound on $H'(\mathbf{V})$ is based on the first moment of the distribution of V_k. It is known that for a stationary zero-mean Gaussian random process with variance σ_{z}^2, the expected number of crossings of the level u in the time interval $T_{\mathrm{sat}} = \lambda^{-1}$ is given by the Rice formula [Ric44]

$$\mu = \mathbb{E}[\mathsf{V}_k] = \frac{1}{\pi}\sqrt{\frac{-s_{\mathsf{zz}}''(0)}{\sigma_{\mathsf{z}}^2}}\exp\left(-\frac{u^2}{2\sigma_{\mathsf{z}}^2}\right)\frac{1}{\lambda} + 1. \tag{3.77}$$

In (3.77) $s_{\mathsf{zz}}(\tau)$ is the ACF of the Gaussian process $\mathsf{z}(t)$ and $s_{\mathsf{zz}}''(\tau) = \frac{\partial^2}{\partial\tau^2}s_{\mathsf{zz}}(\tau)$ is its second derivative. This yields analogously to (3.55)

$$s_{\mathsf{zz}}''(0) = s_{\hat{\mathsf{n}}\hat{\mathsf{n}}}''(0) + s_{\mathrm{ISI}}''(0) = -\frac{4}{3}N_0W^3 + s_{\mathrm{ISI}}''(0) \tag{3.78}$$

where the first term on the right-hand side stems from the AWGN. Moreover, $s_{\mathrm{ISI}}(\tau)$ is the ACF of the ISI. In order for $\mathbb{E}[\mathsf{V}_k]$ to be finite, it is required that $-s''_{\mathsf{zz}}(0) < \infty$ [CL67], which holds for finite bandwidths W since then $s''_{\hat{\mathsf{n}}\hat{\mathsf{n}}}(0)$ and $s''_{\mathrm{ISI}}(0)$ are finite.

The distribution that maximizes the entropy of a discrete random variable on $\mathbb{N}$ for a given mean μ in (3.77) is the geometric distribution, cf. [Kap93, Section 2.1]. Therefore, we upper-bound the entropy of $H(\mathsf{V}_k)$ by

$$H(\mathsf{V}_k) \leq (1-\mu)\log(\mu-1) + \mu\log\mu \tag{3.79}$$

which yields for the entropy rate of the auxiliary process

$$H'(\mathbf{V}) \leq \frac{H(\mathsf{V}_k)}{T_{\mathrm{avg}}} \leq \frac{(1-\mu)\log(\mu-1)+\mu\log\mu}{T_{\mathrm{avg}}} \tag{3.80}$$

as independent V_k maximize the entropy rate. It results that the bound on $H(\mathsf{V}_k)$ is an increasing function in the expected number of level-crossings μ of the Gaussian random process. Here, μ increases with σ_{z}^2. Therefore, an upper bound on σ_{z}^2 and, thus, on σ_{ISI}^2 is required to evaluate (3.77). An upper bound on σ_{ISI}^2 results in a lower bound on $s''_{\mathrm{ISI}}(0)$, cf. Section 3.1.4 and Section 3.1.5, as both parameters depend on the ACF of the distortion process $\tilde{\mathsf{x}}(t)$ and cannot be chosen independently. This in turn yields an upper bound on $-s''_{\mathrm{ISI}}(0)$ and, thus, $-s''_{\mathsf{zz}}(0)$.

3.3 Impact of the Transition Waveform

The choice of transition waveform affects the model in several ways. In principle, a steeper transition increases the noise resistance since the overall distortion $\mathsf{z}(t)$ then translates to a smaller distortion in time. However, a steeper transition is as well subject to a stronger distortion due to filtering, i.e., increasing $\mathsf{z}(t)$ in the first place such that the advantage might be canceled out completely. In order to shed light on this question, we analyze three different transition waveforms

$$\zeta_{\mathrm{sine}}(t) = \begin{cases} \sin\left(\pi\frac{t}{\beta}\right) & \text{for } |t| \leq \beta/2 \\ 0 & \text{otherwise} \end{cases} \tag{3.81}$$

$$\zeta_{\mathrm{lin}}(t) = \begin{cases} \frac{2}{\beta}t & \text{for } |t| \leq \beta/2 \\ 0 & \text{otherwise} \end{cases} \tag{3.82}$$

$$\zeta_{\mathrm{poly},2B+1}(t) = \begin{cases} \sum_{i=0}^{B} b_{2i+1}\left(\frac{t}{\beta}\right)^{2i+1} & \text{for } |t| \leq \beta/2 \\ 0 & \text{otherwise.} \end{cases} \tag{3.83}$$

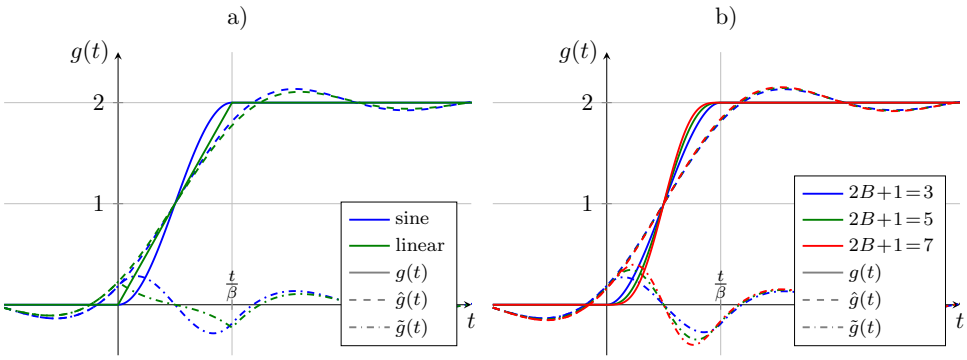

Fig. 3.6. Transition waveforms a) linear and sine, b) polynomial of order $2B+1$

Tab. 3.1. Coefficients for polynomial transitions

$2B+1$	b_7	b_5	b_3	b_1
3	0	0	-4	3
5	0	12	-10	$\frac{15}{4}$
7	-40	42	$-\frac{35}{2}$	$\frac{35}{8}$

In the case of polynomial transitions (polys), polynomials of degree 3 to 7, i.e., $B \in \mathbb{N}, 1 \leq B \leq 3$, are considered. Recall that $\zeta(t)$ is required to be odd such that only odd exponents occur. For smoothness of the overall waveform, the coefficients b_i are chosen such that the first to $(2B+1)$th derivatives at $t = \pm\frac{\beta}{2}$ are zero. This yields the coefficients listed in Table 3.1. All waveforms are depicted in Fig. 3.6. It can already be seen that despite the different characteristics of the original waveforms $g(t)$, the $\hat{g}(t)$ vary only slightly. This is confirmed by the similar slopes $\hat{\zeta}_0$ in Table 3.2, which summarize relevant parameters and properties introduced in Section 3.1 for the different waveforms. Consequently, the parameter Γ_0 capturing the impact of the waveform on the distortion increases for the steeper waveforms. This can be seen as well in Fig. 3.7, where the signal based on the sine-shaped waveform has more out of band energy. This is strong evidence for the assumption that given the structure of the model applied, increasing the slope will not improve the bounds on the achievable rate. We will see this confirmed in Section 3.4.

Note that except of the α in (3.40), all expressions can be given in closed form, which however is omitted in Table 3.2 for the sake of clear presentation. The closed-form expressions for the parameters of all waveforms can be found in Appendix A.6. The numerical integration for obtaining the α has been carried out using Mathematica, which gives us increased accuracy compared to [BDFa], where we used MATLAB. This has a slight impact on optimal value of κ, which is evaluated in the next section. Section 3.4 furthermore gives the results w.r.t. the bounds on the mutual information rate for the different waveforms based on the findings of the previous sections.

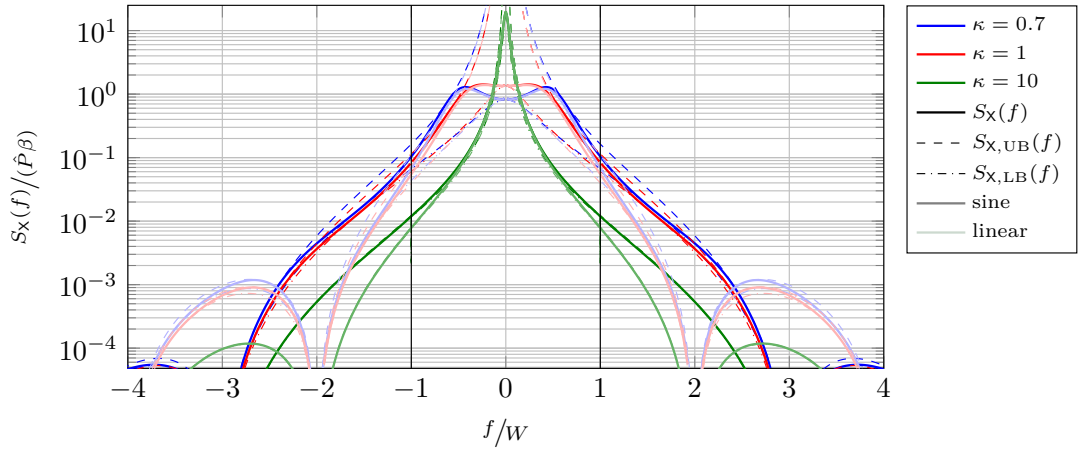

Fig. 3.7. Numerical computation of the spectrum of $\mathsf{x}(t)$ according to (A.8) for $K = 10^4$ and the bounds from (3.21) for linear and sine-shaped transition waveforms

Tab. 3.2. Model parameters of the waveforms, $m = 2W \int_0^\beta \zeta^2\left(t - \frac{\beta}{2}\right) dt$, cf. (3.15)

waveform		$\frac{P}{\hat{P}}$	m	Γ_0	Γ_2	$\frac{u}{\sqrt{\hat{P}}}$	$\hat{\zeta}_0$	α_{TI}	α_{HP}
sine		$\frac{m+2\kappa}{1+2\kappa}$	$\frac{1}{2}$	0.2601β	$\frac{1.1689}{\beta}$	0.81	$\frac{1.8519}{\beta}$	0.3160	0.4283
linear			$\frac{1}{3}$	0.1224β	$\frac{2.8438}{\beta}$	0.77	$\frac{1.7453}{\beta}$	0.4474	0.6924
$2B+1$	3		$\frac{17}{35}$	0.2446β	$\frac{4.3395}{\beta}$	0.8121	$\frac{1.8439}{\beta}$	0.3267	0.4447
	5		$\frac{131}{231}$	0.3399β	$\frac{6.7321}{\beta}$	0.8289	$\frac{1.8872}{\beta}$	0.2726	0.3649
	7		$\frac{797}{1287}$	0.4123β	$\frac{9.0052}{\beta}$	0.8384	$\frac{1.9116}{\beta}$	0.2432	0.3246

3.4 Results and Summary

This section summarizes the results of the previous sections, which leads to the final form of the obtained lower and upper bounds on the mutual information rate. It furthermore discusses the optimal value for the normalized bandwidth κ as well as the spectral efficiency of the system and provides the conclusion of the chapter.

3.4.1 Lower and Upper Bound on the Achievable Rate

In order to obtain the final form of the lower bound on the mutual information rate of the bandlimited 1-bit quantized continuous-time channel, we have to use

$$I'(\mathbf{A};\mathbf{D}) \geq I'(\mathbf{A};\hat{\mathbf{D}}) - H'(\mathbf{V}) \tag{3.84}$$

which is obtained from (3.51) and (3.45). Here, we have to substitute the results of Sections 3.2.2 and 3.2.3 for $I'(\mathbf{A};\hat{\mathbf{D}})$ and $H'(\mathbf{V})$, i.e., (3.76) and (3.80). With the definitions for T_{avg}, β, and κ, i.e., (3.6), (3.9), and (3.11), this yields

$$I'(\mathbf{A};\mathbf{D}) \geq \frac{W}{2\kappa+1}\left[\operatorname{arcosh}\left(\frac{\hat{P}\hat{\zeta}_0{}^2}{2\sigma^2_{\mathsf{z},\mathrm{UB}}\lambda^2}+1\right)+\log\left(\frac{e}{2\pi}\right)\right.$$
$$\left.+2\mu_{\mathrm{UB}}\log\left(\frac{\mu_{\mathrm{UB}}-1}{\mu_{\mathrm{UB}}}\right)-2\log(\mu_{\mathrm{UB}}-1)\right] = I'_{\mathrm{LB}}(\mathbf{A};\mathbf{D}). \tag{3.85}$$

The bound holds for small ratios $\kappa = {}^{W}\!/\!_{\lambda}$ in the order of one due to the limitations of the Gaussian approximation of the LP-distortion. With the indices $(\cdot)_{\mathrm{LB}}$ and $(\cdot)_{\mathrm{UB}}$ we refer to lower and upper bounds on the indexed variable, respectively. The upper bound on the mutual information rate for the given signaling scheme is valid for all κ. It is given by, cf. (3.62),

$$I'(\mathbf{A};\mathbf{D}) \leq \frac{W}{2\kappa+1}\int_{-\frac{1}{2}}^{\frac{1}{2}}\log\left(1+\frac{(\nu-S_{\Delta,\mathrm{LB}}(f))^+}{S_{\Delta,\mathrm{LB}}(f)}\right)df = I'_{\mathrm{UB}}(\mathbf{A};\mathbf{D}). \tag{3.86}$$

Both bounds are valid for $\rho \gtrsim 10\,\mathrm{dB}$, i.e., when $|\mathsf{S}_k| < {}^{\beta}\!/\!_{2}$ with high probability. Thus, the temporal separation of error events (ZC-shifts and additional ZCs) holds.

In (3.85), with (3.15), (3.20), (3.24), (3.6), (3.9), and (3.11) we can express $\sigma^2_{\mathsf{z},\mathrm{UB}}$ as function of $\hat{P}$, ρ, and κ by

$$\sigma^2_{\mathsf{z},\mathrm{UB}} = N_0 W + \alpha_{\mathrm{TI/HP}}\sigma^2_{\mathsf{x},\mathrm{UB}} = \hat{P}\left(\frac{m+2\kappa}{(1+2\kappa)\rho}+\frac{\alpha_{\mathrm{TI/HP}}(1+2c_1)c_0}{2\pi^2(1+2\kappa)}\right) \tag{3.87}$$

Note that c_1 is a function of κ. The same applies to $\frac{-s''_{\mathsf{zz},\mathrm{LB}}(0)}{\lambda^2}$ in (3.77) with (3.78), (3.42), and (3.28) yielding

$$\frac{-s''_{\mathsf{zz},\mathrm{LB}}(0)}{\lambda^2} = \hat{P}\kappa^2\left(\frac{4}{3}\frac{m+2\kappa}{(1+2\kappa)\rho}+\frac{\alpha_{\mathrm{HP}}2(1+2c_1)c_2}{1+2\kappa}\right). \tag{3.88}$$

Furthermore, for a given waveform $\hat{\zeta}_0 \sim \frac{1}{\beta}$, cf. Appendix A.6, such that ${}^{\hat{\zeta}_0{}^2}\!/\!_{\lambda^2} \sim 4\kappa^2$ and $u^2 \sim \hat{P}$. Hence, both μ_{UB} in (3.77) and the normalized lower bound $I'_{\mathrm{LB}}(\mathbf{A};\mathbf{D})/W$ depend solely on κ and ρ. By expanding the fraction in (3.86) with σ_{A}^{-2}, we have $\sigma^2_{\mathsf{S},\mathrm{LB}}\sigma_{\mathsf{A}}^{-2} = \frac{\sigma^2_{\mathsf{z},\mathrm{LB}}}{\hat{P}}\frac{\lambda^2}{\hat{\zeta}_0{}^2}$, cf. (3.64), which shows that the same behavior applies to $I'_{\mathrm{UB}}(\mathbf{A};\mathbf{D})/W$.

An upper bound on the mutual information rate with 1-bit quantization is given by the capacity $C_{\mathrm{AWGN}} = W\log(1+\rho)$ of the AWGN channel without output quantization. The ratio between C_{AWGN} and $I'_{\mathrm{LB}}(\mathbf{A};\mathbf{D})$ is given by

$$\Delta I = \frac{\log(1+\rho)}{\frac{1}{W}I'_{\mathrm{LB}}(\mathbf{A};\mathbf{D})} = f(\kappa,\rho). \tag{3.89}$$

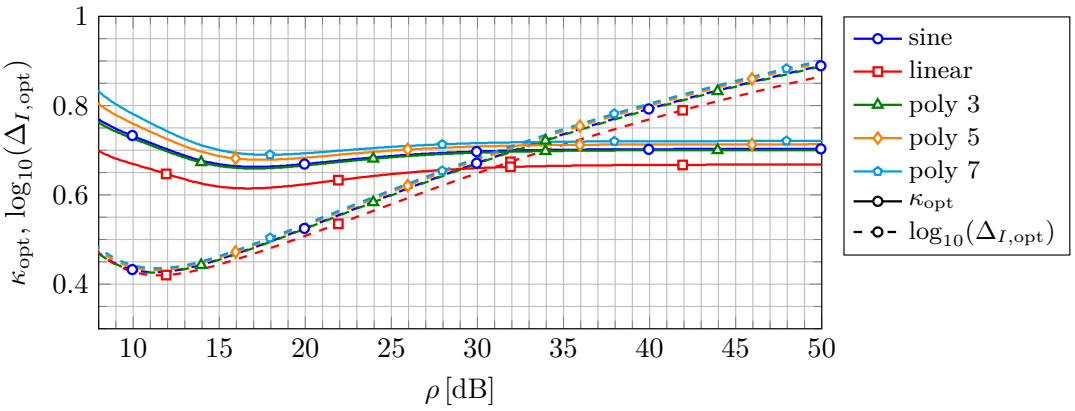

Fig. 3.8. Optimal ratio $\kappa = W\lambda^{-1}$ over the SNR and corresponding ratio $\Delta I = {}^{C_{\mathrm{AWGN}}}/_{I'_{\mathrm{LB}}(\mathbf{A};\mathbf{D})}$, valid for $\rho \gtrsim 10\,\mathrm{dB}$

which is a function of κ and ρ. Thus, for a given ρ we can find the κ minimizing ΔI. The results are shown in Fig. 3.8. We observe that the optimal κ is in the order of one, which means that the randomness at the channel input needs to be matched to the bandwidth of the channel. This can be achieved if λ grows linearly with W. The optimal κ is approximately 0.66 to 0.7 for the steeper transition waveforms (sine and polynomial) and slightly lower with approximately 0.62 to 0.67 for the linear waveform. This values diverge slightly from the ones given in [BDFa] due to the increased accuracy of the numerical integration, cf. Section 3.3, however, the impact on the lower bound on the mutual information rate is negligible. The ratio ΔI is smallest for approximately 12 dB SNR.

Note that these results depend heavily on the connection of the transition time β and the signal bandwidth W given in (3.9). By reducing the bandwidth, for example to $W = {}^{1}/_{(2T_{\mathrm{avg}})}$, the utilization of the spectrum could be improved. However, the minimum symbol duration then not longer corresponds to the coherence time of the noise, based on which we neglect symbol deletions. Thus, in order to model transmission scenarios with $W < {}^{1}/_{(2\beta)}$, deletions have to be included, see Appendix A.2.

3.4.2 Spectral Efficiency Results

The resulting bounds over the SNR ρ in terms of spectral efficiency are given in Fig. 3.9. Due to the normalization with the bandwidth $2W$ they depend solely on κ. Three bounds are compared: $I'_{\mathrm{LB}}(\mathbf{A};\mathbf{D})$ from (3.85), $I'_{\mathrm{UB}}(\mathbf{A};\mathbf{D})$ from (3.86), and the lower bound without ISI, $I'_{\mathrm{LB,noISI}}(\mathbf{A};\mathbf{D})$, which results for $\alpha_{\mathrm{TI}} = \alpha_{\mathrm{HP}} = 0$, cf. (3.41). The latter two bounds hold for any κ, while $I'_{\mathrm{LB}}(\mathbf{A};\mathbf{D})$ is only valid for values of κ in the order of one, cf. Section 3.1.5. However, this restriction is not critical as with κ in the order of one, we obtain the highest mutual information rates and, therefore, the best lower bounds. In Fig. 3.9a) exemplarily the results for a linear transition are depicted.

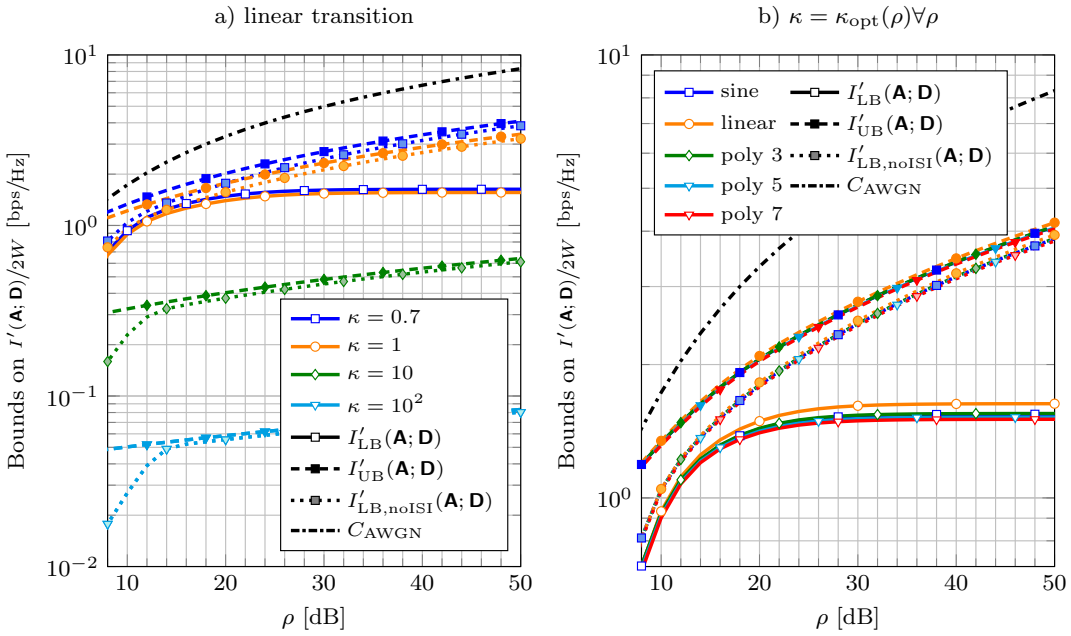

Fig. 3.9. Lower and upper bounds on the spectral efficiency $I'(\mathbf{A};\mathbf{D})/(2W)$ valid for the mid-to-high SNR regime $\rho \gtrsim 10\,\mathrm{dB}$, a) linear transition and b) all waveforms with the optimal κ for every SNR ρ

waveform	sine	linear	poly 3	poly 5	poly 7
spectral efficiency limit [bit/s/Hz]	1.545	1.632	1.551	1.521	1.505

Tab. 3.3. High SNR limits for the spectral efficiency

We observe that below ca. 16 dB SNR, $I'_{\mathrm{LB}}(\mathbf{A};\mathbf{D})$ and $I'_{\mathrm{UB}}(\mathbf{A};\mathbf{D})$ approach each other with increasing SNR. This illustrates the decreasing impact of $H'(\mathbf{V})$ on $I'_{\mathrm{LB}}(\mathbf{A};\mathbf{D})$ as $I'_{\mathrm{UB}}(\mathbf{A};\mathbf{D})$ does not include additional ZCs and their probability decreases with the SNR. In the high SNR domain, on the other hand, $I'_{\mathrm{LB}}(\mathbf{A};\mathbf{D})$ and $I'_{\mathrm{UB}}(\mathbf{A};\mathbf{D})$ diverge again. This is the case when the system becomes dominated by the ISI, which is considered only in $I'_{\mathrm{LB}}(\mathbf{A};\mathbf{D})$. Interestingly, $I'_{\mathrm{LB,noISI}}(\mathbf{A};\mathbf{D})$ does not saturate over the SNR, which coincides with the observation in [AV96] that the achievable rate of the timing channel tends to infinity under noise free conditions ($\rho \to \infty$). Since $I'_{\mathrm{LB,noISI}}(\mathbf{A};\mathbf{D})$ and $I'_{\mathrm{UB}}(\mathbf{A};\mathbf{D})$ are considerably close especially in the high-SNR domain, we conclude that the bounding steps applied, mostly in Section 3.2.2, do not impact the behavior of the bound prohibitively.

Fig. 3.9a) displays the obtained bounds when for every SNR the optimal κ is used. It can be seen that in any case, although the differences are marginal, the linear transition waveforms yields the best results, even for $I'_{\mathrm{UB}}(\mathbf{A};\mathbf{D})$ and $I'_{\mathrm{LB,noISI}}(\mathbf{A};\mathbf{D})$. This is due to the reduced average power P and, for $I'_{\mathrm{LB}}(\mathbf{A};\mathbf{D})$, additionally due to an reduced σ^2_{ISI}. The approximate limits of the lower bound $I'_{\mathrm{LB}}(\mathbf{A};\mathbf{D})$ for high SNR are given in Tab. 3.3.

3.4.3 Conclusions

In this chapter we derived a lower bound on the mutual information rate of the bandlimited 1-bit quantized continuous-time AWGN channel. The focus was on the mid-to-high SNR regime and the lower bound is valid for SNR values above approximately 10 dB and $\kappa = {}^{W}/_{\lambda} \lesssim 3$, in which case the filter distortion can be approximated by a Gaussian distribution. Moreover, for the specific signaling scheme applied, we provided an upper bound on the mutual information rate. We conclude that the bounds provide a valuable characterization of the actual mutual information rate with the given signaling scheme on 1-bit quantized channels in an SNR range between approximately 10 and 20 dB, where they are relatively close. The bounds hold under the following assumptions:

- In order to obtain a closed form expression of the lower bound, the LP-distortion error $\tilde{\mathsf{x}}(t)$ is approximated to be Gaussian. This is suitable for $\kappa \lesssim 3$, for which we obtain the best lower bounds, cf. Section 3.1.5.
- If $W = \frac{1}{2\beta}$, the occurrence of deletions is negligible in the mid-to-high SNR scenario and for the considered input signals, cf. Appendix A.2.
- Per transition interval $[\mathsf{T}_k, \mathsf{T}_k + \beta]$ with high probability only one zero-crossing occurs. This follows from the limited dynamics of the noise due to bandlimitation. It has been verified for an SNR above 5 dB by numerical computation based on curve-crossing problems for Gaussian random processes, cf. Appendix A.3.
- When deriving the upper bound, the individual elements of the process $\mathbf{S}$ are considered to be i.i.d. since the minimum temporal separation of the individual S_k is β, which is matched to the bandwidth of the noise, cf. (3.9). Furthermore, the main contributor to correlation is ISI, which is neglected for upper-bounding.
- For the mid-to-high SNR-domain we have $\mathsf{S}_k \ll \beta$. Thus, the transition can be linearized around the ZC, cf. (3.53). Appendix A.4 shows that this holds for $\rho \gtrsim 10\,\text{dB}$.

In order to maximize the lower bound on the mutual information rate for a given bandwidth, the rate parameter λ of the exponential distribution of the A_k has to grow linearly with the bandwidth W. The optimal coefficient κ depends on the SNR and the waveforms. For the considered system it lies between 0.6 and 0.7 for high SNR. Furthermore, the system model has to incorporate deletions if the filter bandwidth W is smaller than ${}^{1}/_{2\beta}$ as otherwise the spectral efficiency of the system is overestimated. Due to the LP-distortion, the derived lower bound on the mutual information rate saturates for high SNR since the designed signal $\mathsf{x}(t)$ is not strictly bandlimited. This is not the case if the ISI due to filtering is neglected. Thus, the accuracy of the bound could be improved with a more elaborated ISI model, where ISI is not treated as noise.

4 Discrete-Time Channels and Multiple-Input Multiple-Output Channels

This chapter aims to illustrate applications of the previously obtained results to discrete-time scenarios. Due to the digital nature of today's communications systems, it is highly relevant to see which insights carry over and which may not.

4.1 Discrete-Time Runlength Limited Signals

In [BSK+17], it has been shown how the bounding techniques of Section 3.2 can be applied to a discrete-time system. However, spectral efficiencies are only given for non-bandlimited systems w.r.t. the 90 %-power containment bandwidth. Since it has been shown in Section 3.4 that the way of modeling of the intersymbol interference (ISI) plays a central role in this regard, in this section a framework is derived in order to obtain bounds on the spectral efficiency without complex simulations. This includes strictly bandlimited channels and allows also the evaluation of faster-than-Nyquist (FTN) signaling schemes as introduced in Section 2.2.1. As input signaling scheme runlength limited (RLL) sequences of Section 2.2.2 are used, however, the framework can be applied to any 1-bit quantized binary signaling scheme. The bounds on the achievable rate in this chapter are as in Chapter 3 based on an auxiliary process providing genie-aided information to the receiver. The upper bound on the entropy of the auxiliary process is constructed by modeling the ISI as noise as described in Section 2.2.3. The resulting bounds have the advantage of evaluation with low computational complexity for any channel impulse response. The results of this chapter for selected choices of the channel impulse response have been submitted for publication to IEEE Communication Letters [BDFb].

4.1.1 System Model

The system model is depicted in Fig. 4.1a). Again, the elements of the channel input and output vectors $\mathbf{A}^{(K)} = [\mathsf{A}_1, ..., \mathsf{A}_K]^T$ and $\mathbf{D}^{(M)} = [\mathsf{D}_1, ..., \mathsf{D}_M]^T$ represent the temporal distances of two consecutive zero-crossings (ZCs) of the signal $\mathsf{x}(t)$ and the received signal $\mathsf{r}(t)$, respectively. However, in this case due to sampling the

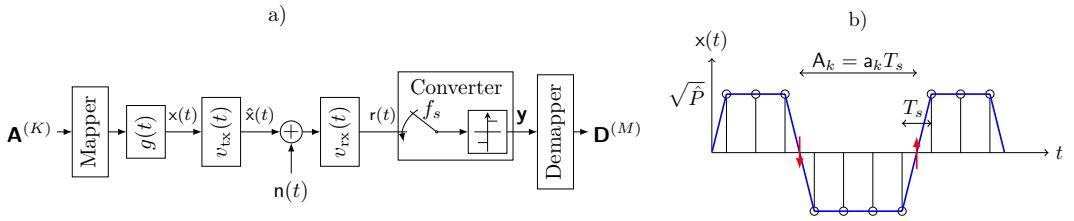

Fig. 4.1. Discrete-time runlength encoded system: a) system model, b) mapping between A_k and $\mathsf{x}(t)$ for a triangular pulse $g(t)$ of width $2T_s$

resolution of A_k and D_m is finite and limited to multiples of the sampling interval T_s. Nevertheless, again K is not necessarily equal to M since insertions and deletions can occur. Furthermore, the A_k are independent and identically distributed (i.i.d.). The channel input symbols A_k are converted into a_k consecutive Dirac-pulses, which yields a sequence $\mathsf{x}_0^{N_{\text{samp}}-1}$, where $\mathsf{x}_n \in \{-1, 1\}$ is the weight of the Dirac-pulse and N_{samp} denotes the total number of samples. In order to obtain $\mathsf{x}(t)$, this sequence is then convolved with the impulse response $g(t)$ of a transmit pulse former yielding

$$\mathsf{x}(t) = \sqrt{\hat{P}} \sum_{n=0}^{N_{\text{samp}}-1} \mathsf{x}_n g(t - nT_s). \tag{4.1}$$

Fig. 4.1b) shows this mapping for a triangular pulse $g(t)$ of duration $2T_s$.

The number of consecutive alike samples a_k will be limited to a minimum runlength $a_{\min}$, which yields an RLL sequence as described in Section 2.2.2 without a constraint on the maximum runlength. Based on the fact that for max-entropic RLL sequences the runlength a_k is geometrically distributed [Imm90], the entropy of an input symbol is given by

$$H(\mathsf{A}_k) = H(\mathsf{a}_k) = (1 + \tilde{\mu}) \log(1 + \tilde{\mu}) - \tilde{\mu} \log(\tilde{\mu}) \tag{4.2}$$

where $\tilde{\mu} = \mathbb{E}[\mathsf{a}_k] - a_{\min}$. Since the input symbols A_k are i.i.d., the entropy rate of the input process becomes

$$H'(\mathbf{A}) = \lim_{K \to \infty} \frac{H(\mathsf{A}_1, \mathsf{A}_2, ..., \mathsf{A}_K)}{K\, \mathbb{E}[\mathsf{A}_k]} = \frac{H(\mathsf{A}_k)}{\mathbb{E}[\mathsf{a}_k] T_s} = \frac{C_{\text{rll}}(a_{\min})}{T_s} \tag{4.3}$$

where C_{rll} is the capacity in bits per channel use (bpcu) of the RLL sequence [Imm90]. It is given in Table 2.2 for some $a_{\min} = d_{\text{rll}} + 1$. The spectrum of the signal $\mathsf{x}(t)$ is

$$S_{\mathsf{x}}(\omega) = \frac{|G(\omega)|^2}{T_s} S_{\text{rll}}(\omega) \tag{4.4}$$

where $S_{\text{rll}}(\omega)$ is the spectrum of the RLL sequence given in (2.10) and $G(\omega)$ is the Fourier transform of $g(t)$.

If not mentioned otherwise, we assume ideal lowpasses (LPes) with one-sided bandwidth W and amplitude one as channel filters. With the combined impulse response of transmit and receive filter

$$h_{\mathrm{LP}}(t) = (v_{\mathrm{tx}} * v_{\mathrm{rx}})(t) \tag{4.5}$$

the overall channel impulse $h(t)$ response is

$$h(t) = (g * h_{\mathrm{LP}})(t). \tag{4.6}$$

The receiver consists of a data converter, which samples the received signal with sampling rate $f_s = \frac{1}{T_s}$, and a 1-bit quantizer. The received signal $\mathsf{r}(t)$ becomes

$$\mathsf{r}(t) = (v_{\mathrm{rx}} * \hat{\mathsf{x}})(t) + \hat{\mathsf{n}}(t) \tag{4.7}$$

where $\hat{\mathsf{x}}(t)$ is the transmit signal. In the case of the ideal LPes, $\hat{\mathsf{x}}(t)$ is the noise-free receive signal as well, i.e., $\hat{\mathsf{x}}(t) = (v_{\mathrm{tx}} * \mathsf{x})(t) = (h_{\mathrm{LP}} * \mathsf{x})(t)$. Moreover, $\hat{\mathsf{n}}(t)$ is the filtered version of the additive white Gaussian noise (AWGN) $\mathsf{n}(t)$ with mean zero and power spectral density (PSD) $N_0/2$. After LP filtering the noise power at the receiver becomes

$$\sigma^2_{\hat{\mathsf{n}}} = \frac{1}{2\pi}\frac{N_0}{2}\int_{-\infty}^{\infty} |V_{\mathrm{rx}}(\omega)|^2 d\omega \tag{4.8}$$

where $V_{\mathrm{rx}}(\omega)$ is the Fourier transform of $v_{\mathrm{rx}}(t)$. For ideal LPes, we have $\sigma^2_{\hat{\mathsf{n}}} = N_0 W$. The amount of the LP filter distortion

$$\tilde{\mathsf{x}}(t) = \hat{\mathsf{x}}(t) - \mathsf{x}(t) \tag{4.9}$$

depends on $h_{\mathrm{LP}}(t)$ as

$$\sigma^2_{\tilde{\mathsf{x}}} = \mathbb{E}\left[\,|\hat{\mathsf{x}}(t) - \mathsf{x}(t)|^2\,\right] = \frac{1}{2\pi}\int_{-\infty}^{\infty} (|H_{\mathrm{LP}}(\omega)|^2 - 1) S_{\mathsf{x}}(\omega) d\omega \tag{4.10}$$

$$= \frac{1}{\pi}\int_{2\pi W}^{\infty} S_{\mathsf{X}}(\omega) d\omega. \tag{4.11}$$

where $H_{\mathrm{LP}}(\omega)$ is the Fourier transform of $h_{\mathrm{LP}}(t)$ and (4.11) holds only for the ideal LP. For the transmit power P_{Tx} of $\hat{\mathsf{x}}(t)$ we have

$$P_{\mathrm{Tx}} = \mathbb{E}\left[\,|\hat{\mathsf{x}}(t)|^2\,\right] = \frac{1}{2\pi}\int_{-\infty}^{\infty} |V_{\mathrm{tx}}(\omega)|^2 S_{\mathsf{X}}(\omega) d\omega \tag{4.12}$$

where $V_{\mathrm{tx}}(\omega)$ is the Fourier transform of $v_{\mathrm{tx}}(t)$. For the special case of the ideal LPes this results to $P_{\mathrm{Tx}} = P - \sigma^2_{\tilde{\mathsf{x}}}$ where the average power $P = \mathbb{E}\left[\,|\mathsf{x}(t)|^2\,\right]$ depends on $\hat{P}$

and the waveform $g(t)$. In this case we have $P_{\mathrm{Rx}} = P_{\mathrm{Tx}}$ for the received power, while in general holds

$$P_{\mathrm{Rx}} = \mathbb{E}\left[\,|\hat{\mathsf{x}}(t)|^2\,\right] = \frac{1}{2\pi}\int_{-\infty}^{\infty} |H_{\mathrm{LP}}(\omega)|^2 S_{\mathsf{X}}(\omega) d\omega. \tag{4.13}$$

Thus, we have for the signal-to-noise ratio (SNR)

$$\mathrm{SNR} = \frac{P_{\mathrm{Rx}}}{\sigma^2_{\hat{\mathsf{n}}}}. \tag{4.14}$$

The received signal $\mathsf{r}(t)$ is sampled with sampling frequency f_s and quantized by a 1-bit quantizer. For the sampled received vector we have

$$\begin{aligned}\mathbf{y} &= [\mathsf{y}_0, \mathsf{y}_1, ..., \mathsf{y}_n, ..., \mathsf{y}_{N_{\mathrm{samp}}-1}] && (4.15)\\ &= [Q_1(\mathsf{r}(0)), Q_1(\mathsf{r}(T_s)), ..., , Q_1(\mathsf{r}(nT_s)), ..., Q_1(\mathsf{r}((N_{\mathrm{samp}}-1)T_s))]. && (4.16)\end{aligned}$$

The sampled received sequence $\mathbf{y}$ is then mapped to a sequence $\mathsf{D}_m = \mathsf{d}_m T_s$ of ZC distances. Due to the spectral shaping properties of RLL sequences, cf. Section 2.2.2, oversampling is considered here w.r.t. to the signal bandwidth. The oversampling ratio (OSR) is

$$M_{\mathrm{osr}} = M_{\mathrm{osr},f} = \frac{f_s}{2W}. \tag{4.17}$$

The possible error events are again shift, insertion and deletions, such that with a suitable choice on an auxiliary process $\mathbf{V}$, the bound in (3.51) can be applied to lower-bound the mutual information rate.

4.1.2 Bound on the Achievable Rate and Auxiliary Process

As described above, while the different input symbols A_k are independent, the samples of the received signal are not. The number a_k of consecutive alike samples encodes the transmitted information. Therefore, a flipped sample can change the number of received symbols and – as in Chapter 3 – we are dealing with a insertion- and deletion-channel. The capacity for insertion- and deletion-channels is in general unknown. In [FDE11] an approach is introduced how the achievable rate can be lower-bounded by using an auxiliary process. It is therefore required that we are able to define an appropriate discrete auxiliary process $\mathbf{V}$ to provide genie-aided information about insertions and deletions to the receiver. We then can use (3.51) for lower-bounding the achievable rate, i.e.,

$$I'(\mathbf{A};\mathbf{D}) \geq I'(\mathbf{A};\mathbf{D},\mathbf{V}) - H'(\mathbf{V}). \tag{4.18}$$

Note that in contrast to the continuous-time case, the coding theorem exists and has been proven by Dobrushin [Dob67]. In the simplest case, $\mathbf{V}$ is a binary sequence (± 1) that is multiplied with the received vector. It indicates which samples are to be flipped (-1) and which not ($+1$). In that case, the auxiliary process allows for perfect reconstruction of the transmitted sequence and we have

$$I'(\mathbf{A}; \mathbf{D}, \mathbf{V}) = I'(\mathbf{A}; \mathbf{A}) = H'(\mathbf{A}). \tag{4.19}$$

The entropy rate of the auxiliary process can then be bounded based on the average sample flipping probability p_b. It results

$$H'(\mathbf{V}) \leq \frac{1}{N_{\text{samp}} T_s} \sum_{n=0}^{N_{\text{samp}}-1} H(\mathsf{V}_n) = \frac{1}{N_{\text{samp}} T_s} \sum_{n=0}^{N_{\text{samp}}-1} H_b(p_{b,n}) \tag{4.20}$$

$$\leq \frac{1}{T_s} H_b\left(\frac{1}{N_{\text{samp}}} \sum_{n=0}^{N_{\text{samp}}-1} p_{b,n} \right) = \frac{1}{T_s} H_b(p_b) \tag{4.21}$$

where $H_b(\cdot)$ is the binary entropy function and $p_{b,n}$ denotes the flipping probability of the nth sample. The first inequality holds since independence maximizes the entropy (4.20) while the second is based on Jensen's inequality as $H_b(\cdot)$ is a concave function (4.21). Contributions to p_b result from the filtered noise as well as the ISI. Despite being deterministic when conditioned on $\mathsf{x}(t)$, $\tilde{\mathsf{x}}(t)$ is still a random process. With (4.9), the received signal at time nT_s can be written as

$$\mathsf{r}(nT_s) = \mathsf{x}(nT_s) + \tilde{\mathsf{x}}(nT_s) + \hat{\mathsf{n}}(nT_s) = \mathsf{x}_n + \tilde{\mathsf{x}}_n + \hat{\mathsf{n}}_n. \tag{4.22}$$

where

$$\tilde{\mathsf{x}}_n = (h_0 - 1)\mathsf{x}_n + \sum_{\substack{i=0 \\ i \neq n}}^{N_{\text{samp}}-1} h((i-n)T_s)\mathsf{x}_i = (h_0 - 1)\mathsf{x}_n + \mathsf{z}_{\text{ISI},n}. \tag{4.23}$$

In order to bound the average sample flipping probability p_b, we condition on the current sample x_n. Then $\tilde{\mathsf{x}}_n$ consists of a deterministic and a random contribution. The deterministic contribution is $(h_0 - 1)\mathsf{x}_n$, where $h_0 = h(0)$ is the maximum value of $h(t)$, yielding a deterministic reduction of the power per sample. The random contribution, $\mathsf{z}_{\text{ISI},n}$, represents the ISI, which depends on the realization of the neighboring symbols. In order to upper-bound $H'(\mathbf{V})$ in (4.21), we assumed independent samples. Thus, we model the ISI as additional noise, which allows us to obtain bounds the achievable rate that can be evaluated with low computational complexity. However, the sequence

information remains captured in $I'(\mathbf{A};\mathbf{D},\mathbf{V})$, cf. (4.18). The flipping probability p_b of the nth sample is

$$\begin{aligned} p_b &= \Pr(\mathsf{y}_n \neq \mathsf{x}_n) && (4.24)\\ &= \Pr(\mathsf{y}_n = 1|\mathsf{x}_n = -1)\Pr(\mathsf{x}_n = -1) + \Pr(\mathsf{y}_n = -1|\mathsf{x}_n = 1)\Pr(\mathsf{x}_n = 1) && (4.25)\\ &= \frac{1}{2}(\Pr(\mathsf{z}_{\mathrm{ISI},n} + \hat{\mathsf{n}}_n > h_0|\mathsf{x}_n = -1) + \Pr(\mathsf{z}_{\mathrm{ISI},n} + \hat{\mathsf{n}}_n < -h_0|\mathsf{x}_n = 1)) && (4.26)\\ &= \Pr(\mathsf{z}_{\mathrm{ISI},n} + \hat{\mathsf{n}}_n < -h_0|\mathsf{x}_n = 1)) && (4.27) \end{aligned}$$

where we have used that RLL sequences are zero-mean, i.e., $\Pr(\mathsf{x}_n = 1) = \Pr(\mathsf{x}_n = -1) = \frac{1}{2}$. Furthermore, due to the symmetry of the setup $\Pr(\mathsf{z}_{\mathrm{ISI},n} + \hat{\mathsf{n}}_n > h_0|\mathsf{x}_n = -1) = \Pr(\mathsf{z}_{\mathrm{ISI},n} + \hat{\mathsf{n}}_n < -h_0|\mathsf{x}_n = 1)$. An upper bound on the expression in (4.26) can be given by

$$p_b \leq \frac{1}{2}\Pr(|\mathsf{z}_{\mathrm{ISI},n} + \hat{\mathsf{n}}_n| > h_0) = \tilde{p_b}. \tag{4.28}$$

Bounds for (4.28) are given, e.g., in [Sal68; Lug69] for i.i.d. samples, in [Gla72] and, more completely, in [Mat73] for correlated samples. The basic idea is to find the distribution for $\mathsf{z}_{\mathrm{ISI},n}$ that maximizes (4.28) based on Chebyshev's inequality; knowing that

$$\tilde{p_b} = \int\limits_{|z_n|>h_0} p_{\mathsf{z}_{\mathrm{ISI}}}(z_n) * p_{\hat{\mathsf{n}}}(z_n)dz_n = \int\limits_{-Z}^{Z} p_{\mathsf{z}_{\mathrm{ISI}}}(z_n)\Omega(z_n)dz_n \tag{4.29}$$

with

$$\Omega(z_n) = \frac{1}{2}\left(\operatorname{erfc}\left(\frac{h_0 - z_n}{\sqrt{2}\sigma_{\hat{\mathsf{n}}}}\right) + \operatorname{erfc}\left(\frac{h_0 + z_n}{\sqrt{2}\sigma_{\hat{\mathsf{n}}}}\right)\right) \tag{4.30}$$

due to the AWGN channel and under the conditions that the ISI is zero-mean and

$$\mathsf{z}_{\mathrm{ISI},n} \in [-Z, Z] \qquad \int\limits_{-Z}^{Z} p_{\mathsf{z}_{\mathrm{ISI}}}(z_n)dz_n = 1 \quad \int\limits_{-Z}^{Z} z_n^2 p_{\mathsf{z}_{\mathrm{ISI}}}(z_n)dz_n = \sigma^2_{\mathsf{z}_{\mathrm{ISI}}}. \tag{4.31}$$

Here, $p_{\mathsf{z}_{\mathrm{ISI}}}(z_n)$ denotes the probability density function of $\mathsf{z}_{\mathrm{ISI},n}$, where the time index n in the subscript is omitted for notational convenience. The resulting bounds are

$$p_b \leq \frac{1}{2}\begin{cases} \Omega(\sigma_{\mathsf{z}_{\mathrm{ISI}}}), & h_0 \leq \sqrt{3}\sigma_{\hat{\mathsf{n}}} \\ \Omega(\sigma_{\mathsf{z}_{\mathrm{ISI}}}), & h_0 > \sqrt{3}\sigma_{\hat{\mathsf{n}}}, z_0 < \sigma_{\mathsf{z}_{\mathrm{ISI}}} \\ \left(1 - \frac{\sigma^2_{\mathsf{z}_{\mathrm{ISI}}}}{z_0^2}\right)\Omega(0) + \frac{\sigma^2_{\mathsf{z}_{\mathrm{ISI}}}}{z_0^2}\Omega(z_0), & h_0 > \sqrt{3}\sigma_{\hat{\mathsf{n}}}, \sigma_{\mathsf{z}_{\mathrm{ISI}}} \leq z_0 \leq Z \\ \left(1 - \frac{\sigma^2_{\mathsf{z}_{\mathrm{ISI}}}}{Z^2}\right)\Omega(0) + \frac{\sigma^2_{\mathsf{z}_{\mathrm{ISI}}}}{Z^2}\Omega(Z), & h_0 > \sqrt{3}\sigma_{\hat{\mathsf{n}}}, z_0 > Z \end{cases} \tag{4.32}$$

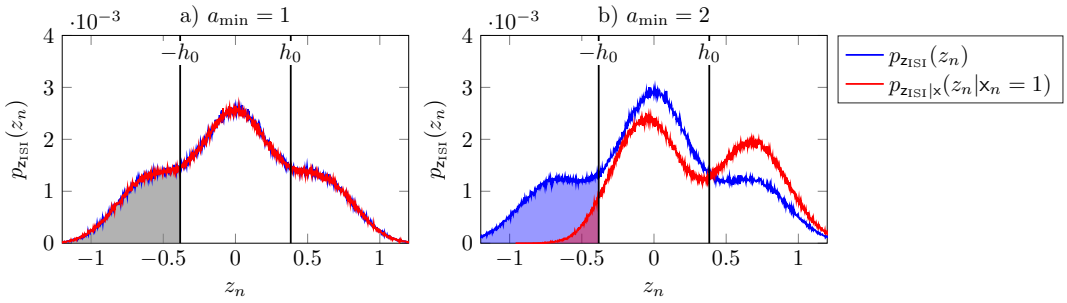

Fig. 4.2. Histogram of $p_{\mathsf{z}_{\mathrm{ISI}}}(z_n)$ for a strong ISI scenario $2WT_s = 0.4$ and minimum runlength $a_{\min} = \{1, 2\}$, $h_0 = 0.3830$

where

$$z_0 = \arg\max \frac{\Omega(z_n) - \Omega(0)}{z_n^2}. \tag{4.33}$$

For details on the derivation the reader is referred to [Mat73]. While holding, these bounds are very loose in our case due to the strong constructive ISI, which is occurring since we use RLL sequences. This is illustrated in Fig. 4.2 for a strong ISI scenario. Assume $\sigma_{\hat{\mathsf{n}}}^2 \to 0$, such that $p_{\mathsf{z}_{\mathrm{ISI}}+\hat{\mathsf{n}}}(z_n) \approx p_{\mathsf{z}_{\mathrm{ISI}}}(z_n)$. We observe that for i.i.d. samples, i.e., $a_{\min} = 1$, (4.27) and (4.28) are equivalent (shaded in gray) since both distributions, the marginal distribution $p_{\mathsf{z}_{\mathrm{ISI}}}(z_n)$ and the conditional distribution $p_{\mathsf{z}_{\mathrm{ISI}}|\mathsf{x}}(z_n|\mathsf{x}_n = 1)$ coincide and are zero-mean. On the other hand, for $a_{\min} = 2$ bounding (4.28) yields an upper bound on the blue area while we are looking for an upper bound on the red area. We see that when $a_{\min} > 1$, the conditional distribution $p_{\mathsf{z}_{\mathrm{ISI}}|\mathsf{x}}(z_n|\mathsf{x}_n = 1)$ is not longer symmetric nor zero-mean, which remains valid for the marginal distribution. In order to reduce the overestimation of p_b, we thus attempt to bound (4.27) directly. We have

$$p_b = \int_{Z_1}^{Z_2} p_{\mathsf{z}_{\mathrm{ISI}}|\mathsf{x}}(z_n|\mathsf{x}_n = 1)\frac{1}{2}\operatorname{erfc}\left(\frac{h_0 + z_n}{\sqrt{2}\sigma_{\hat{\mathsf{n}}}}\right) dz_n \tag{4.34}$$

assuming that that

$$\begin{aligned} &\mathsf{z}_{\mathrm{ISI}} \in [Z_1, Z_2], && \int_{Z_1}^{Z_2} z_n^2 p_{\mathsf{z}_{\mathrm{ISI}}|\mathsf{x}}(z_n|\mathsf{x}_n = 1)dz_n = \mathbb{E}[z_{\mathrm{ISI},n}^2|\mathsf{x}_n = 1] \\ &\int_{Z_1}^{Z_2} p_{\mathsf{z}_{\mathrm{ISI}}|\mathsf{x}}(z_n|\mathsf{x}_n = 1)dz_n = 1, && \int_{Z_1}^{Z_2} z_n p_{\mathsf{z}_{\mathrm{ISI}}|\mathsf{x}}(z_n|\mathsf{x}_n = 1)dz_n = \mathbb{E}[\mathsf{z}_{\mathrm{ISI},n}|\mathsf{x}_n = 1]. \end{aligned} \tag{4.35}$$

Note that the conditional ISI distribution $p_{\mathsf{z}_{\mathrm{ISI}}|\mathsf{x}}(z_n|\mathsf{x}_n = 1)$ is unknown. Furthermore, it is neither symmetric nor zero-mean, which is why additional parameters are needed in order to describe the support and the mean of the ISI. A simple method to obtain

estimates on Z_1, Z_2, $\mathbb{E}[\mathsf{z}_{\mathrm{ISI},n}|\mathsf{x}_n = 1]$, and $\mathbb{E}[z_{\mathrm{ISI},n}^2|\mathsf{x}_n = 1]$ is a low-complexity Monte-Carlo simulation by simply generating a long RLL sequence, which is filtered by $h(t)$. Furthermore, a method for obtaining Z and $\sigma_{\mathsf{z}_{\mathrm{ISI}}}^2$ is given in [Big74] for finite impulse response (FIR) filters and in [Gla72]

$$\mathbb{E}[z_{\mathrm{ISI},n}^2] = \frac{1}{2\pi T_s} \int_{-\frac{\pi}{T_s}}^{\frac{\pi}{T_s}} S_{\mathrm{rll}}(\omega) |\tilde{H}(\omega) - h_0 T_s|^2 d\omega \tag{4.36}$$

$$\tilde{H}(\omega) = \begin{cases} \sum_{i=-\infty}^{\infty} H\left(\omega + \frac{2\pi i}{T_s}\right), & |\omega| \leq \frac{\pi}{T_s} \\ 0, & |\omega| > \frac{\pi}{T_s} \end{cases} \tag{4.37}$$

is given, when the spectrum of the data sequence is known and the ISI. Hereby $H(\omega)$ is the Fourier transform of the channel impulse response $h(t)$.

Since $\mathrm{erfc}(x)$ is monotonically decreasing in x, we deduce from Fig. 4.2 that in order to obtain an upper bound on (4.34) as much weight as possible has to be put onto z_n close to Z_1 while the conditions in (4.35) have to hold. Therefore, we upper-bound p_b based on a worst-case $p_{\mathsf{z}_{\mathrm{ISI}}|\mathsf{x}}(z_n|\mathsf{x}_n = 1)$ with two mass points at z_1 and z_2 of probability p_{z_1} and p_{z_2}. This yields for the upper bound on the average sample flipping probability

$$p_b \leq \frac{1}{2} \left(p_{z_1}^* \, \mathrm{erfc}\left(\frac{h_0 + z_1^*}{\sqrt{2}\sigma_{\hat{\mathsf{n}}}} \right) + (1 - p_{z_1}^*) \, \mathrm{erfc}\left(\frac{h_0 + z_2^*}{\sqrt{2}\sigma_{\hat{\mathsf{n}}}} \right) \right) \tag{4.38}$$

where z_1^*, z_2^*, and $p_{z_1}^*$ denote the values that maximize the right-hand side of (4.38). The maximization in (4.38) can be carried out as shown in Appendix B.1 or directly using a nonlinear programming solver.

4.1.3 Achievable Rate Results

The system analyzed in this section combines the concepts of FTN signaling and RLL coding. This combination can be beneficial in terms of spectral efficiency, cf. Section 2.2.2. At the same time the signal is oversampled w.r.t. its bandwidth. We will analyze the performance bounds w.r.t. the spectral efficiency by

- evaluating the bounds on p_b obtained previously, which are then used to obtain a bound $\mathrm{SE}_{\mathrm{LB,an}}$ on the spectral efficiency that can be evaluated foremost by analytical and numerical computations
- evaluating p_b based on a long realization of $\hat{\mathsf{x}}(t)$ via

$$p_{b\mathrm{sim}} = \frac{1}{N_{\mathrm{samp}}} \sum_{n=1}^{N_{\mathrm{samp}}} \frac{1}{2} \, \mathrm{erfc}\left(\frac{\hat{\mathsf{x}}_n \mathsf{x}_n}{\sqrt{2}\sigma_{\hat{\mathsf{n}}}} \right) \tag{4.39}$$

  yielding a simulation-aided bound $\mathrm{SE}_{\mathrm{LB,sim}}$ on the spectral efficiency

Tab. 4.1. Waveform overview

waveform	$g(t)$	$G(\omega)$	$P/\hat{P}$
rectangular	$\sqrt{\hat{P}}\,\mathrm{rect}\left(\frac{t}{T_s}\right)$	$\frac{2\sqrt{\hat{P}}\sin\left(\frac{\omega T_s}{2}\right)}{\omega}$	1
triangular	$\sqrt{\hat{P}}\,\mathrm{tri}\left(\frac{t}{T_s}\right)$	$\frac{2\sqrt{\hat{P}}(1-\cos(\omega T_s))}{\omega^2 T_s}$	$\left(1-\frac{2}{3\,\mathbb{E}[\mathsf{a}_k]}\right)$
cosine	$\frac{\sqrt{\hat{P}}}{2}\left(1+\cos\left(\frac{\pi t}{T_s}\right)\right)\mathbb{1}_{[-T_s,T_s]}(t)$	$\frac{\sqrt{\hat{P}}\pi^2\sin(\omega T_s)}{\omega(\pi^2-(\omega T_s)^2)}$	$\left(1-\frac{1}{2\,\mathbb{E}[\mathsf{a}_k]}\right)$

With respect to the waveform $g(t)$ in what follows we will consider a rectangular pulse of duration T_s, a triangular pulse of duration $2T_s$ (as in [BSK+17]), and a cosine wave of frequency $1/2T_s$ and duration $2T_s$ (as in Chapter 3 and [BDFa; BDF17a]). The waveforms are summarized in Table 4.1. The spectral efficiency of the system w.r.t. to a chosen bandwidth measure B_x is then given as

$$\mathrm{SE} = \frac{I'(\mathbf{A};\mathbf{D})}{B_x} \geq \frac{C_{\mathrm{rll}}(a_{\min}) - H_b(p_b)}{B_x T_s} = \mathrm{SE}_{\mathrm{LB}}. \tag{4.40}$$

Application to Strictly Bandlimited Channels

We now evaluate the spectral efficiency assuming ideal LP filters at transmitter and receiver. The bandwidth measure is therefore the occupied two-sided bandwidth $B_x = 2W$. In Fig. 4.3, the average bit flipping probability p_b and the resulting spectral efficiencies are depicted using a cosine waveform. The indices $(\cdot)_{\mathrm{sim}}$, $(\cdot)_{\mathrm{cheb}}$, $(\cdot)_{\mathrm{cond}}$, $(\cdot)_{\mathrm{noISI}}$ denote the results based on (4.39), (4.32), (4.38), and without filtering (i.e., $v_{\mathrm{tx}}(t) = v_{\mathrm{rx}}(t) = \delta(t)$), respectively. The following observations can be made

- For $a_{\min} = 1$, $p_{b,\mathrm{cheb}}$ gives the tighter bound on the average flipping probability compared to $p_{b,\mathrm{cond}}$. This is due to the fact that $p_{\mathsf{z}_{\mathrm{ISI}}|\mathsf{x}}(z_n|x_n) = p_{\mathsf{z}_{\mathrm{ISI}}}(z_n)$, whose symmetry is captured in $\Omega(z_n)$, which gives additional information in the maximization of p_b yielding a symmetric set of mass points. However, for all $a_{\min} > 1$, $p_{b,\mathrm{cond}}$ yields the tighter bounds. This is due to the better treatment of the constructive interference inherent to the RLL sequence. In what follows we, thus, consider
$$p_{b,\mathrm{UB}} = \begin{cases} p_{b,\mathrm{cheb}}, & a_{\min} = 1 \\ p_{b,\mathrm{cond}}, & a_{\min} > 1 \end{cases}. \tag{4.41}$$
- The analytically and numerically obtained bounds $p_{b,\mathrm{UB}}$ are especially valuable for describing low to medium ISI scenarios. If the ISI becomes too strong (cf. Fig. 4.3 for $2WT_s = 0.5$ and $a_{\min} = \{1, 2\}$) $p_{b,\mathrm{UB}}$ is not converging to zero for high SNR. Thus, the corresponding bounds on the spectral efficiency, $\mathrm{SE}_{\mathrm{LB,an}}$, will not converge against $\mathrm{SE}_{\mathrm{LB,sim}}$. If $a_{\min} > 1$, the bounds will converge if $z_1^* > -h_0$ as then a bit cannot be flipped purely due to ISI. Table 4.2 gives the

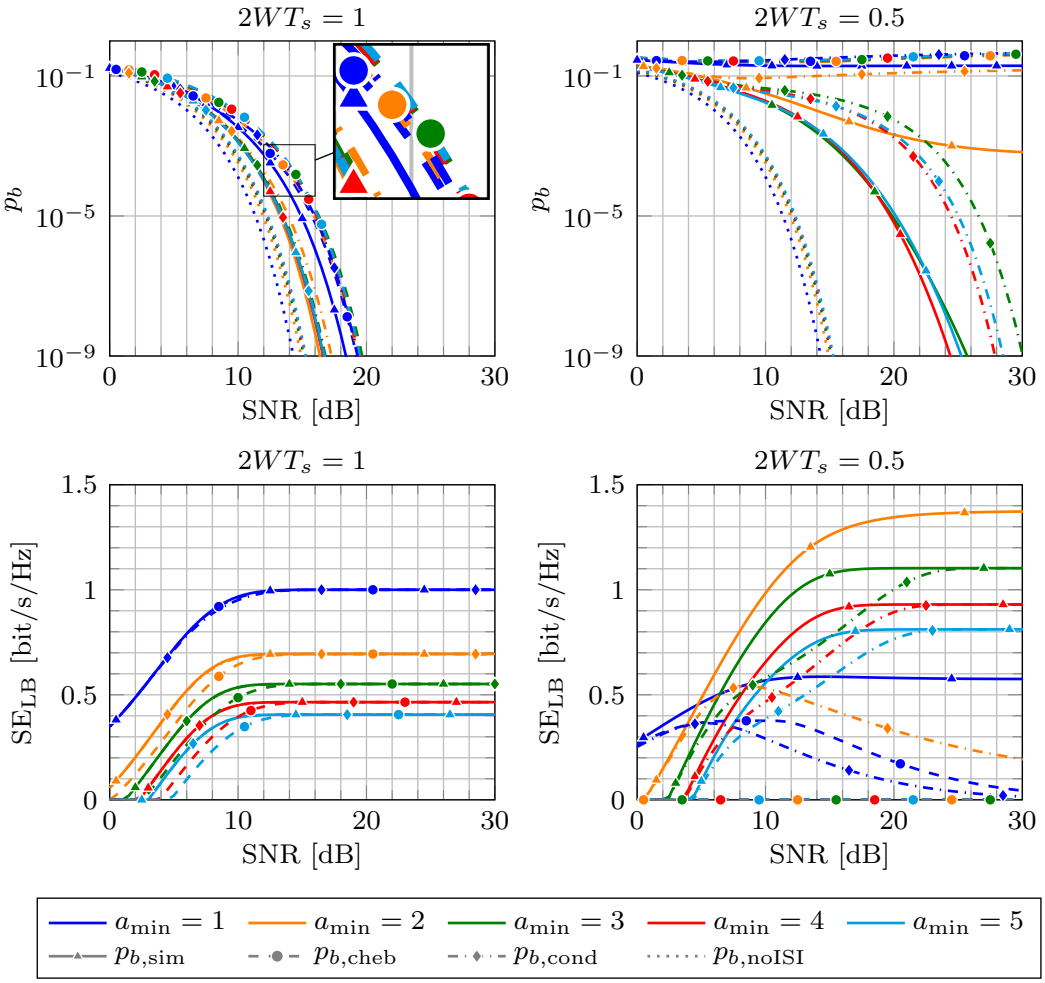

Fig. 4.3. Bounds on p_b and corresponding spectral efficiency for the cosine waveform and Nyquist signaling ($2WT_s = 1$) as well as two-fold FTN signaling ($2WT_s = 0.5$)

minimum bandwidth $(2WT_s)_{\min}$ for which this criterion is fulfilled. Since the ISI is dominated by the channel filters, there is no notable difference between the waveforms $g(t)$.

- Differently to $p_{b,\mathrm{sim}}$, it is possible that the bounds $p_{b,\mathrm{UB}}$ increase when the SNR increases. This occurs for strong ISI scenarios, i.e., if $z_1^* < -h_0$. Then the probability decreases that the error introduced by the ISI – which itself is bounded to be at a mass point close to its worst case value – is "repaired" by the noise.

- Where they converge, both bounds $\mathrm{SE_{LB,sim}}$ and $\mathrm{SE_{LB,an}}$ show increased spectral efficiencies for $2WT_s = 0.5$ and $a_{\min} > 1$. Thus, they confirm the benefit of applying FTN signaling and RLL sequences to 1-bit quantized channels and allow for insights on the performance of a concrete system implementation via numerical computations and very low complexity simulations.

Tab. 4.2. Minimum bandwidth $(2WT_s)_{\min}$ for which bounds and simulation results converge w.r.t. minimum runlength $a_{\min}$

$a_{\min}$	1	2	3	4	5	6
$(2WT_s)_{\min}$	0.96	0.57	0.44	0.36	0.33	0.33

In Fig. 4.4 the application of the spectral efficiency bounds to FTN transmission is depicted. For different bandwidths $2WT_s$ and SNR values the spectral efficiency is given for the $a_{\min}$ that maximizes $\mathrm{SE}_{\mathrm{LB}}$. The $a_{\min}$, at which the maximum is attained, is indicated by markers. The choice of the normalized bandwidth $2WT_s$ of $h_{\mathrm{LP}}(t)$ determines the degree of FTN signaling. Hereby, strictly speaking $2WT_s = 1$ is not Nyquist-signaling as $h(t)$ is not orthogonal w.r.t. $\frac{1}{2W}$ since $h_{\mathrm{LP}}(t)$ is convolved with $g(t)$. However, roughly we have

$$M_{\mathrm{FTN}} \approx \frac{1}{2WT_s} \tag{4.42}$$

i.e., $2WT_s = 1$ can be considered as Nyquist signaling, $2WT_s = 0.5$ as twofold FTN and so on.

We observe that, as in Chapter 3, due to the ideal LP-filters, the actual shape of the waveform plays a subordinate role and the achievable rate is mainly defined by $2WT_s$. Also, similarly, the maximum spectral efficiency in the high-SNR limit is observed using a triangular pulse as $\lim_{SNR\to\infty} \mathrm{SE}_{\mathrm{LB,sim}} \approx 1.53\,\mathrm{bit/s/Hz}$ for $a_{\min} = 4$ and $2WT_s = 0.29$. On the other hand, for low SNR values FTN signaling is not beneficial due to the increased ISI, which was to be expected. Furthermore, we observe that the $a_{\min}$ maximizing $\mathrm{SE}_{\mathrm{LB,sim}}$ is often different from the one maximizing $\mathrm{SE}_{\mathrm{LB,an}}$. This is due to the fact that the ISI is pushed to its limits if the spectral efficiency is maximized. In this case we have $z_1^* < -h_0$ and $p_{b,\mathrm{UB}}$ exhibits an error floor such that $\mathrm{SE}_{\mathrm{LB,an}}$ becomes loose and does not converge to $\mathrm{SE}_{\mathrm{LB,sim}}$ as discussed above. This can already be observed in Fig. 4.3 for $a_{\min} = 2$ for two-fold FTN. Exceptions are $2WT_s = \{0.6, 0.7, 1\}$ where $\mathrm{SE}_{\mathrm{LB,an}}$ and $\mathrm{SE}_{\mathrm{LB,sim}}$ reach the maximum for the same $a_{\min}$ and converge to the same values.

For the triangular and the cosine pulse, the bounds from Chapter 3 (c.f. Fig. 3.9) are plotted over the respective SNR range. In the latter case, they fit well to the envelope of the achievable rate curves for $\mathrm{SNR} \geq 10\,\mathrm{dB}$, while in the former, there is about 0.1 bit/s/Hz difference in the high-SNR limit. Moreover, the AWGN-spectral efficiency $\frac{1}{2}\log(1+\mathrm{SNR})$ is given for comparison with $\mathrm{SE}_{\mathrm{LB,sim}}$.

Performance with Excess Bandwidth

Excess bandwidth is known to notably increase the performance of FTN signaling. A small amount of out of band power is usually tolerable in a mobile communication

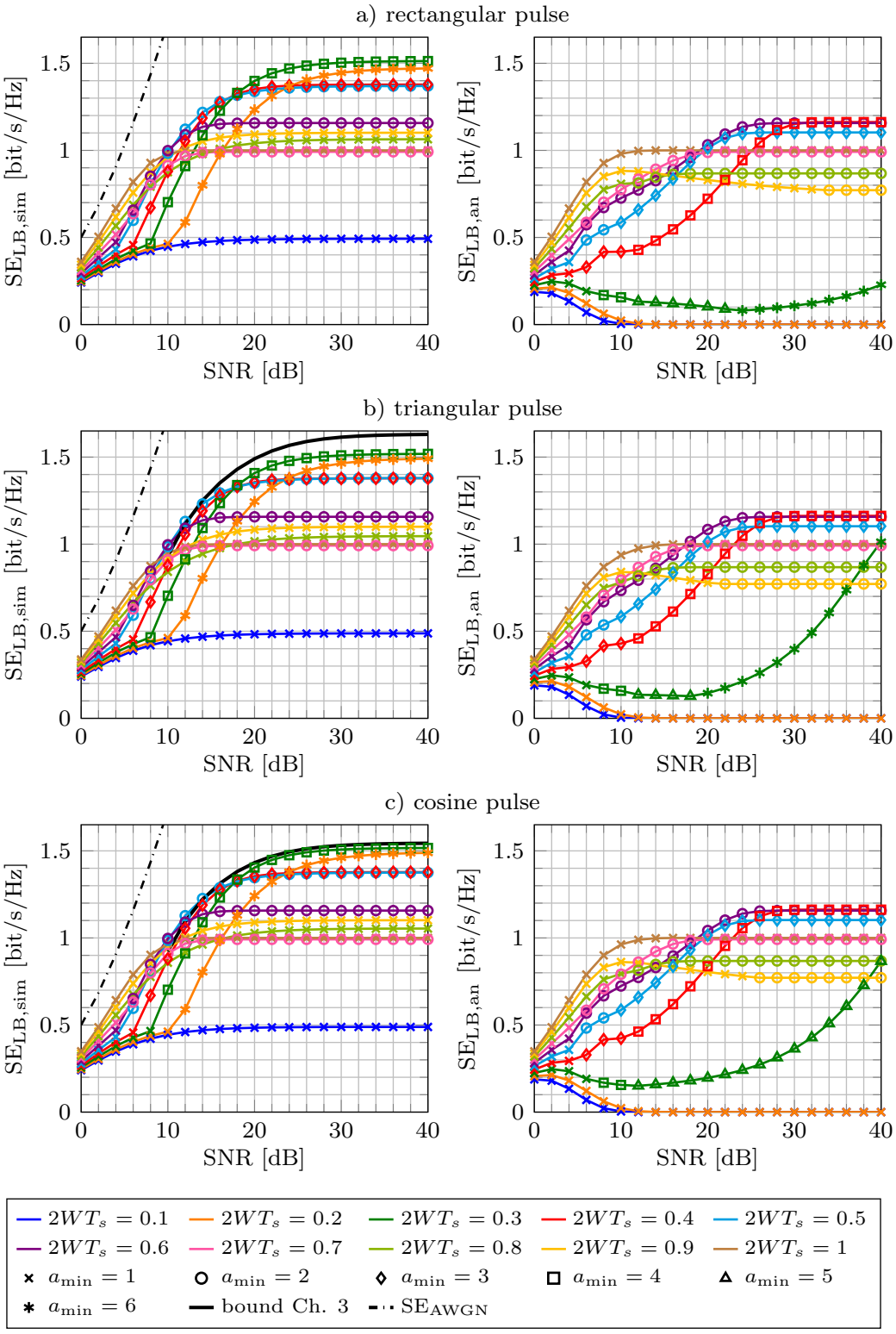

Fig. 4.4. Lower bounds on the maximum spectral efficiencies for strictly bandlimited channels and FTN signaling: simulation-aided bound $\mathrm{SE}_{\mathrm{LB,sim}}$ (left) and analytical bound $\mathrm{SE}_{\mathrm{LB,an}}$ (right) for three different pulses $g(t)$, a) rectangular pulse, b) triangular pulse, and c) cosine pulse

system. Therefore, we continue with analyzing the performance bounds of the 1-bit quantized FTN system using other filters than ideal LPes for $v_{\mathrm{tx}}(t)$ and $v_{\mathrm{rx}}(t)$. We assume that the pulse shape $h(t)$ is solely determined by $h_{\mathrm{LP}}(t)$, i.e., $g(t) = \delta(t)$. The choices of $h_{\mathrm{LP}}(t)$ that we analyze comprise

- raised cosine (RC) filters, i.e., $v_{\mathrm{tx}}(t)$ and $v_{\mathrm{rx}}(t)$ are root raised cosine (RRC) filters
- RRC filters, i.e., the convolution of $v_{\mathrm{tx}}(t)$ and $v_{\mathrm{rx}}(t)$ is an RRC filter
- squared raised cosine (SRC) filters, i.e., $v_{\mathrm{tx}}(t)$ and $v_{\mathrm{rx}}(t)$ are RC filters

with different roll-off factors α_{RC}. Note that here, except for $\alpha_{\mathrm{RC}} = 0$, $\hat{\mathrm{x}}(t)$ is not the noise-free receive signal anymore. As bandwidth measure B_x we use the 95 % power containment bandwidth $B_{95\%}$ of the transmit signal $\hat{\mathrm{x}}(t)$ and in order to define the degree of FTN signaling we continue to refer to the bandwidth $2W$ of an ideal LP, i.e., the bandwidth of the filter for $\alpha_{\mathrm{RC}} = 0$.

The spectral efficiency bounds for RC filters are depicted in Fig. 4.5. Here, for $2WT_s = 1$, the pulse $h(t)$ is truly orthogonal and

$$M_{\mathrm{FTN}} = \frac{T_{\mathrm{Nyq}}}{T_{\mathrm{symb}}} = \frac{1}{2WT_s}. \tag{4.43}$$

It can be observed that when increasing α_{RC} from 0 to 0.5 and further to 1, the convergence between $\mathrm{SE}_{\mathrm{LB,sim}}$ and $\mathrm{SE}_{\mathrm{LB,an}}$ improves with increased α_{RC} due to the decreasing ISI. Increasing the roll-off factor of the RC filter furthermore reduces the SNR that is required for the bounds to achieve a given spectral efficiency. While for example approximately 24 dB are required to achieve 1.8 bit/s/Hz for $\alpha_{\mathrm{RC}} = 0$, this reduces to 18 dB for $\alpha_{\mathrm{RC}} = 0.5$ and further to 17 dB for $\alpha_{\mathrm{RC}} = 1$. Again the AWGN-spectral efficiency $\frac{1}{2B_{95\%}T_s}\log(1+\mathrm{SNR})$ is given for comparison with $\mathrm{SE}_{\mathrm{LB,sim}}$. In the high SNR limit, the maximum values for the lower bounds on the spectral efficiency observed are 1.8, 2.23, and 2.09 bit/s/Hz for $\alpha_{\mathrm{RC}} = \{0, 0.5, 1\}$, respectively, requiring 3.33- and 5-fold FTN signaling. This confirms the expected benefit of excess bandwidth for FTN signaling.

Since for the considered configurations we observe a maximum of the spectral efficiency at $\alpha_{\mathrm{RC}} = 0.5$, there is clearly a trade-off between ISI and the bandwidth occupied w.r.t. the chosen bandwidth measure. While the former reduces with the roll-off factor of an RC filter, the latter increases. Besides the roll-off factor, which quantifies the additional bandwidth spent compared to an ideal LP, also the slope of the filter has an impact on the 95 % power containment bandwidth.

We therefore compare as a last step the maximum of $\mathrm{SE}_{\mathrm{LB,sim}}$ over $a_{\min}$ for the RRC, RC, and SRC filter, respectively, as mentioned above. The results are depicted in Fig. 4.6. For the sake of presentation, indicating the $a_{\min}$ for which these spectral

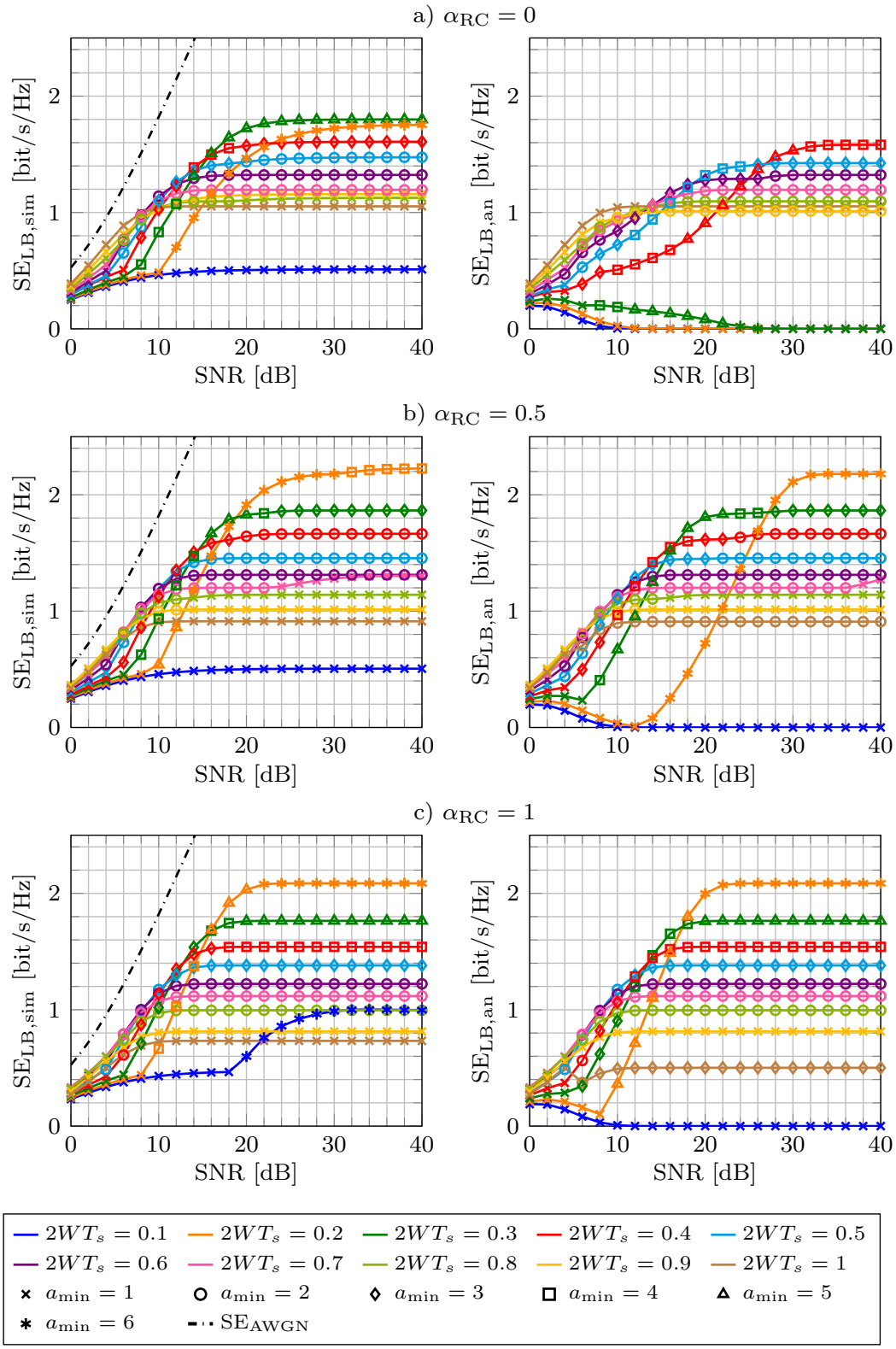

Fig. 4.5. Lower bounds on the maximum spectral efficiencies for channels with FTN signaling and excess bandwidth (RC-pulses) w.r.t. 95 %-power containment bandwidth: simulation-aided bound $\mathrm{SE}_{\mathrm{LB,sim}}$ (left) and analytical bound $\mathrm{SE}_{\mathrm{LB,an}}$ (right) for three different roll-off factors, a) $\alpha_{\mathrm{RC}} = 0$, b) $\alpha_{\mathrm{RC}} = 0.5$, and c) $\alpha_{\mathrm{RC}} = 1$

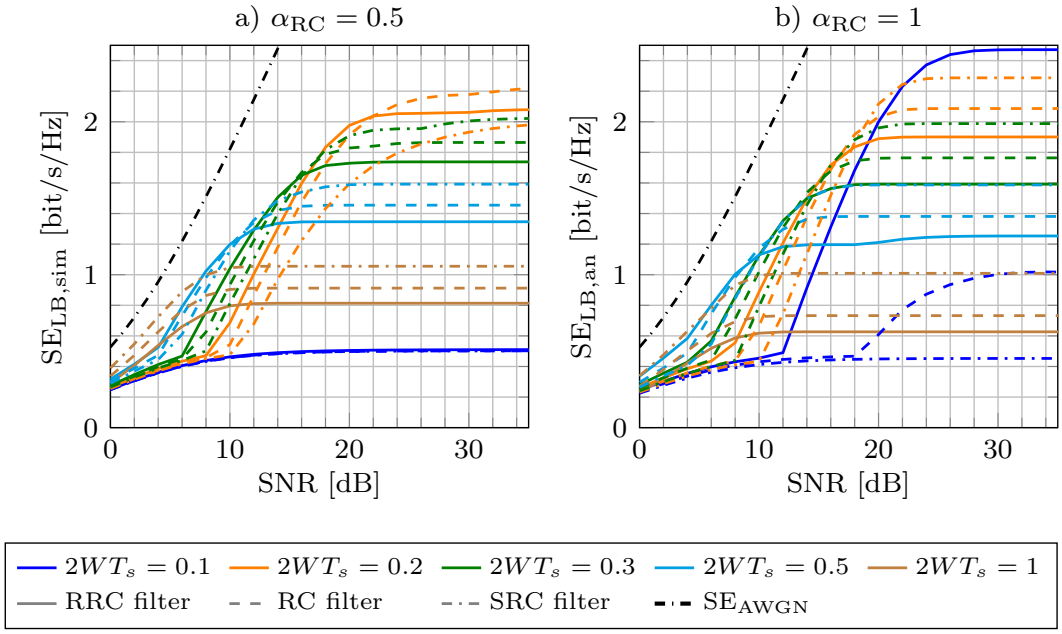

Fig. 4.6. Simulation-aided lower bound $\mathrm{SE}_{\mathrm{LB,sim}}$ on the maximum spectral efficiencies for channels with FTN signaling and excess bandwidth w.r.t. 95 %-power containment bandwidth for the roll-off factors a) $\alpha_{\mathrm{RC}} = 0.5$, b) $\alpha_{\mathrm{RC}} = 0.5$

efficiencies are obtained is omitted. The largest values for $\mathrm{SE}_{\mathrm{LB,sim}}$ are obtained for $a_{\min} = \{4, 5, 6\}$. There is no need to consider $\alpha_{\mathrm{RC}} = 0$ since then all three filters are equivalent and the spectral efficiencies are already given in Fig. 4.5a). At high SNR and for a fixed $2WT_s$ it can be seen in Fig. 4.6 that in most cases the spectral efficiency of the SRC filter exceeds the one of the RC filter, which in turn exceeds the spectral efficiency of the RRC filter. This is true as long as sufficient bandwidth is available to accommodate the spectrum of the RLL sequence without major distortion since the bandwidth measure $B_{95\%}$ increases from SRC to RC to RRC filter for a fixed $2WT_s$. However, for large FTN factors, the steeper slopes of the RC and RRC filter reduce the ISI, which results beneficial and leads to an $\mathrm{SE}_{\mathrm{LB,sim}}$ up to the previously observed 2.23 bit/s/Hz for $\alpha_{\mathrm{RC}} = 0.5$ and up to 2.47 bit/s/Hz for $\alpha_{\mathrm{RC}} = 1$. For low SNR values, the reduced ISI of RRC filters results in larger spectral efficiencies for $2WT_s < 1$.

4.2 Application to MIMO Channels

In [BDF17b] we analyzed power allocation schemes for oversampled multiple-input multiple-output (MIMO) channels based on previously obtained lower bounds on the achievable rate for a continuous and a discrete-time scenario. We found that assuming full channel state information at the transmitter (CSIT), simple precoding and power allocation schemes can deliver good performance results. Precoding plays an important role for 1-bit quantized MIMO transmission since due to the 1-bit quantizer any post-

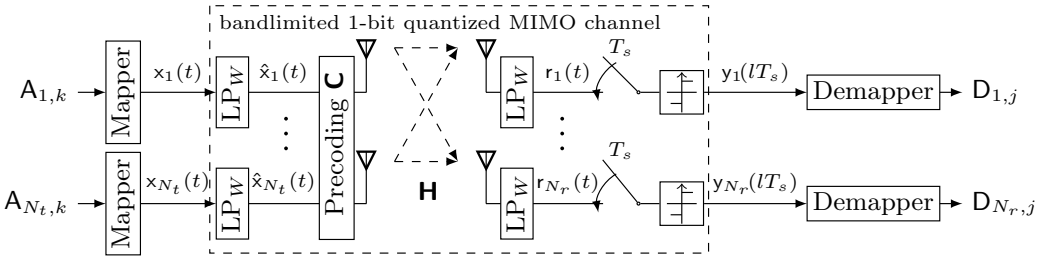

Fig. 4.7. System model of the 1-bit quantized multiple-input multiple-output channel with runlength modulation

processing at the receiver in the digital domain becomes more difficult. So while the assumption of full CSIT is arguable, it is a common starting point for analysis. In this section we briefly review the system model and approach of [BDF17b] and then evaluate the most suitable precoding and power allocation schemes w.r.t. the new results obtained for the bounds in Chapter 3 and Section 4.1.

4.2.1 System Model

The considered channel is a bandlimited complex $N_r \times N_t$ MIMO channel with 1-bit quantization and oversampling at the receiver. Here, N_r and N_t denote the number of receive and transmit antennas, respectively. Fig. 4.7 depicts the respective system model. Note that the in-phase and quadrature components of the signal are processed separately at mapper, analog-to-digital converter (ADC), and demapper. The information continues to be encoded in the position of the ZCs of the real and imaginary part of the transmitted signal, respectively.

For every input i, there are two input vectors $\mathbf{A}_{i,\Re}^{(K)}$ and $\mathbf{A}_{i,\Im}^{(K)}$ containing the ZC-distances of the real and the imaginary part of the signal $\mathsf{x}_i(t)$, respectively. The design of the signal $\mathsf{x}_i(t)$ corresponds to the approaches from Chapter 3 and Section 4.1, i.e.,

$$\mathsf{x}_{i,\Re/\Im}(t) = \Big(\sum\nolimits_{k=1}^{K} \sqrt{\hat{P}_i}(-1)^k g(t - \mathsf{T}_{i,\Re/\Im,k}) \Big) + \sqrt{\hat{P}_i} \tag{4.44}$$

considering a continuous-time channel model and

$$\mathsf{x}_{i,\Re/\Im}(t) = \sqrt{\hat{P}_i} \sum_{n=0}^{N_{\text{samp}}-1} \mathsf{x}_{i,\Re/\Im,n}\, g(t - nT_s) \tag{4.45}$$

for the discrete-time model. All signal parameters are defined in the respective sections. In what follows, we use the matrix notation, where the sampling period T_s is used for defining the time instants. Note that the continuous-time model in (4.44) implies that $T_s \to 0$.

Before transmission the signal $\mathsf{x}_i(t)$ is filtered by an ideal LP with one-sided bandwidth W. This yields the transmit signal $\hat{\mathsf{x}}_i(t)$, which is transmitted via an $N_r \times N_t$

MIMO channel with precoding. The received signal $\mathbf{y} = [\mathsf{y}_1(lT_s), \mathsf{y}_2(lT_s), ...\mathsf{y}_{N_r}(lT_s)]^T$ after sampling and quantization is given by

$$\mathbf{y} = Q_1\left(\mathbf{H}\mathbf{C}\hat{\mathbf{x}} + \hat{\mathbf{n}}\right). \tag{4.46}$$

Here, $\mathbf{H}$ is the time-variant $N_r \times N_t$ channel matrix containing the complex Gaussian distributed channel coefficients at time lT_s and $\hat{\mathbf{x}} = [\hat{\mathsf{x}}_1(lT_s), \hat{\mathsf{x}}_2(lT_s), ...\hat{\mathsf{x}}_{N_t}(lT_s)]^T$. Furthermore, and $\hat{\mathbf{n}}$ is the filtered and sampled complex AWGN vector with mean zero and variance $\sigma^2_{\hat{\mathsf{n}}}$. The noise samples are temporally correlated due to bandlimitation and oversampling. Moreover, $Q_1(\cdot)$ denotes the 1-bit quantization function and $\mathbf{C}$ is the precoding operation. We will omit the sampling time index l at $\mathbf{H}$, $\hat{\mathbf{x}}$, $\hat{\mathbf{n}}$ and $\mathbf{y}$ for notational convenience.

As motivated in Chapter 1, we are interested in high-speed short-range communication scenarios. For very short range board-to-board communication with line-of-sight component it was found that such channels can be considered largely frequency flat despite their large bandwidth [FuHLF13]. Therefore, we assume full channel state information (CSI) at transmitter and receiver. Channel estimation with 1-bit ADCs is a topic under active research, e.g., [IN07; MAN10; MSPH14; ZSK12; SLd19].

At the receiver, the positions of the ZCs of the real and imaginary component are mapped back onto vectors of distances $\mathbf{D}^J_{j,\Re}$ and $\mathbf{D}^J_{j,\Im}$ for every receive signal y_j. Recall that the length K of the input vector is not necessarily equal to the length J of the output vector since the noise can insert or delete ZCs. The signals $\mathsf{x}_i(t)$ in (4.44) and (4.45) are not bandlimited such that LP-filtering introduces a distortion to the signal. This yields to a reduced achievable rate due to ISI, which has been modeled in Chapter 3 and Section 4.1 accordingly. We assume that the noise is i.i.d. for all receive antennas. The sum transmit power is

$$P = \begin{cases} \mathbb{E}[\|\mathbf{C}\hat{\mathbf{x}}\|^2] & \text{for the discrete-time model} \\ \mathbb{E}[\|\mathbf{C}\mathbf{x}\|^2] & \text{for the continuous-time model}^1 \end{cases}. \tag{4.47}$$

4.2.2 Channel Decomposition and Power Allocation Schemes

It is known that the capacity of the MIMO channel without output quantization is achieved by circular-symmetric complex Gaussian input symbols with covariance matrix $\mathbf{R}_{\mathsf{x}} = \mathbf{W}\boldsymbol{P}\mathbf{W}^H$. Here, $\mathbf{W}$ is obtained by singular value decomposition (SVD) of $\mathbf{H} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{W}^H$, where $\mathbf{U}$ and $\mathbf{W}$ are unitary. The diagonal matrix $\boldsymbol{\Sigma}$ contains the weights σ_v for every channel, which are the singular values of $\mathbf{H}$. Based on SVD we can pre- and post-process the transmit and receive signals in order to obtain

[1] Refer to Section 3.1.2 for the reasoning behind this definition.

$\nu = \mathrm{rank}(\mathbf{H}) \leq \min(N_t, N_r)$ independent and non-interfering channels. We assume that channel coefficients in $\mathbf{H}$ are generated from $\mathcal{CN}(0,1)$. Then the receive SNR on the vth channel becomes $\rho_{v,\mathrm{rx}} = \sigma_v^2 \frac{P_v}{\sigma_{\hat{\mathrm{n}}}^2}$. It then remains to find an appropriate power distribution $\boldsymbol{P} = \mathrm{diag}(P_1, ..., P_\nu)$ among all channels given that

$$P = \sum_{v=1}^{\nu} P_v. \tag{4.48}$$

With 1-bit quantization, however, it is not possible to apply the post-processing the receiver since multiplication of the coarsely quantized signal with $\mathbf{U}^H$ would not yield the desired result. Thus, precoding is often used to shape the signal already at the transmitter. One option is to carry out channel equalization already at the transmitter by channel inversion (CI). This comes at the cost of potentially increased transmit power or a decreased receive SNR, respectively, since the different channel weights σ_v cannot be leveraged and $\rho_{v,\mathrm{rx}} = \frac{P_v}{\sigma_{\hat{\mathrm{n}}}^2}$.

In any case, based on the corresponding processes for $K \to \infty$ the achievable rate is given by the sum rate of the achievable rates on the individual channels

$$I'(\mathbf{A}; \mathbf{D}) = \sum_{v=1}^{\nu} I'(\mathbf{A}_v; \mathbf{D}_v) \tag{4.49}$$

where $\mathbf{A} = [\mathbf{A}_{1,\Re}, \mathbf{A}_{1,\Im}, ..., \mathbf{A}_{\nu,\Re}, \mathbf{A}_{\nu,\Im}]$ and $\mathbf{D} = [\mathbf{D}_{1,\Re}, \mathbf{D}_{1,\Im}, ..., \mathbf{D}_{\nu,\Re}, \mathbf{D}_{\nu,\Im}]$. With the techniques obtained previously, (4.49) can be evaluated in terms of lower bounds, cf. (3.51) and (4.18).

One can think of different power allocation schemes that can be applied to MIMO systems. Fig. 4.8 summarizes the schemes analyzed in [BDF17b], which are

- SVD with waterfilling (WF) (SVD WF)
- SVD with equal power allocation (EP) (SVD EP)
- SVD with a maximum effective channel SNR (SVD $\rho_{\max}$), i.e., no more power is assigned to a channel that has already the maximum receive SNR $\rho_{v,\mathrm{rx}} = \rho_{\max}$ until every channel has this maximum receive SNR
- CI at the transmitter with with equal power allocation (CI Tx)
- CI at the transmitter using only a subset of the ν channels such that the sum rate is maximized (CI Tx opt).

Out of the above, only the CI based power allocation schemes can be implemented straightforward under 1-bit quantization and CI Tx opt has been shown in [BDF17b] to perform close to or even outperform the SVD based schemes. Nevertheless, the following evaluation of (4.49) will consider all schemes for the sake of comparison.

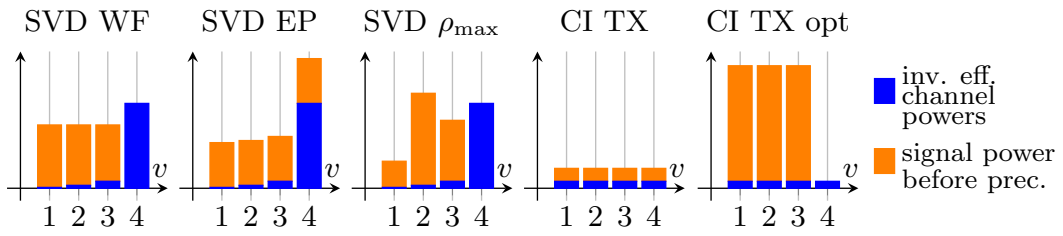

Fig. 4.8. Schematic illustration of the possible power allocation schemes [BDF17b]

4.2.3 Results on the Achievable Rate

The results on the achievable rate are depicted in Fig. 4.9 for the 2×2 and 4×4 MIMO channel. For comparison existing bounds for Nyquist sampling are given: an upper bound (UB) [MH15] as well as lower bounds (LBs) based on Bussgang decomposition [MN12], CI and an additive quantization noise model (AQNM) [MH15]. Furthermore, the Gaussian capacity C_{AWGN} using waterfilling is given. We assume that the complex noise is of unit variance, i.e., $\sigma_{\mathsf{n}}^2 = 1$, such that the achievable rate can be given over the amount of transmit power P that can be allocated.

We depict the lower bound on the spectral efficiency of the continuous-time channel for the linear transition waveform and the near-optimal $\kappa = 0.7$. The low SNR range is omitted due to the restrictions on the validity of the bound. The lower bound on the spectral efficiency of the discrete-time channel is given for the linear transition waveform and $a_{\min} = 4$. We observe that the spectral efficiencies achieved with the discrete-time model are only about 0.5 bit below the ones of the continuous-time channel for the 2×2 MIMO channel, which increases 1 bit for the 4×4 MIMO channel. Furthermore, except classical CI all schemes perform comparably. This is due to the noise enhancement of CI. However, it also shows that with the adapted channel inversion scheme CI Tx opt spectral efficiencies comparable to SVD can be achieved without post-processing the 1-bit quantized receive signal. Note that this requires full CSI at transmitter and receiver, which is also true for SVD. Furthermore, we observe that oversampling improves the achievable spectral efficiencies compared to Nyquist sampling especially in the high-SNR regime, while for the 2×2 MIMO channel in the low-SNR regime the lower bound based on CI shows higher spectral efficiencies. Overall, we conclude that for 1-bit quantized oversampled MIMO channels, intelligent precoding is sufficient to achieve a good performance.

4.3 Summary

We have applied the insights obtained in Chapter 3 to more practical scenarios. We used the bounding techniques based on the genie-aided receiver for discrete channels with finite OSRs. With this approach, lower bounds on the achievable rate of the

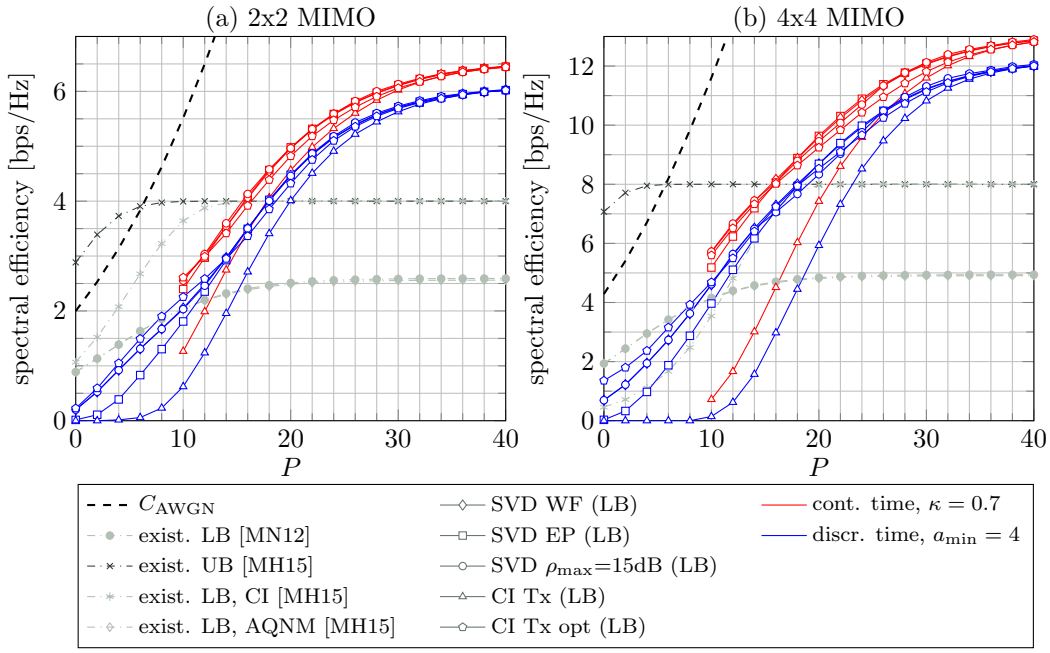

Fig. 4.9. Lower bounds on the spectral efficiencies of the oversampled strictly bandlimited channel MIMO channel for discrete and continuous-time scenarios with a linear waveform

runlength-modulated 1-bit quantized discrete-time AWGN channel using FTN signaling have been derived. Here, FTN signaling corresponds to oversampling w.r.t. the channel bandwidth. Two lower bounds are given based on maximizing the average sample flipping probability p_b, which in turn maximizes the entropy of the auxiliary process. First, an analytical expression for p_b is obtained by deriving a worst-case ISI distribution. Secondly, an estimate of p_b based on a sufficiently long realization of the filtered transmit signal yields a simulation-aided lower bound on the spectral efficiency. With increasing ISI, the bounds diverge until the analytical bound collapses. This is due to the fact that it is based on a worst-case assumption for the probability density function (pdf) of the ISI. However, for low to moderate ISI, both bounds converge. Given a hard bandlimited channel, the achievable rates derived match the results of the continuous-time channel well achieving up to approximately 1.53 bit/s/Hz. When excess bandwidth is tolerable, spectral efficiencies up to 2.47 bit/s/Hz in the real domain are achievable w.r.t. the 95 %-power containment bandwidth.

We show that these bounds as well as the ones for the continuous-time scenario can be applied to the oversampled 1-bit quantized MIMO channel using appropriate power allocation schemes. Especially in the high-SNR regime the gain by oversampling or FTN signaling is significant.

5 The Potential of Continuous Phase Modulation for 1-bit Quantized Communication

As constant envelope modulation scheme, continuous phase modulation (CPM) enables the use of highly efficient non-linear, e.g., class-E, power amplifiers (PAs). Originally designed to reduce spectral side lobes, the continuous phase trajectory provides the opportunity to extract additional information by oversampling the signal. Thus, it appears to be a natural match to oversampled 1-bit quantized communication. In this chapter, basic properties of CPM are introduced and system model of the oversampled 1-bit quantized CPM system is given. We analyze the limiting performance in terms of distinguishable symbols in a noise-free scenario as well as the bit error rate (BER) performance under noise given channel filters and faster-than-Nyquist (FTN) signaling. Parts of the presented results are published in [BDF19; FDB+19].

5.1 Background on Continuous Phase Modulation

A passband CPM-signal can be described as [AAS86, Chapter 2.2]

$$\mathsf{s}(t) = \Re\left\{\sqrt{\frac{2E_s}{T_{\text{symb}}}} e^{j(2\pi f_0 t + \phi(t))}\right\} \tag{5.1}$$

where T_{symb} is the symbol duration, E_s is the symbol energy of the bandpass signal, and f_0 is the carrier frequency. The phase term $\phi(t)$ is the superposition of all previous phase transitions up to the present time t. It depends on the corresponding CPM-symbols $\alpha_k \in \mathcal{A}$ as

$$\phi(t) = 2\pi h_{\text{cpm}} \sum_{k=0}^{K-1} \alpha_k q(t - kT_{\text{symb}}) + \phi_0. \tag{5.2}$$

Here, ϕ_0 is the initial phase at $t = 0$, K is the number of symbols, and h_{cpm} is the modulation index. Moreover, $|\mathcal{A}| = M_{\text{cpm}}$ is the order of the modulation alphabet. Conventionally, M_{cpm} is chosen to be even with the symbol alphabet being

$$\mathcal{A} = \{\pm 1, \pm 3, ..., \pm(M_{\text{cpm}} - 1)\}. \tag{5.3}$$

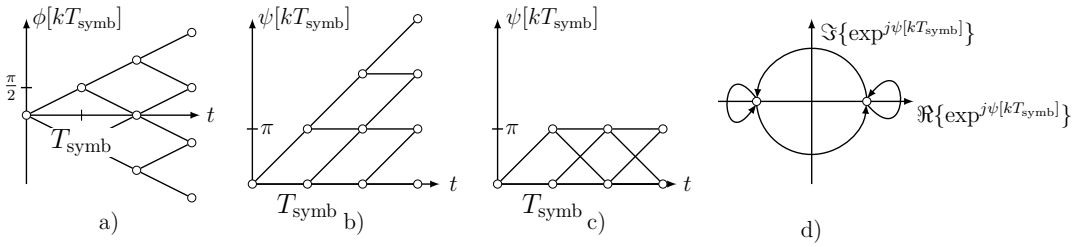

Fig. 5.1. Trellis diagram for $M_{\text{cpm}} = 2$, $h_{\text{cpm}} = 1/2$, and $\phi_0 = 0$: a) time variant trellis based on $\phi(t)$, b) time invariant trellis based on $\psi(t)$, c) wrapped trellis and d) constellation diagram

Furthermore, $q(t)$ is the phase smoothing response, for which it holds

$$q(t) = \begin{cases} 0, & t \leq 0 \\ \frac{1}{2}, & t > L_{\text{cpm}} T_{\text{symb}} \end{cases}. \tag{5.4}$$

Here, L_{cpm} is the memory of the CPM-symbol in terms of the symbol duration T_{symb}. The choice of $q(t)$ for $0 < t \leq L_{\text{cpm}} T_{\text{symb}}$ defines the phase transition. The derivative in time $q'(t)$ of $q(t)$ represents the instantaneous frequency of the pulse. It can be used to define CPM families. Examples are the L_{cpm}-REC family based on a rectangular $q'(t)$, i.e., a linear $q(t)$, or the L_{cpm}-(S)RC families based on a (spectral) raised cosine frequency pulse. A special case of the L_{cpm}-REC family is continuous phase frequency shift keying (CPFSK) corresponding to 1-REC, which in turn becomes minimum shift keying (MSK) when $h_{\text{cpm}} = 1/2$, $M_{\text{cpm}} = 2$. Classical phase shift keying (PSK) can be represented in this framework based on a step function as phase smoothing response $q(t)$, i.e., based on the Dirac delta function as frequency pulse $q'(t)$.

The phase term $\phi(t)$ can be represented using a trellis diagram. With the definition of the alphabet $\mathcal{A}$ in (5.3), a time variant trellis results as given for $M_{\text{cpm}} = 2$ in Fig. 5.1a). It is possible to transform a time variant into a time invariant trellis by adding a frequency offset [Rim88] as depicted in Fig. 5.1b), i.e.,

$$\psi(t) = \phi(t) + 2\pi\Delta f \tag{5.5}$$

with

$$\Delta f = \frac{h_{\text{cpm}}(M_{\text{cpm}} - 1)}{2T_{\text{symb}}}. \tag{5.6}$$

Fig. 5.1c) depicts the wrapped trellis, which results when taking into account the periodicity of the phase. Furthermore, in Fig. 5.1d) the constellation diagram is given, which comprises n_{const} constellation points in the complex plane. From each constellation point M_{cpm} transitions are possible. Note that n_{const} is not necessarily equal to M_{cpm} and that in order for n_{const} to be finite, h_{cpm} needs to be a rational number.

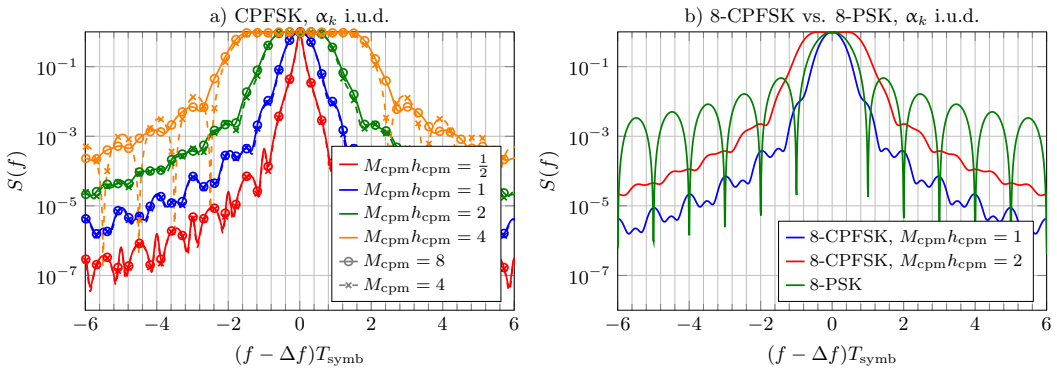

Fig. 5.2. CPFSK spectrum a) for different M_{cpm} and h_{cpm} and b) 8-CPFSK vs. 8-PSK

Compared to pulse amplitude modulation schemes obtaining the spectrum of a general CPM signal is more involved. Closed form solution exist for example for PSK, CPFSK, and MSK. For more general configurations and independent and uniformly distributed (i.u.d.) symbols, an algorithm for computing the power spectral density (PSD) based on the auto-correlation function (ACF) is given in [AAS86, Chapter 4.2]. The closed form expression available for M_{cpm}-ary PSK is [AAS86, Chapter 4.3]

$$S(f) = \frac{T_{\text{symb}}}{\log_2 M_{\text{cpm}}} \frac{\sin^2\left(\pi f \frac{T_{\text{symb}}}{\log_2 M_{\text{cpm}}}\right)}{\left(\pi f \frac{T_{\text{symb}}}{\log_2 M_{\text{cpm}}}\right)^2} \tag{5.7}$$

while the CPFSK spectrum is given as

$$\begin{aligned} S(f) &= \frac{T_{\text{symb}}}{\log_2 M_{\text{cpm}}} \frac{2}{M_{\text{cpm}}} \sum_{i=1}^{M_{\text{cpm}}} \left(\frac{1}{2} \frac{\sin^2 \nu_i}{\nu_i^2} + \frac{1}{M_{\text{cpm}}} \sum_{j=1}^{M_{\text{cpm}}} \frac{B_{ij} \sin \nu_i}{\nu_i} \frac{\sin \nu_j}{\nu_j} \right) \\ \nu_i &= \pi \left(f \frac{T_{\text{symb}}}{\log_2 M_{\text{cpm}}} - \frac{(2i - M_{\text{cpm}} - 1)h_{\text{cpm}}}{2} \right) \\ B_{ij} &= \frac{\cos(\nu_i + \nu_j) - \xi_\alpha \cos\left(\nu_i + \nu_j - 2\pi f \frac{T_{\text{symb}}}{\log_2 M_{\text{cpm}}}\right)}{1 - 2\xi_\alpha \cos\left(2\pi f \frac{T_{\text{symb}}}{\log_2 M_{\text{cpm}}}\right) + \xi_\alpha^2} \\ \xi_\alpha &= \frac{1}{M_{\text{cpm}}} \frac{\sin M_{\text{cpm}} \pi h_{\text{cpm}}}{M_{\text{cpm}} \pi h_{\text{cpm}}}. \end{aligned} \tag{5.8}$$

Fig. 5.2b) illustrates that the difference between (5.7) and (5.8) is mainly in the height of the spectral side lobes while Fig. 5.2a) shows the spectrum of CPFSK depending on M_{cpm} and h_{cpm}. The bandwidth occupied is mainly determined by the product $h_{\text{cpm}} M_{\text{cpm}}$. This is due to the fact that with (5.3) the maximum difference in instantaneous frequency between two symbols at time k is $h_{\text{cpm}}{}^{(M_{\text{cpm}}-1)}/_{T_{\text{symb}}}$. The price to pay for decreasing the bandwidth by reducing h_{cpm} is a decreased Euclidean distance $d_{\min}$. The distance between two paths trough the trellis can be obtained by using a

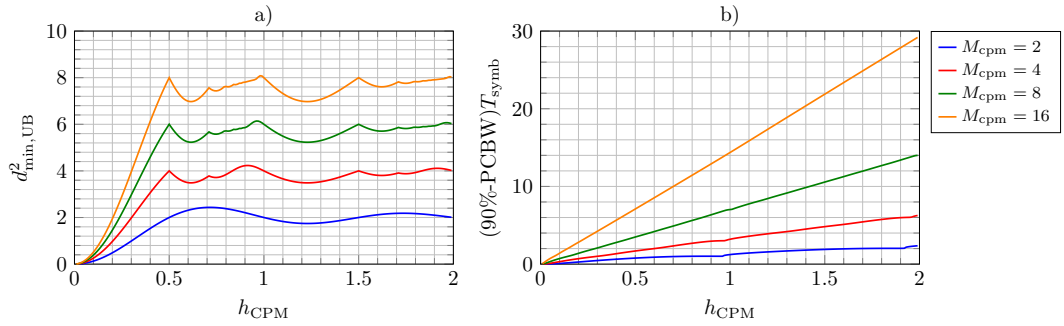

Fig. 5.3. Comparison of a) upper bound on the squared Euclidean distance and b) 90 %-power containment bandwidth for CPFSK and different M_{cpm} depending on h_{cpm}

difference sequence $\boldsymbol{\gamma} = \boldsymbol{\alpha} - \tilde{\boldsymbol{\alpha}}$, where $\boldsymbol{\alpha}$ and $\tilde{\boldsymbol{\alpha}}$ are any two possible distinct sequences of CPM symbols. The minimum distance results to [AAS86, Chapter 3]

$$d^2_{\min} = \log_2 M_{\mathrm{cpm}} \min_{\boldsymbol{\gamma}_K} \left[K - \frac{1}{T_{\mathrm{symb}}} \int_0^{KT_{\mathrm{symb}}} \cos\phi(t, \boldsymbol{\gamma}_K) dt \right]. \tag{5.9}$$

Since testing all possible difference sequences $\boldsymbol{\gamma}$ quickly becomes tedious, an upper bound can be obtained choosing any of the possible distance sequences $\boldsymbol{\gamma}$. Good candidates for tight upper bounds are infinitely long sequences merging as soon as possible preferably independently of h_{cpm}. Based on these so-called inevitable mergers of paths, e.g., $\boldsymbol{\gamma} = [\gamma_0, -\gamma_0, \ldots.]$, this yields for M_{cpm}-ary CPFSK with $L_{\mathrm{cpm}} = 1$ and constant h_{cpm}

$$d^2_{\min,\mathrm{UB}} = \log_2 M_{\mathrm{cpm}} \min_{1 \le i \le M_{\mathrm{cpm}}-1} 2 \left(1 - \frac{\sin(2\pi i h_{\mathrm{cpm}})}{2\pi i h_{\mathrm{cpm}}} \right). \tag{5.10}$$

Fig. 5.3 depicts the upper bound in (5.10) on the minimum distance as well as the normalized 90 %-power containment bandwidth (PCBW) depending on h_{cpm}, which shows that up to $h_{\mathrm{cpm}} = 0.5$ robustness can be obtained at the cost of bandwidth.

5.2 System Model

The system model of the 1-bit quantized oversampled CPM system is depicted in Fig. 5.4. In every time interval T_{symb}, $\log_2(M_{\mathrm{cpm}})$ bits are mapped to a CPM-symbol α_k using Gray mapping. The phase modulated signal is then generated using the phase smoothing response $q(t)$, cf. (5.2). After transmit filtering with $v_{\mathrm{tx}}(t)$, the signal is up-converted to the carrier frequency f_0. This gives the transmit signal $\hat{\mathsf{s}}(t)$ as filtered version of $\mathsf{s}(t)$, cf. (5.1), which is transmitted over an additive white Gaussian noise (AWGN) channel. The AWGN signal $\mathsf{n}(t)$ has noise power density N_0. After

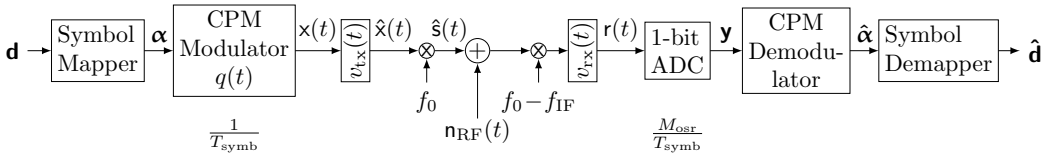

Fig. 5.4. System model

down-conversion to an intermediate frequency f_{IF} and receive filtering with $v_{\mathrm{rx}}(t)$, the received signal becomes

$$\mathsf{r}(t) = v_{\mathrm{rx}}(t) * (v_{\mathrm{tx}}(t) * \mathsf{x}(t) + \mathsf{n}(t)) \tag{5.11}$$

$$= h(t) * \mathsf{x}(t) + \hat{\mathsf{n}}(t). \tag{5.12}$$

It is oversampled with sampling rate $f_s = \frac{M_{\mathrm{osr}}}{T_{\mathrm{symb}}}$ and quantized with two 1-bit quantizers, one for each dimension. Thus, we have $\mathbf{y} = [\mathbf{y}_0, ..., \mathbf{y}_{K-1}]$ and $\mathbf{y}_k = [\mathsf{y}_{k,1}, ..., \mathsf{y}_{k,M_{\mathrm{osr}}}]$. With

$$\mathsf{r}_{k,m} = \mathsf{r}\left(\left(k + \frac{m}{M_{\mathrm{osr}}}\right) T_{\mathrm{symb}}\right),\ m = 1, ..., M_{\mathrm{osr}} \tag{5.13}$$

it results for the complex received sample

$$\mathsf{y}_{k,m} = Q_1(\Re\{\mathsf{r}_{k,m}\}) + jQ_1(\Im\{\mathsf{r}_{k,m}\}). \tag{5.14}$$

Note that oversampling is defined here w.r.t. the symbol duration T_{symb}. Thus, $M_{\mathrm{osr}} = M_{\mathrm{osr},t}$, cf. (2.16), and $T_s = \frac{T_{\mathrm{symb}}}{M_{\mathrm{osr}}}$ is the sampling period. Here, $Q_1(\cdot)$ is the 1-bit quantization function. Defining

$$\mathsf{x}_{k,m} = \mathsf{x}\left(\left(k + \frac{m}{M_{\mathrm{osr}}}\right) T_{\mathrm{symb}}\right), \tag{5.15}$$

$$\mathsf{n}_{k,m} = \mathsf{n}\left(\left(k + \frac{m}{M_{\mathrm{osr}}}\right) T_{\mathrm{symb}}\right) \tag{5.16}$$

and

$$\hat{\mathsf{n}}_{k,m} = \hat{\mathsf{n}}\left(\left(k + \frac{m}{M_{\mathrm{osr}}}\right) T_{\mathrm{symb}}\right) \tag{5.17}$$

the vectors $\mathbf{x}$, $\hat{\mathbf{x}}$, $\mathbf{n}$, and $\hat{\mathbf{n}}$ as well as $\mathbf{x}_k$, $\hat{\mathbf{x}}_k$, $\mathbf{n}_k$, and $\hat{\mathbf{n}}_k$ can be expressed accordingly. The estimated received symbols $\hat{\boldsymbol{\alpha}} = \hat{\alpha}_0^{K-1}$ are mapped to the bit sequence $\hat{\mathbf{d}}$. Both, the transmit and the receive filter are root raised cosine (RRC) filters with bandwidth time product $2WT_{\mathrm{symb}}$ and roll-off factor α_{RC}. Thus, the channel impulse response $h(t)$ is a raised cosine (RC)-filter.

In what follows, we focus on parameter settings that results into time invariant trellis diagrams and constellation diagrams that are symmetric w.r.t. the quantizer thresholds, i.e., the axes in the complex plane. The former facilitates detection while the latter is favorable in the presence of noise since the average distance to the quantization

threshold is maximized. In order to obtain a symmetric constellation diagram, the number of constellation points n_{const} needs to be 1, 2, or a multiple of four, such that there can exist a ϕ_0 yielding said symmetry. From (5.3), (5.4), and (5.2) one can see that the difference in phase rotation caused by any two symbols is an integer multiple of $2\pi h_{\text{cpm}}$. Thus, $2\pi h_{\text{cpm}}$ is the minimum angle between two constellation points. Assuming $h_{\text{cpm}} \leq 1$, we therefore have[1]

$$n_{\text{const}} = \frac{1}{h_{\text{cpm}}} = 2^{c_1},\ c_1 \in \mathbb{N}_0 \tag{5.18}$$

constellation points on the unit circle. For a symmetric constellation diagram we then need

$$\phi_0 = \frac{\pi/2}{2\lceil n_{\text{const}}/4 \rceil} = \begin{cases} \frac{\pi}{4} & h_{\text{cpm}} > \frac{1}{4} \\ \pi h_{\text{cpm}} & h_{\text{cpm}} \leq \frac{1}{4} \end{cases} \tag{5.19}$$

where $\lceil \cdot \rceil$ is the ceiling function. One common configuration fulfilling this is that M_{cpm} is a multiple of four, $h_{\text{cpm}} = 1/M_{\text{cpm}}$, and $\phi_0 = \frac{\pi}{M_{\text{cpm}}}$. Furthermore, we assume $L_{\text{cpm}} = 1$.

The signal-to-noise ratio (SNR) is defined using the energy of the samples after the receive filter according to [Lap17, Theorem 8.4.3] as

$$\text{SNR} = \frac{T_s \sum_{k=0}^{K-1} \sum_{m=1}^{M_{\text{osr}}} |\mathsf{r}_{k,m}|^2}{T_s \sum_{k=0}^{K-1} \sum_{m=1}^{M_{\text{osr}}} |\hat{\mathsf{n}}_{k,m}|^2} = \frac{\sum_{k=0}^{K-1} \sum_{m=1}^{M_{\text{osr}}} |\mathsf{r}_{k,m}|^2}{\sum_{k=0}^{K-1} \sum_{m=1}^{M_{\text{osr}}} |\hat{\mathsf{n}}_{k,m}|^2}. \tag{5.20}$$

5.3 Maximum Achievable Rate in a Noise-free Scenario and Intermediate Frequencies

In this section, we aim to understand the limits of oversampled CPM under 1-bit quantization in an ideal transmission scenario, i.e., without any noise and filter distortion. This is equivalent to $h\left(n\frac{T_{\text{symb}}}{M_{\text{osr}}}\right) = \delta[n]$, $n \in \mathbb{Z}$ and $\hat{\mathsf{n}}_{k,m} = 0$ for all $k = 0, ..., K-1$ and $m = 1, ..., M_{\text{osr}}$, i.e.,

$$\mathsf{r}(t) = \mathsf{x}(t). \tag{5.21}$$

In contrast to classic baseband PSK with $M_{\text{cpm}} > 4$, with CPM the continuous phase change over time can be tracked by oversampling the received signal, even when the amplitude information is discarded due to 1-bit quantization. With the additional observations it is possible to more accurately determine the speed of the phase change, which in turn corresponds to a transmit symbol α_k. We thus examine the number of

[1] For $h_{\text{cpm}} > 1$ the denominator would need to be $\mod(h_{\text{cpm}}, 1)$, which complicates notation unnecessarily since this case is of very limited relevance here due to bandwidth efficiency reasons (c.f. Section 5.1).

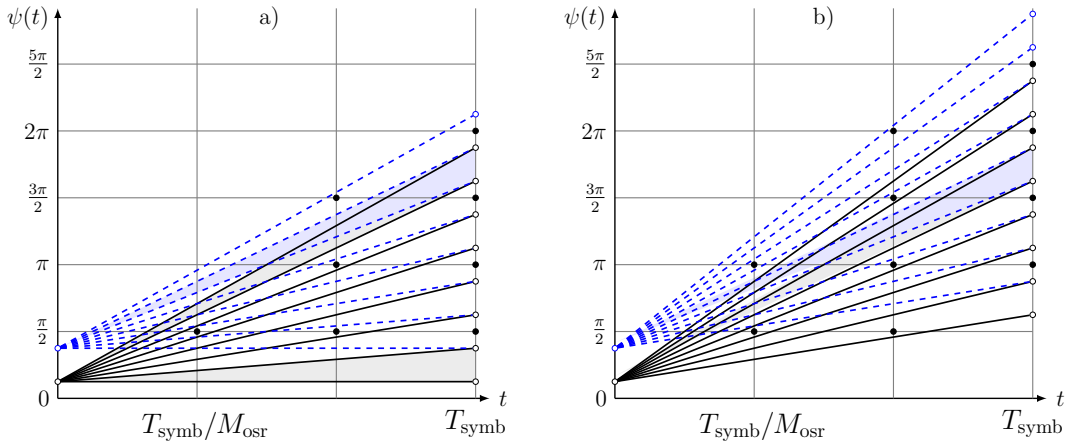

Fig. 5.5. State transitions for $\phi_0 = {}^{\pi}/_{M_{\text{cpm}}}$ (black) and $\phi_0 = {}^{3\pi}/_{M_{\text{cpm}}}$ (blue) with $M_{\text{cpm}} = 8$ and $M_{\text{osr}} = 3$ for a) $f_{\text{IF}} = \Delta f$ and b) $f_{\text{IF}} = \Delta f + {}^{1}/_{(4T)}$

resolvable symbols (or paths in the trellis) N_d that can be distinguished despite 1-bit quantization.

In order to illustrate the problem, Fig. 5.5 depicts phase transitions (corresponding to transmitted symbols) of $\psi(t)$ for an 8-CPFSK constellation on a 3-fold oversampled grid. Two phase transitions are distinguishable if they produce a different sequence of samples, i.e., if at least one sample differs after 1-bit quantization. Fig. 5.5a) shows the phase trajectories if ϕ_0 is in the first quadrant and depending on ϕ_0 either 6 or 7 symbols can be distinguished. For distinguishable trajectories one of the grid points of the oversampled grid, marked with black dots, has to be in between the lines representing the phase trajectories. If not, the two trajectories are ambiguous, which is marked by filled areas. Due to the symmetry of the constellation diagram w.r.t. the quadrants, it is sufficient to consider ϕ_0 to be in the first quadrant. Therefore, in this scenario for i.u.d. input symbols we can distinguish on average 6.5 symbols, i.e., $\log_2 6.5 \approx 2.7$ bits per channel use (bpcu). By increasing M_{osr}, we would be able to resolve the blue and the upper gray ambiguity, however, the lower gray ambiguity remains as both trajectories are in the same quadrant.

Recall that $\psi(t)$ is obtained from $\phi(t)$ by adding a frequency offset Δf. Thus, in contrast to $e^{j\phi(t)}$, the signal $e^{j\psi(t)}$ is not a baseband signal anymore but its center frequency is at an intermediate frequency (IF). Increasing this IF as depicted in Fig. 5.5b) provides a solution to the remaining ambiguity in the first quadrant, which yields $\log_2 7 \approx 2.8$ bpcu. Thus, the parameters M_{osr} and f_{IF} determine the number of distinguishable phase transitions N_d[2].

[2]Note that for a given oversampled grid like the one in Fig. 5.5, even more paths can be distinguished if ϕ_0, h_{cpm}, and the phase changes $\pi h_{\text{cpm}} \alpha_k$ can be chosen freely. However, in general this yields the already mentioned disadvantages of an unsymmetrical constellation diagrams and a time invariant trellis.

For analyzing the impact of the IF, we adopt the notations

$$\psi(t) \quad = \phi(t) + 2\pi f_{\mathrm{IF}} t \tag{5.22}$$

and

$$f_{\mathrm{IF}} \quad = \Delta f + \tfrac{n_{\mathrm{IF}}}{T_{\mathrm{symb}}} \tag{5.23}$$

where we restrict n_{IF} to

$$n_{\mathrm{IF}} = c_2 h_{\mathrm{cpm}},\ c_2 \in \mathbb{Z}. \tag{5.24}$$

This retains the property of a time invariant trellis since the additional phase turn is always an integer multiple of the angle between two constellation points. From the discussion on Fig. 5.5a) we observed that in order to resolve all ambiguities given a sufficiently high oversampling factor, it must be avoided that more than one transition starts and ends in the same quadrant. Since Δf is already positive, we consider phase changes counterclockwise, i.e., $n_{\mathrm{IF}} > 0$. With this and the assumptions in Section 5.2, in Appendix C.1 the minimum IF for CPM to resolve all transitions in the trellis despite 1-bit quantization is shown to be

$$n_{\mathrm{IF,min}} = h_{\mathrm{cpm}} \left(\left\lceil \frac{1}{4h_{\mathrm{cpm}}} \right\rceil - 1 \right) = h_{\mathrm{cpm}} c_{2,\mathrm{min}}. \tag{5.25}$$

This is shown in Fig. 5.6 for $M_{\mathrm{cpm}} = 8$ and CPFSK. Different parameters $h_{\mathrm{cpm}} = \{1/4, 1/8, 1/16\}$ correspond to different $h_{\mathrm{cpm}} M_{\mathrm{cpm}} = \{2, 1, 0.5\}$ and, thus, different bandwidths, cf. Fig. 5.2. The minimum IFs according to (5.25) are 0, 0.125, and 0.1875, respectively, i.e., $c_2 = \{0, 1, 3\}$. This is confirmed in parts a) and b) of Fig. 5.6, where the number of resolvable bits is given for different oversampling factors and IFs. The corresponding required oversampling factors to resolve $\log_2(M_{\mathrm{cpm}})$ bits are 2, 5 and 13, respectively. Note that for $h_{\mathrm{cpm}} M_{\mathrm{cpm}} = 0.5$ in Fig. 5.6b) $\log_2(M_{\mathrm{cpm}})$ bits can be already resolved for 11-fold oversampling using an IF larger than $n_{\mathrm{IF,min}}$.

Furthermore, Fig. 5.6a) shows that for a given M_{osr}, e.g., $M_{\mathrm{osr}} = 5$, a larger n_{IF} is not necessarily better. Indeed, while for $c_2 = \{1, 2\}$ all paths can be resolved, this is not the case for $c_2 > 2$. In other words, for a given oversampling ratio (OSR) the phase changes can be too fast for the 1-bit quantizer, which then has to be resolved by increasing M_{osr}.

For CPFSK we investigate in Appendix C.2 an upper bound $n_{\mathrm{IF\,max}}$ on the interval $[n_{\mathrm{IF\,min}}, n_{\mathrm{IF\,max}}]$, which can serve as a restriction of the searching space when looking for reasonable choices of n_{IF}. Thus, $n_{\mathrm{IF\,max}}$ represents an IF beyond which increasing n_{IF} further has no additional benefit in terms of the number of resolvable bits. The derivation is based on the assumption that when the phase rotates through a quadrant without a sample being generated, the distinction of the paths in the trellis becomes

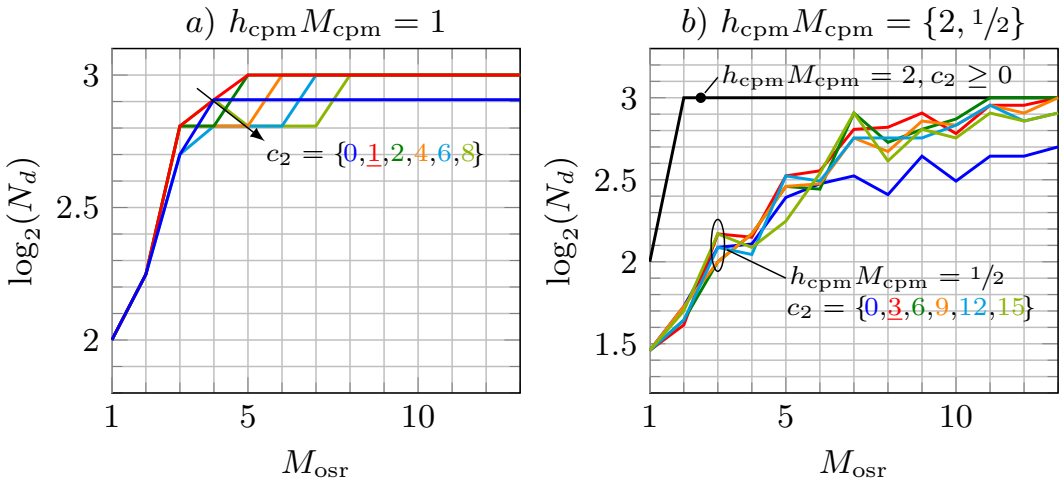

Fig. 5.6. Number of bits $\log_2(N_d)$ resolvable for CPFSK and $M_{\mathrm{cpm}} = 8$ and $h_{\mathrm{cpm}} = \{1/4, 1/8, 1/16\}$, i.e., $h_{\mathrm{cpm}} M_{\mathrm{cpm}} = \{2, 1, 0.5\}$

more difficult for the 1-bit quantizer. Given that every symbol in $\mathcal{A}$ is associated with a frequency f_i, this yields for the minimum and maximum values of α_i in $\mathcal{A}$, cf. (5.3), two possible upper bounds on n_{IF} as

$$\Delta\psi = 2\pi f_i \frac{T_{\mathrm{symb}}}{M_{\mathrm{osr}}} < \frac{\pi}{2} \tag{5.26}$$

$$\begin{aligned} n_{\mathrm{IF}} &< \frac{M_{\mathrm{osr}}}{4} - h_{\mathrm{cpm}} \left(\frac{\alpha_i + M_{\mathrm{cpm}} - 1}{2} \right) \\ &= \begin{cases} \frac{M_{\mathrm{osr}}}{4} = n_{\mathrm{IF,max,1}}, & \alpha_i = -(M_{\mathrm{cpm}} - 1) \\ \frac{M_{\mathrm{osr}}}{4} - \frac{h_{\mathrm{cpm}}(M_{\mathrm{cpm}}-1)}{2} = n_{\mathrm{IF,max,2}}, & \alpha_i = M_{\mathrm{cpm}} - 1 \end{cases} . \end{aligned} \tag{5.27}$$

We evaluated the above expression for different OSRs and compared the results to the number of resolvable paths for different M_{osr} and n_{IF}, cf. Appendix C.2. We find that for all non-trivial cases the performance is periodic over n_{IF} and that the period increases with M_{osr}. Interestingly, the performance in the interval $[n_{\mathrm{IF,max,2}}, n_{\mathrm{IF,max,1}}]$ is particularly bad, which leads us to the conclusion that our initial assumption was correct and that passing a quadrant without sampling – even for just one out of the M_{cpm} symbols, i.e., $\max_{\alpha_i}(2\pi f_i \frac{T_{\mathrm{symb}}}{M_{\mathrm{osr}}}) \geq \frac{\pi}{2}$ – represents a significant loss of information. We thus conclude that for CPFSK a valid description for the desired searching space is

$$n_{\mathrm{IF,min}} \leq n_{\mathrm{IF,CPFSK}} < n_{\mathrm{IF,max,2}} \tag{5.28}$$

which, however, does not mean that below $n_{\mathrm{IF,max,2}}$ all symbols can be resolved but that there is an IF below n_{IF} that achieves at least the same $\log_2(N_d)$ bits as any IF above. This and the low OSRs used also explain the limited benefit observed by increasing n_{IF} to one in [LDdF18].

The minimum OSR to resolve all paths $M_{\text{osr,min}}$ can be found by a simple search. However, for CPFSK, also a closed-form expression can be given. Then, the minimum OSR to distinguish all symbols can be given based on the set

$$\mathcal{B}_{p,s,n,(i,j)} = \left\{ m \middle| m \in \mathbb{N} \wedge m \in \left[\left\lceil \frac{2s\tilde{f}_i}{\frac{n}{2} - \frac{\phi_{0,p}}{\pi}} + \epsilon \right\rceil, \left\lfloor \frac{2s\tilde{f}_j}{\frac{n}{2} - \frac{\phi_{0,p}}{\pi}} \right\rfloor \right] \wedge m \geq s \right\} \tag{5.29}$$

$$\tilde{f}_i = h_{\text{cpm}} \left(\frac{\alpha_i + M_{\text{cpm}} - 1}{2} + c_2 \right) \tag{5.30}$$

$$\phi_{0,p} = \phi_0 + 2\pi(p-1)h_{\text{cpm}},\ p = 1, 2, ..., \lceil n_{\text{const}}/4 \rceil \tag{5.31}$$

which contains all OSRs that allow distinction between the ith and the jth symbol in $\mathcal{A}$ at the sth sample using a threshold n. Hereby, $i < j$ such that the symbols are ordered w.r.t. increasing frequency and $m \geq s$ since the distinguishing sample must be smaller then the OSR. Furthermore, n represents a multiple of $\pi/2$, i.e., the border between two quadrants, ϵ is an arbitrary small positive number, $p \geq 1$ is an integer referring to the pth constellation point in the quadrant, and $\lfloor \cdot \rfloor$ is the floor function. For more details on (5.29) to (5.34) refer to Appendix C.3. Denoting as $\text{count}_m(\cdot)$ the number of occurrences of element m in a set, the set

$$\mathcal{B}_{p,(i,j)} = \bigcup_{s=1}^{\infty} \left(\bigcup_{n=1}^{n_{\text{max}}} \mathcal{B}_{p,s,n,(i,j)} \setminus \left\{ m \middle| \text{count}_m \left(\bigcup_{n=1}^{n_{\text{max}}} \mathcal{B}_{p,s,n,(i,j)} \right) \bmod 4 = 0 \right\} \right) \tag{5.32}$$

$$n_{\text{max}} = \left\lfloor \frac{1}{\pi/2} \left(\phi_0 + 2\pi \left(\lceil n_{\text{const}}/4 \rceil - 1 \right) h_{\text{cpm}} + 2\pi f_{M_{\text{cpm}}} T_{\text{symb}} \right) \right\rfloor \tag{5.33}$$

contains all OSRs that can distinguish between the ith and the jth symbol when starting at $\phi_{0,p}$ and

$$M_{\text{osr,min}} = \min \left(\bigcap_{p=1}^{n_{\text{const}}/4} \left(\bigcap_{(i,j),i<j} \mathcal{B}_{p,(i,j)} \right) \right) \tag{5.34}$$

is the minimum OSR that allows to resolve all symbols for all starting points in the constellation diagram. Fig. 5.7 depicts the results of (5.34) for different modulation orders M_{cpm} and modulation indices h_{cpm} given that $n_{\text{IF}} \geq n_{\text{IF,min}}$. We observe that with increasing OSR higher modulation orders can be resolved, e.g., 8-CPFSK with 5-fold oversampling and 16-CPFSK with 14-fold oversampling at $n_{\text{IF,min}}$ for $h_{\text{cpm}} M_{\text{cpm}} = 1$. When the bandwidth is small ($h_{\text{cpm}} M_{\text{cpm}} = 1/2$), higher OSRs are required while a larger bandwidth can decrease the required OSR. However, the latter does not hold always as can be seen for $M_{\text{cpm}} = 4$ and $h_{\text{cpm}} M_{\text{cpm}} = 2$. In this case the frequencies associated with the symbols become so large that, due to the wrapping of the phase with modulo 2π, at least two-fold oversampling is required to resolve all symbols. Note, that in the presence of noise odd oversampling factors are to be

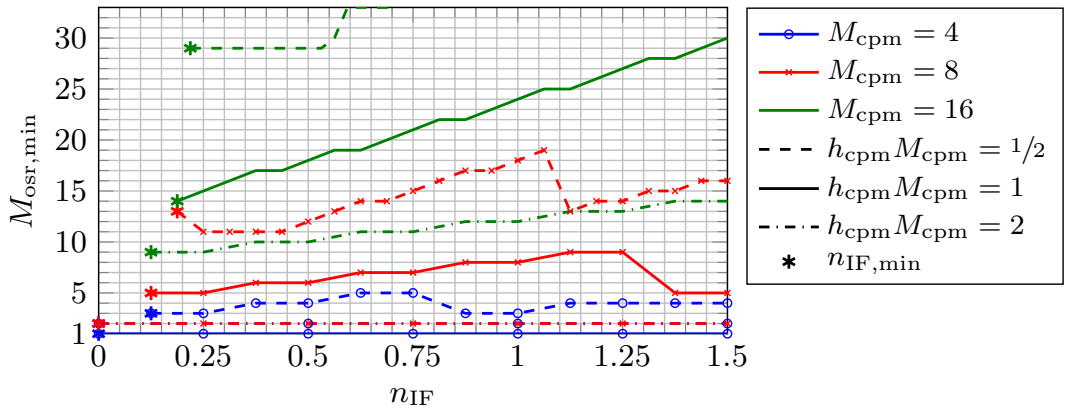

Fig. 5.7. Minimum oversampling ratio of CPFSK to resolve M_{cpm} symbols according to (5.34) depending on the intermediate frequency

preferred as this, due to the symmetry of the constellation diagram, avoids samples on the decision thresholds of the quantizers.

For waveforms other than CPFSK, the number of resolvable bits can be obtained by a simple search. Fig. 5.8 shows the number of resolvable bits for modulation orders 4, 8, and 16 for PSK, CPFSK and a sine shaped transition waveform using the IF $n_{\mathrm{IF\,min}}$ obtained for CPM in (5.25). In principle, the considerations leading to (5.25) for CPM also apply to PSK, whose phase trajectories are linear with a slope of $2\pi f_{\mathrm{IF}}$ and an offset of height $\pi h_{\mathrm{cpm}} \alpha_k$. However, in Fig. 5.8c) we see that – in contrast to CPM – for PSK $n_{\mathrm{IF\,min}}$ this is not the only criterion for resolving all symbols. The problem is again the wrapping of the phase with modulo 2π. Since the slope of the phase trajectory only depends on f_{IF} and the information is only encoded in the offset, for $h_{\mathrm{cpm}} M_{\mathrm{cpm}} = 2$ trajectories for different symbols overlap and become indistinguishable. In other cases PSK can reduce the required OSR to resolve all bits, cf. Fig. 5.8a) and b), however, it is not continuous in phase, which means increased side-lobes in the spectrum compared to CPM. The difference between the sine shaped transition and CPFSK is marginal. We therefore focus on CPFSK in the next sections due to the better spectral characteristics given the smaller variations of instantaneous frequencies in $q'(t)$. Looking at the results for different $h_{\mathrm{cpm}} M_{\mathrm{cpm}}$, the interesting question is, given the spectra in Fig. 5.2a), which configuration provides more spectral efficiency while remaining robust under filtering and noise. This question shall be addressed in the next sections.

5.4 Noisy Analysis

Due to oversampling and FTN signaling, the impact of intersymbol interference (ISI) has to be modeled besides the AWGN in order to characterize the BER-performance.

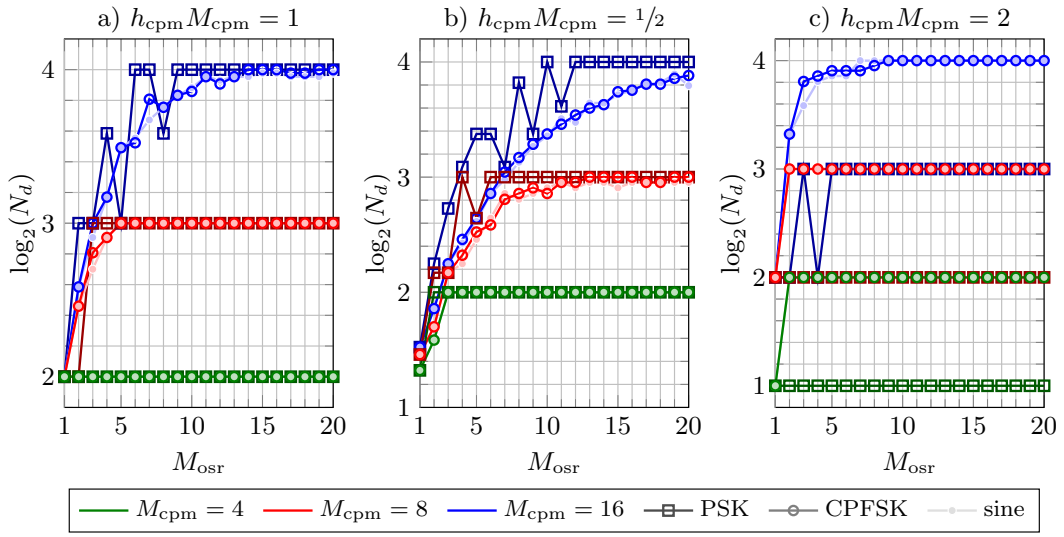

Fig. 5.8. Number of bits $\log_2(N_d)$ resolvable for CPFSK,PSK and sine-shaped transition, $M_{\mathrm{cpm}} = \{4, 8, 16\}$, $h_{\mathrm{cpm}}M_{\mathrm{cpm}} = \{2, 1, 0.5\}$ and $n_{\mathrm{IF}} = n_{\mathrm{IF\,min}}$

CPM is already a trellis based modulation, where the trellis only has to be extended based on the channel filters. In this section, the conditional probability density function (pdf) of the channel is given, which is required for the Bahl-Cocke-Jelinek-Raviv-algorithm (BCJR). Furthermore, a simplified Viterbi-based decoder is proposed.

5.4.1 Conditional Probabilities of the Received Signal

With oversampling the complex receive vector that corresponds to the kth symbol α_k is given by $\mathbf{r}_k = [\mathsf{r}_{k,1}, ..., \mathsf{r}_{k,M_{\mathrm{osr}}}]$. Since oversampling and FTN signaling are used, noise correlation has to be considered. With $\check{\mathbf{r}}_k = [\Re\{\mathbf{r}_k\}, \Im\{\mathbf{r}_k\}]^T$ being the stacked vector of real and imaginary part of dimension $2M_{\mathrm{osr}} \times 1$, we can express the conditional probability of the received given the transmitted vector via the multivariate Gaussian as distribution

$$p_{\check{\mathbf{r}}_k|\check{\mathbf{x}}_k}(\check{\boldsymbol{r}}_k|\check{\boldsymbol{x}}_k) = \frac{e^{-\frac{1}{2}(\check{r}_k - \check{x}_k)^T \mathbf{R}_{\check{\mathbf{n}}}^{-1}(\check{r}_k - \check{x}_k)}}{\sqrt{(2\pi)^{2M_{\mathrm{osr}}}|\mathbf{R}_{\check{\mathbf{n}}}|}} \tag{5.35}$$

where $\check{\mathbf{x}}_k$ is the expected receive signal in the absence of noise and $\mathbf{R}_{\check{\mathbf{n}}}$ is covariance matrix of the oversampled noise vector $\check{\mathbf{n}}_k = [\Re\{\hat{\mathbf{n}}_k\}, \Im\{\hat{\mathbf{n}}_k\}]^T$. The expected receive signal is given by

$$\check{\mathbf{x}}_k = \begin{bmatrix} \Re\{\hat{\mathbf{x}}_k\} \\ \Im\{\hat{\mathbf{x}}_k^T\} \end{bmatrix} = \begin{bmatrix} \Re\{\boldsymbol{H}(\mathbf{x}_{k-L_h}^k)\} \\ \Im\{\boldsymbol{H}(\mathbf{x}_{k-L_h}^k)^T\} \end{bmatrix}. \tag{5.36}$$

Here, $\boldsymbol{H}$ is a $M_{\mathrm{osr}} \times M_{\mathrm{osr}}(L_h + 1)$ Toeplitz matrix, whose first row is $[\boldsymbol{h}, \mathbf{0}_{M_{\mathrm{osr}}-1}]$ using the oversampled channel filter $\boldsymbol{h} = [h(-\frac{L_h M_{\mathrm{osr}} T_s}{2}), ..., h(0), h(T_s), ..., h(\frac{L_h M_{\mathrm{osr}} T_s}{2})]$ of size $1 \times M_{\mathrm{osr}} L_h + 1$. Furthermore, L_h is the length of $h(t)$ w.r.t. the symbols duration

and $\mathbf{0}_a$ is an all-zero row vector of length a. Thus, the probability that the complex 1-bit quantized observation $\boldsymbol{y}_i$ is received can be given by integrating over (5.35) as

$$\Pr(\mathbf{y}_k = \boldsymbol{y}_i | \check{\mathbf{x}}_k = \check{\boldsymbol{x}}_i) = \int_{\mathbb{Y}_i^{M_{\mathrm{osr}}}} \cdots \int p_{\check{\mathbf{r}}_k|\check{\mathbf{x}}_k}(\check{\boldsymbol{r}}|\check{\boldsymbol{x}}_k) d\check{\boldsymbol{r}} \tag{5.37}$$

where $\mathbb{Y}_i^{M_{\mathrm{osr}}}$ denotes the quantization region to which $\boldsymbol{y}_i$ belongs. Note that real and imaginary part are correlated since the filters $v_{\mathrm{tx}}(t)$ and $v_{\mathrm{rx}}(t)$ have their center frequency at f_{IF}. Thus, we cannot compute (5.35) separately for real and imaginary part. Therefore, $\mathbf{R}_{\check{\mathbf{n}}}$ is of dimension $2M_{\mathrm{osr}} \times 2M_{\mathrm{osr}}$ and is is given by

$$\mathbf{R}_{\check{\mathbf{n}}} = \mathbb{E}\left\{ \begin{bmatrix} \Re\{\boldsymbol{V}\} & -\Im\{\boldsymbol{V}\} \\ \Im\{\boldsymbol{V}\} & \Re\{\boldsymbol{V}\} \end{bmatrix} \begin{bmatrix} \Re\{(\mathbf{n}_{k-L_v}^k)^T\} \\ \Im\{(\mathbf{n}_{k-L_v}^k)^T\} \end{bmatrix} \begin{bmatrix} \Re\{(\mathbf{n}_{k-L_v}^k)^T\} \\ \Im\{(\mathbf{n}_{k-L_v}^k)^T\} \end{bmatrix}^T \begin{bmatrix} \Re\{\boldsymbol{V}\} & -\Im\{\boldsymbol{V}\} \\ \Im\{\boldsymbol{V}\} & \Re\{\boldsymbol{V}\} \end{bmatrix}^T \right\}. \tag{5.38}$$

Here, $\boldsymbol{V}$ is an $M_{\mathrm{osr}} \times (L_v+1)M_{\mathrm{osr}}$ Toeplitz matrix, whose first row is $[\boldsymbol{v}_{\mathrm{rx}}, \mathbf{0}_{M_{\mathrm{osr}}-1}]$ using the oversampled receive filter $\boldsymbol{v}_{\mathrm{rx}} = [v_{\mathrm{rx}}(-\frac{L_v M_{\mathrm{osr}} T_s}{2}), ..., v_{\mathrm{rx}}(0), v_{\mathrm{rx}}(T_s), ..., v_{\mathrm{rx}}(\frac{L_h M_{\mathrm{osr}} T_s}{2})]$ of size $1 \times M_{\mathrm{osr}} L_v + 1$, where L_v is the length of $v_{\mathrm{rx}}(t)$ in terms of the symbol duration. The integration in (5.37) can be evaluated by a quasi-Monte Carlo integration algorithm as, e.g., implemented in the MATLAB-function mvncdf$(\cdot)$. The conditional probabilities can then be used in a trellis based detection algorithm as for example the BCJR. However, as mentioned before, the computational effort required is very high. The number of evaluations n_{ev} of (5.37) depends on the length of the phase response L_{cpm}, the number of constellation points n_{const}, the filter length L_h, the alphabet size $|\mathcal{A}| = M_{\mathrm{cpm}}$, and the OSR M_{osr}. It is given by

$$n_{\mathrm{ev}} = 4^{M_{\mathrm{osr}}} \frac{n_{\mathrm{const}}}{4} {M_{\mathrm{cpm}}}^{L_h + L_{\mathrm{cpm}}} \tag{5.39}$$

where $4^{M_{\mathrm{osr}}}$ is the number of possibly observed complex vectors $\mathbf{y}_i$. Furthermore, the factor of $1/4$ represents the symmetry w.r.t. the quadrants, due to which the evaluation for the constellation points of the first quadrant is sufficient. With the parameters $M_{\mathrm{cpm}} = 8$, $M_{\mathrm{osr}} = 5$, $L_{\mathrm{cpm}} = 1$, and for a short filter length $L_h = 2$, this already yields $1024^2 \approx 10^6$ evaluations of a 10-dimensional multivariate integral. This clearly becomes infeasible for higher OSRs. Therefore, this method is suitable for small temporal OSRs or for low modulations orders M_{cpm}. In other cases, however, it is reasonable to consider a simplified decision criterion.

5.4.2 Simplified Decision Metric

The phase of the oversampled receive vector is $\theta_k = [\theta_{k,1}, ..., \theta_{k,M_{\mathrm{osr}}}] = \arg(\mathbf{y}_i)$ with $\theta_{k,m} \in \{\frac{\pi}{4}, \frac{3\pi}{4}, \frac{5\pi}{4}, \frac{7\pi}{4}\}$. It represent the sequence of quadrants in which the signal is observed at the sampling instants. A simplified metric needs to capture the periodicity

of the phase and it has to be able to operate on the 1-bit quantized samples. For example, we can choose the following mapping $\mathcal{M}$ inspired by the concept of Hamming distance

$$\tilde{\theta}_{k,m} = \mathcal{M}(\theta_{k,m}) = \begin{cases} [0,0], & \text{if } \theta_{k,m} = \frac{\pi}{4} \\ [0,1], & \text{if } \theta_{k,m} = \frac{3\pi}{4} \\ [1,1], & \text{if } \theta_{k,m} = \frac{5\pi}{4} \\ [1,0], & \text{if } \theta_{k,m} = \frac{7\pi}{4} \end{cases} \tag{5.40}$$

where $\tilde{\theta}_k = [\tilde{\theta}_{k,1}, ..., \tilde{\theta}_{k,M_{\text{osr}}}]$ is of size $1 \times 2M_{\text{osr}}$. Then, the phases of the received vector at time k can be compared with the possible transmit vector $\hat{\boldsymbol{x}}_i$ belonging to the ith symbol α_i in $\mathcal{A}$ using the Hamming distance

$$d(\mathbf{y}_k, \alpha_i) = d_H \left(\mathcal{M}(\boldsymbol{\theta}_k), \mathcal{M}(\arg(Q_1(\hat{\boldsymbol{x}}_i^T))) \right). \tag{5.41}$$

This is straight-forward to implemented in a trellis based algorithm, e.g., the Viterbi-algorithm, using the XOR-operation.

5.5 Faster-than-Nyquist Phase Modulation

From Section 5.3 we took away the question how to most efficiently configure an oversampled 1-bit quantized phase modulated system in the presence of noise and channel filters. Could a 4-CPFSK modulation scheme with $h_{\text{cpm}} M_{\text{cpm}} = 1/2$ operate under a channel filter channel filter using half the bandwidth and, thus, increasing the spectral efficiency by 33 % compared to 8-CPFSK with $h_{\text{cpm}} M_{\text{cpm}} = 1$? Or is the benefit marginal as decreasing h_{cpm} does not decrease the main lobe of the spectrum sufficiently, cf. Fig. 5.2? From [BDF19] we know that the temporally oversampled CPFSK setup is very sensitive to noise and filter distortion. This is due to the fact, that the samples that distinguish between two phase trajectories are naturally close to the decision threshold. One option to resolve this problem could be using discontinuous, e.g., PSK based, phase smoothing responses, however, this increases the side lobes and the advantage we saw in Section 5.3 could be mitigated.

Another option would be to maximize the distance to the quantization threshold by allowing the constellation points to take on only values of $\{\frac{\pi}{4}, \frac{3\pi}{4}, \frac{5\pi}{4}, \frac{7\pi}{4}\}$ and making sure that the signal is always sampled at these points. In order to mitigate losses in spectral efficiency, the concept of FTN signaling can be applied to the CPM signal. Traditionally, FTN signaling is used with minimum phase filters [ARO13] in order to use reduced search methods in the trellis. This is not ideal for phase modulation since minimum phase filters do not have a linear phase response. However, in [BDF19] we

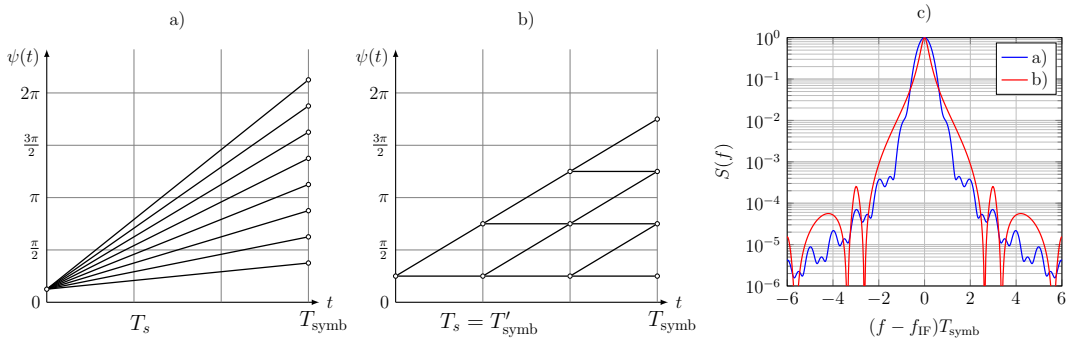

Fig. 5.9. Concept of FTN-CPM for $n_{\text{IF}} = n_{\text{IF,min}}$. Phase transitions for a) 8-CPFSK, $h_{\text{cpm}}M_{\text{cpm}} = 1$, b) 2-CPFSK, $h_{\text{cpm}}M_{\text{cpm}} = 1/2$ with reduced symbol duration T'_{symb}, and c) comparison of the signal spectra of a) and b)

proposed an adapted waveform for 1-bit quantized CPM based on shorter 2-CPFSK symbols with $h_{\text{cpm}} = 1/4$ using a filter with a linear phase response. This drastically improves the BER performance due to the increased distance to the quantization threshold. More generally described, this can be seen as a form of coded FTN 2-CPFSK. This idea was further studied in [dAL19] using uncoded FTN 2-CPFSK with $h_{\text{cpm}} = 1/4$ and a wideband channel filter. The BER performance was shown to be reasonable as well with a BCJR sequence decoder as with a simple differential decoder.

5.5.1 Uncoded FTN Phase Modulation

There are two possibilities to model FTN transmission. One is depicted in Fig. 5.9 for $n_{\text{IF}} = n_{\text{IF,min}}$. Part a) shows the phase transitions of an 8-CPFSK symbol with $h_{\text{cpm}}M_{\text{cpm}} = 1$. From Section 5.3 we know that this can be resolved under 5-fold temporal oversampling. Part b) depicts 2-CPFSK symbols with $h_{\text{cpm}} = 1/4$, i.e., $h_{\text{cpm}}M_{\text{cpm}} = 1/2$, with a reduced symbol duration T'_{symb}, which ideally only requires 3-fold oversampling w.r.t. T_{symb} to transmit 3 bits. However, in part c) of the figure the wider spectrum can be observed as already suspected above. The other option is using $T_{\text{symb}} = T'_{\text{symb}}$ but to reduce the bandwidth $2W$ of the RC-filter available for transmission.

In any case, we have for the FTN factor

$$M_{\text{FTN}} = \frac{1}{2WT'_{\text{symb}}}. \tag{5.42}$$

Furthermore, since the objective is to sample at maximum distance to the quantization thresholds, i.e., at odd integer-multiples of $\pi/4$, the temporal oversampling factor w.r.t. the symbol duration is

$$M_{\text{osr}} = M_{\text{osr},t} = \frac{T'_{\text{symb}}}{T_s} = 1 \tag{5.43}$$

while the oversampling factor w.r.t. the system bandwidth is

$$M_{\mathrm{osr},f} = \frac{f_s}{2W} = \frac{1}{2WT_s} = \frac{1}{2WT'_{\mathrm{symb}}} = M_{\mathrm{FTN}}. \tag{5.44}$$

If all possible symbol combinations, i.e., all trajectories in Fig. 5.9b), are allowed, the maximum attainable spectral efficiency is

$$\mathrm{SE}_{\max} = \frac{\log_2(M_{\mathrm{cpm}})}{2WT'_{\mathrm{symb}}}. \tag{5.45}$$

This can be understood as uncoded FTN signaling, which was applied in [dAL19].

5.5.2 Coded FTN Phase Modulation

In contrast to [dAL19], the approach [BDF19] can be seen as coded FTN transmission. Coded FTN transmission means the selection of a subset $\mathcal{S}$ of the FTN signals [ARO13] such that the α_k are no longer data symbols. One information symbol will be encoded in a certain subsequence $\boldsymbol{\alpha}_l$. In order to transmit the same number of bits compared to uncoded FTN within T_{symb}, T'_{symb} has to be further reduced. For example, if ${}^{T_{\mathrm{symb}}}/_{T'_{\mathrm{symb}}} = 5$, then 8 out of the $2^5 = 32$ possible FTN signals had to be selected in order to transmit $\log_2(8) = 3$ bit. The temporal OSR w.r.t. to an FTN symbol α_k remains one, cf. (5.43), however, the information symbol is now encoded in $\boldsymbol{\alpha}_l$ of duration T_{symb} such that

$$M_{\mathrm{osr}} = \frac{T_{\mathrm{symb}}}{T'_{\mathrm{symb}}} \tag{5.46}$$

with $l = \lfloor {}^{k}/_{M_{\mathrm{osr}}} \rfloor$ and $\boldsymbol{\alpha}_l$ is of length M_{osr}. Fig. 5.10 illustrates the mapping from α_k to $\boldsymbol{\alpha}_l$ using the example from [BDF19] for $M_{\mathrm{osr}} = 5$ and 8 code sequences, i.e., 3 bit.

Criteria for designing a code can be manifold. For example, one could maximize the number of trajectories $|\mathcal{S}|$ given a minimum distance in terms of differing samples between the trajectories. The criterion used in [BDF19] is that trajectories connecting the same constellation points should be separated by the maximum possible number of samples. The systematic design of good codes for coded FTN-CPM has remained out of scope of this work. It would be interesting to see how known tools of coding theory could be applied to this problem. For example, concatenated codes and generalized concatenated codes (GCCs), see for example [ZSB99] for an introduction, have been applied successfully in order to derive communication systems with low decoding complexity and good error-control capabilities. Since the inner code can also be considered to be the modulation scheme, i.e., CPFSK in our case, the outer code would select the valid subsequences for coded FTN transmission. For example, in [HBSU98], GCCs have been successfully applied to reduce the BER by designing outer

a)

i	$\boldsymbol{\alpha}_i$
0	$[-1,-1,-1,-1,\ 1]$
1	$[\ 1,-1,-1,-1,-1]$
2	$[-1,-1,-1,\ 1,\ 1]$
3	$[\ 1,\ 1,-1,-1,-1]$
4	$[-1,-1,\ 1,\ 1,\ 1]$
5	$[\ 1,\ 1,\ 1,-1,-1]$
6	$[-1,\ 1,\ 1,\ 1,\ 1]$
7	$[\ 1,\ 1,\ 1,\ 1,-1]$

b)

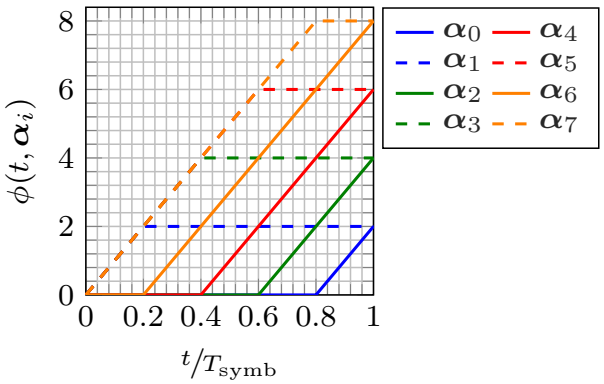

Fig. 5.10. Coded FTN example for $M_{\text{osr}} = 5$: a) mapping between $\alpha_k \in \mathcal{A} = \{-1, 1\}$ and $\boldsymbol{\alpha}_l$ for the symbol indices $i = 0, ..., |\mathcal{S}| - 1$ ans b) resulting phase trajectories

convolutional codes for CPFSK transmission over an unquantized AWGN channel with Nyquist sampling.

The maximum attainable spectral efficiency for coded FTN signaling is given as

$$\text{SE}_{\max} = \frac{\log_2(|\mathcal{S}|)}{2WT_{\text{symb}}}. \tag{5.47}$$

while

$$M_{\text{FTN}} = \frac{1}{2WT'_{\text{symb}}}. \tag{5.48}$$

5.6 Numerical Results

For performance evaluation we apply the decoders derived in Section 5.4. We evaluate the BER for temporally oversampled phase modulation using Nyquist signaling and $h_{\text{cpm}}M_{\text{cpm}} = 1$ as well as coded and uncoded FTN transmission using 2-CPFSK with $h_{\text{cpm}}M_{\text{cpm}} = 1/2$. This corresponds to the configurations shown in Fig. 5.9. For all simulations we have $n_{\text{IF}} = n_{\text{IF,min}}$, $f_{\text{IF}} = \Delta f + \frac{n_{\text{IF}}}{T'_{\text{symb}}}$ and use truncated RC- and RRC-filters with $L_h = L_v = 10$. In order to limit the complexity of the decoder and the number of states in the trellis, we use a shorter truncated version of $h(t)$ with $L_h = 2$ for the computation of the $\hat{\boldsymbol{x}}_i$ and for obtaining the conditional probabilities of the BCJR and the states of the trellis. Hence, there is a small mismatch between the decoder and the channel, such that the BCJR-decoder is not strictly optimal. We consider blocks of length $K = 10^4$ symbols and simulate as many blocks until we observe at least 100 symbol error events. For the computation of the spectral efficiencies, we consider the 3 dB-bandwidth of the RRC transmit filter such that the roll-off factor α_{RC} directly relates to the percentage of excess or out-of-band bandwidth. We compare the

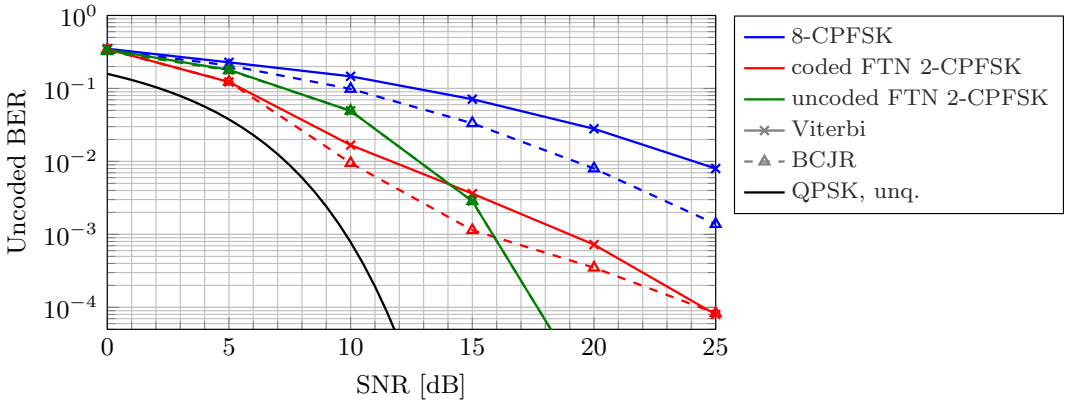

Fig. 5.11. BER results for oversampled 8-CPFSK ($M_{\mathrm{osr}} = 5$, $2WT_{\mathrm{symb}} = 1.5$), uncoded FTN 2-CPFSK ($M_{\mathrm{osr}} = 1$, $2WT_{\mathrm{symb}} = 1/2$), and coded FTN 2-CPFSK ($M_{\mathrm{osr}} = 5$, $2WT_{\mathrm{symb}} = 1.5$), in all cases $\mathrm{SE}_{\max} = 2\,\mathrm{bit/s/Hz}$, $\alpha_{\mathrm{RC}} = 1/2$, and $n_{\mathrm{IF}} = n_{\mathrm{IF,min}}$

results to the theoretical bit error probabilities of unquantized quadrature phase-shift keying (QPSK) and 8-PSK.

In Fig. 5.11 we compare the BCJR-based decoder, cf. Section 5.4.1, with the simplified metric described in Section 5.4.2 used in a Viterbi decoder for one example of each approach, i.e., oversampled 8-CPFSK, coded and uncoded FTN 2-CPFSK signaling. For the systems using temporal oversampling, i.e., oversampled 8-CPFSK and coded FTN 2-CPFSK, the BCJR outperforms the Viterbi decoder with the simplified decision metric. For uncoded FTN 2-CPFSK with no temporal oversampling and independent symbols and, thus, a fully connected trellis, the BCJR and the Viterbi performs equally. Overall, both decoders give reasonable results such that we use the simpler Viterbi decoder for further analysis due to the lower complexity. Furthermore, we see that out of the investigated approaches coded FTN 2-CPFSK can achieve BERs closest to the theoretical results of a conventional unquantized QPSK system with Nyquist signaling

$$p_{b,\mathrm{QPSK}} = \frac{1}{2}\operatorname{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right) = \frac{1}{2}\operatorname{erfc}\left(\sqrt{\frac{\mathrm{SNR}}{2}}\right) \tag{5.49}$$

up to an SNR of ca. 15 dB. On the other hand, oversampled 8-CPFSK is very sensitive to noise and the BER only decreases slowly over the SNR. The performance of uncoded FTN 2-CPFSK lies in-between these two up to ca. 15 dB SNR after which it begins to outperform the coded FTN 2-CPFSK example. This is due to the fact that the coded FTN 2-CPFSK uses 3.33-fold FTN while the uncoded example uses only 2-fold FTN. It is, thus, less affected by the ISI introduced by the channel filters, which is the reason for the superior performance in the high-SNR regime. All results given have a spectral efficiency in the noise and distortion free limit of $\mathrm{SE}_{\max} = 2\,\mathrm{bit/s/Hz}$ and a roll-off factor $\alpha_{\mathrm{RC}} = 1/2$ was used.

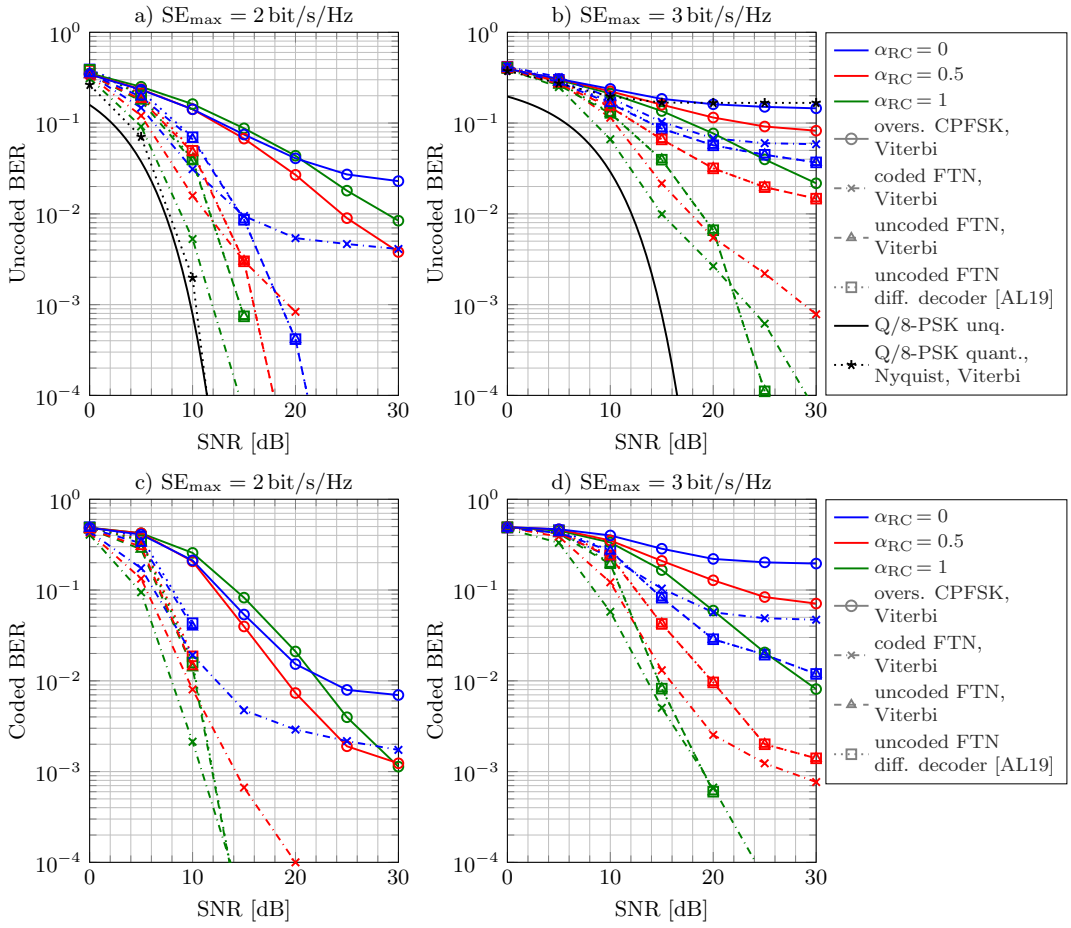

Fig. 5.12. Uncoded and coded BER results for configurations providing spectral efficiencies of $\mathrm{SE}_{\max} = 2\,\mathrm{bit/s/Hz}$ and $\mathrm{SE}_{\max} = 3\,\mathrm{bit/s/Hz}$, respectively, for $n_{\mathrm{IF}} = n_{\mathrm{IF,min}}$

Fig. 5.12 shows the uncoded and coded BERs for configurations yielding at most 2 and 3 bit/s/Hz spectral efficiency, respectively, for different amounts of excess bandwidth given by $\alpha_{\mathrm{RC}} = \{0, 0.5, 1\}$. We use a rate one-half convolutional code with memory 5 and generator polynomial $(62, 57)$ in octal notation. This code was found to perform well for iterative decoding of ISI for FTN signals in [AZ12]. Note that the notion of $\mathrm{SE}_{\max}$ always refers to the transmitted and not to the information bits. As expected, the BER performance improves with increasing excess bandwidth.[3] Furthermore, coded and uncoded FTN 2-CPFSK provide a significant gain in BER compared to oversampled CPFSK, due to the advantage of larger distance to the quantization thresholds and only one constellation point per quadrant. Depending on the (excess) bandwidth available, i.e., $2WT_{\mathrm{symb}}$ and α_{RC}, the relative performance of coded and uncoded FTN 2-CPFSK varies. If sufficient bandwidth is available (cf.

[3] A slight deviation from this is observed in Fig. 5.12a) for oversampled CPFSK. Since $2WT_s = 1.5$ there is sufficient bandwidth available and the distortion of the main lobe due to $\alpha_{\mathrm{RC}} = 1$ is more severe than what is gained by the excess bandwidth.

Fig. 5.12a) or Fig. 5.12b) $\alpha_{\mathrm{RC}} = 1$), uncoded FTN gains on coded FTN in the high-SNR regime due to the lower M_{FTN} and, thus, lower ISI. For comparison, in Fig. 5.12a) the QPSK bit error probability in (5.49) is shown and in Fig. 5.12b) an approximation of the bit error probability of unquantized 8-PSK Nyquist-signaling using Gray mapping is given as [PS08, Chapter 4.3-2]

$$p_{b,M_{\mathrm{cpm}}\mathrm{PSK}} \approx \frac{1}{\log_2 M_{\mathrm{cpm}}} \operatorname{erfc}\left(\sqrt{\frac{2\log_2(M_{\mathrm{cpm}})\sin^2\left(\frac{\pi}{M_{\mathrm{cpm}}}\right)}{2}\frac{E_b}{N_0}}\right) \tag{5.50}$$

$$= \frac{1}{\log_2 M_{\mathrm{cpm}}} \operatorname{erfc}\left(\sqrt{\sin^2\left(\frac{\pi}{M_{\mathrm{cpm}}}\right)\mathrm{SNR}}\right). \tag{5.51}$$

It can be observed that, using uncoded or coded FTN 2-CPFSK, for a given BER the SNR difference between 1-bit quantized and unquantized schemes can be significantly reduced compared to temporally oversampled CPFSK – especially if excess bandwidth is tolerable ($\alpha_{\mathrm{RC}} = \{0.5, 1\}$). It is not surprising that for $\mathrm{SE}_{\mathrm{max}} = 2$ bit/s/Hz it is of little benefit to go through the trouble of FTN signaling since this spectral efficiency can be easily achieved with Nyquist-QPSK under 1-bit quantization yielding an even better BER performance. In Fig. 5.12a) we see that for coded FTN 2-CPFSK with $\alpha_{\mathrm{RC}} = 0.5$ (50 % excess bandwidth), there is approximately 10 dB loss in the SNR required to reach a BER of 10^{-3} compared to the unquantized case. The SNR needed to achieve a BER of 10^{-3} can be reduced with the applied code by ca. 6 dB, cf. Fig. 5.12c). Increasing α_{RC} to 1 allows BERs close to the theoretical results for unquantized QPSK transmission at the expense of bandwidth. For uncoded FTN 2-CPFSK, we have only 6 dB and 4 dB gaps in the SNR for the uncoded BER for $\alpha_{\mathrm{RC}} = \{0.5, 1\}$.

Note that we have not carried out an investigation w.r.t. to the best convolutional codes for this transmission scenario. In [AZ12] only a subset of codes was found suitable for the faster-than-Nyquist ISI in the considered transmission scenario and decoder architecture, which differs from the one considered here. Thus, there is potential for improvement if the codes are adapted to the CPM-signal. For example, in [GBSS08] gains were reported when using the Gram-Schmidt orthogonalization transform on CPM-signals in order investigate good codes for the transmission with space time convolutional codes.

When higher spectral efficiencies are needed, FTN signaling provides gains w.r.t. the BER performance. The price to pay for the increased spectral efficiency is SNR as in the best case there is still around 7 and 4 dB more SNR required compared to unquantized 8-PSK in the uncoded and coded cases, respectively, to achieve a BER of 10^{-3}, cf. Fig. 5.12b) and d). However, the FTN CPFSK system can be used with a 1-bit quantizer, which is not the case for classic 8-PSK.

Tab. 5.1. Parameter Configuration of the Simulation

	oversampled CPFSK		coded FTN 2-CPFSK		uncoded FTN 2-CPFSK	
$M_{\mathrm{osr},t}$	5		5		1	
M_{cpm}	8		2		2	
$n_{\mathrm{IF,min}}$	$9/16$		0		0	
$2WT_{\mathrm{symb}}$	1	1.5	1	1.5	$1/3$	$1/2$
$2WT'_{\mathrm{symb}}$	1	1.5	$1/5$	$3/10$	$1/3$	$1/2$
M_{FTN}	1	$2/3$	5	$10/3$	3	2
$\mathrm{SE}_{\mathrm{max}}$	3	2	3	2	3	2

Furthermore, the BERs of the simple differential decoder from [dAL19] are given for uncoded FTN 2-CPFSK. For this low complexity modulation it was seen in Fig. 5.11 that the results of the Viterbi decoder coincide with the BCJR decoder results. As observed in [dAL19] their differential decoder performance also matches the BCJR results since there is no temporal oversampling and no information in the sequence of the symbols to be exploited. Consequently, the performance of the differential decoder equals the one of the Viterbi decoder with almost negligible performance loss in the low-SNR domain. However, the decoding rule is specifically designed for 2-CPFSK and could not be applied to coded FTN 2-CPFSK or oversampled CPFSK. Table 5.1 summarizes the parameters used for the different setups in Fig. 5.12. Note that coded FTN 2-CPFSK is the only case where $T'_{\mathrm{symb}} \neq T_{\mathrm{symb}}$.

Lastly, we compare the BER of the Viterbi decoder of temporally oversampled CPFSK and temporally oversampled PSK in Fig. 5.13. In Section 5.3 we saw that PSK ideally requires a smaller OSR than CPFSK, which we seek to investigate for noisy scenarios as well. It can be observed that given the same OSR, indeed the discontinuities of PSK improve the BER performance despite the increased spectral side lobes. The gain is even larger for the narrower filters in Fig. 5.13b), however, it does not match the improvements achieved by coded and uncoded FTN 2-CPFSK. Furthermore, decreasing the temporal OSR for PSK to $M_{\mathrm{osr,min}}$ does increase the BER compared to CPFSK at its respective $M_{\mathrm{osr,min}}$. Thus, temporally oversampled PSK can enable a reduction in sampling rate compared to temporally oversampled CPFSK but not without a BER degradation.

5.7 Summary

In this chapter the advantage of sampling a phase modulated signal at an IF has been demonstrated given a 1-bit quantizer at the receiver. The IF can be used to manipulate the phase trajectories associated with the transmit symbols in a way that makes them distinguishable even under 1-bit quantization if a sufficiently high oversampling ratio

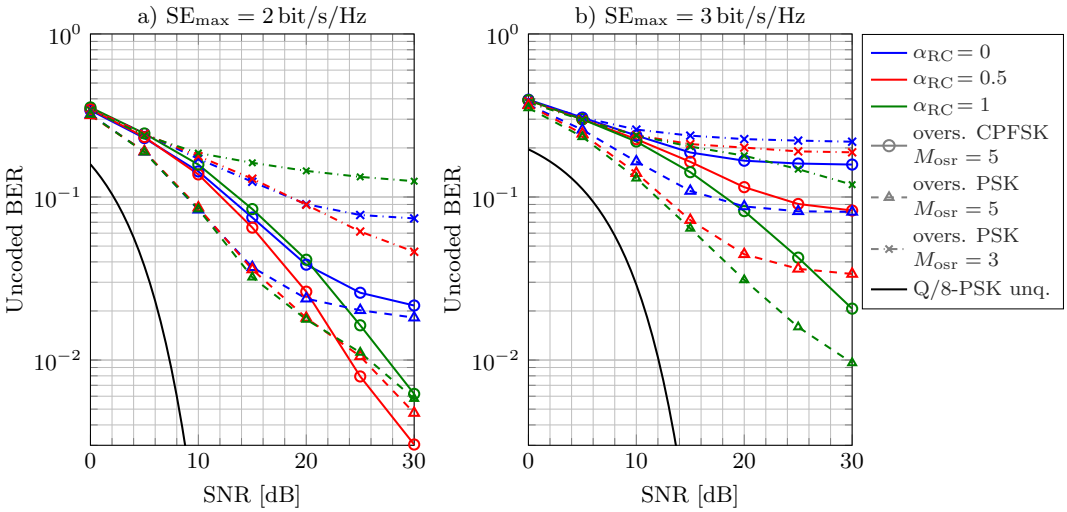

Fig. 5.13. BER results for oversampled 8-CPFSK and 8-PSK providing spectral efficiencies of a) $\mathrm{SE}_{\max} = 2\,\mathrm{bit/s/Hz}$, $2WT_{\mathrm{symb}} = 1.5$ and b) $\mathrm{SE}_{\max} = 3\,\mathrm{bit/s/Hz}$, , $2WT_{\mathrm{symb}} = 1$ for $n_{\mathrm{IF}} = n_{\mathrm{IF,min}}$

can be applied. Even for the non-continuous PSK modulation, the IF creates linear phase trajectories that can be resolved by temporal oversampling.

Under the assumption of a time-invariant trellis and a symmetric constellation diagram, the requirements for resolving all possible transmit symbols in an ideal, i.e., noise and distortion free, setup have been derived. This includes analytical expressions for the minimum IF required for resolving all symbols for CPM and the minimum OSR for resolving all symbols for CPFSK. For other modulation schemes than CPFSK, namely PSK and a sine-wave phase transition, the minimum OSR for resolving all symbols has been found by a simple search. While the results for the two continuous-phase scheme, CPFSK and sine, barely differ, the discontinuous PSK modulation could potentially require lower OSRs.

Two trellis-based decoder architectures have been presented in order to evaluate the BER performance of 1-bit quantized phase modulation: a BCJR decoder based on the conditional probabilities between channel input and output and a Viterbi decoder based on a Hamming-distance like decision metric. All results consider bandlimited channels with RRC transmit and receive filters. It results that temporally oversampled CPFSK is very sensitive to noise and filter distortion since the samples distinguishing between two symbols can be arbitrary close to the transition threshold. This can be mitigated by using PSK modulation. However, if PSK modulation is used to reduce the oversampling ratio, a penalty w.r.t. the BER is incurred.

Another option to improve the BER performance is applying CPFSK with a low modulation order jointly with FTN signaling. In this context, the notion of coded and uncoded FTN signaling has been introduced using 2-CPFSK with modulation index $h_{\mathrm{cpm}} = 1/4$. Comparing these two, uncoded FTN-CPFSK has shown to perform better

in the high-SNR domain while coded FTN-CPFSK provides lower BERs in the low-to-mid SNR domain. The gain in SNR required for achieving a given BER compared to temporally oversampled CPFSK exceeds the improvements obtained by PSK and can in extreme cases be up to 15 or 20 dB. The price for boosting the spectral efficiency by applying FTN signaling to CPFSK is an amount of excess bandwidth as well as an increased SNR to achieve a certain BER. However, the resulting system can be implemented using 1-bit quantizer and highly efficient non-linear PAs. For example, the PA in [GURP16] provides when used between 71 to 95 GHz a peak power added efficiency of 40 % while at the 1 dB compression point this reduces to 10 %. Roughly, this corresponds to a factor of four – or 6 dB – in power required at the PA, which can compensate for the larger SNR required compared to traditional schemes. Additional power savings are achieved by using the 1-bit quantizers.

6 Conclusions and Future Work

Based on a bottom-up approach where hardware restrictions imply system design choices in order to increase the energy efficiency of the communications system, the performance of a bandlimited 1-bit quantized wireless communication system has been investigated. Here, 1-bit quantization is applied to reduce the analog-to-digital converter (ADC) power consumption at the receiver. Using performance metrics like the achievable rate or mutual information rate and the bit error rate (BER), we studied

- the hard bandlimited continuous-time 1-bit quantized additive white Gaussian noise (AWGN) channel,
- the oversampled discrete-time bandlimited 1-bit quantized channel with and without excess bandwidth,
- the oversampled discrete- and continuous-time 1-bit quantized multiple-input multiple-output (MIMO) channels, and
- temporally oversampled and faster-than-Nyquist (FTN) phase modulated communication over bandlimited channels.

The continuous-time 1-bit quantized channel was studied as a limiting case representing an infinitely oversampled system, i.e., $M_{\mathrm{osr}} \to \infty$. We derived analytic expressions for bounds on the mutual information rate when transmitting over an AWGN channel in the mid-to-high signal-to-noise ratio (SNR) regime. The two lower bounds, one considering intersymbol interference (ISI) due to hard bandlimitation and one neglecting the ISI, are based on real and runlength encoded input signals with an exponentially distributed runlength. An upper bound on the ISI-free channel – and, thus, implicitly also on the ISI channel – for the specific signaling scheme shows that the mathematical bounding steps applied besides treating the ISI as noise do not have a significant impact on the tightness of the bounds. Comparing the two lower bounds on the mutual information rate, however, shows that modeling the ISI as noise has a crucial impact: the lower bound on the mutual information rate saturates over the SNR only if a hard bandlimitation is considered. In this case, we observed that the lower and the upper bound are close to each other in the mid-to-high SNR regime as long as the impact of the hard bandlimitation does not dominate, i.e., in the SNR range of approximately 10 to 20 dB. In general, for a given SNR the spectral efficiencies solely depend on the ratio of channel bandwidth and the rate parameter of the exponential

distribution at the channel input. For real signals the maximum value of the lower bound on the spectral efficiency in the high-SNR limit was found to be approximately 1.63 bit/s/Hz in the bandlimited case while in the ISI-free case it grows without limit. In order to obtain analytical results, assumptions and approximations were required based on which the bounds are valid in the mid-to-high SNR domain above 10 dB.

In a practical scenario, however, the oversampling ratio remains finite. Therefore, lower bounds on the achievable rate of the runlength modulated 1-bit quantized discrete time AWGN channel using FTN signaling have been derived. Here, FTN signaling corresponds to oversampling w.r.t. the channel bandwidth. Two lower bounds are derived, one is based on an analytical expression and one is simulation-aided in the sense that it requires a sufficiently long realization of the filtered transmit signal. With increasing intersymbol interference, the bounds diverge until a point where the analytical bound collapses. However, for low to moderate ISI both bounds are close to each other. Given a hard bandlimited channel, the achievable rates derived match the results of the continuous-time channel well and achieve up to approximately 1.53 bit/s/Hz in th high-SNR limit. When excess bandwidth is tolerable, e.g., for raised cosine-filters with roll-offs larger than zero, spectral efficiencies up to 2.23 bit/s/Hz in the real domain are achievable w.r.t. the 95 %-power containment bandwidth. For a root raised cosine-filter with roll-off $\alpha_{\mathrm{RC}} = 1$ even 2.47 bit/s/Hz have been observed. We show that these bounds as well as the ones for the continuous-time scenario can be applied to the oversampled 1-bit quantized MIMO channel using appropriate power allocation schemes. Especially in the high-SNR regime the gain by oversampling or FTN signaling is significant.

In both scenarios, discrete-time and continuous-time, we demonstrated that applying the concepts of FTN signaling and oversampling, which are closely related, can significantly improve the achievable rates of the bandlimited 1-bit quantized channel compared to Nyquist-sampling and -signaling. The bounds on the achievable rate in the discrete-time case match the corresponding mutual information rates in the coninuous-time case remarkably well, which indicates that the mutual information rates derived in Chapter 3 can be achievable. Furthermore, the actual gain in achievable rates depends strongly on the type of bandlimitation considered. In the discrete-time model, the relative difference between the maximum spectral efficiencies of the hard bandlimited channel and the channel with excess bandwidth reported above is significant. We therefore conclude that knowing the restrictions on spectral characteristics is crucial for predicting the achievable data rates when designing a high-speed 1-bit quantized transmission system. If the assumed spectral constraints are too strict, potential is wasted while the achievable rates are overestimated if the spectral constraints are underestimated.

As a constant envelope modulation scheme, continuous phase modulation (CPM) provides the opportunity of using highly efficient non-linear, e.g., class-E, power ampli-

fiers. Thus, synergies between CPM and oversampled 1-bit quantized communication were investigated. Analysis of the limiting performance in terms of distinguishable symbols in a noise-free scenario shows that processing the signal at an intermediate frequency (IF) is advantageous under 1-bit quantization. We determined the minimum IF for temporally oversampled CPM, for which all transmitted symbols can be resolved given a sufficiently high oversampling ratio. Based on this, we quantified the minimum required oversampling ratio for a special case of CPM, continuous phase frequency shift keying (CPFSK), with an analytical expression. Since we found that temporally oversampled CPFSK is very sensitive to noisy transmission, we investigated FTN signaling for CPFSK in order to increase the robustness w.r.t. noise without compromising on the spectral efficiency. We introduced the notions of uncoded and coded FTN signaling for CPFSK where in the latter case the message is encoded into a subsequence of symbols. We found that both approaches can greatly improve the BER performance for a given spectral efficiency. For example, with coded FTN 2-CPFSK, the gain in SNR required for achieving an uncoded BER of 10^{-2} compared to temporally oversampled CPFSK is approximately 12 dB (19 dB) for a roll-off factor of 0.5 (1) at 2 bit/s/Hz with a Viterbi decoder using a simplified distance metric.

Future Work

There are several topics remaining that are useful extensions to this work or were beyond the scope. These are summarized in the following:

- So far, only AWGN channels with ISI have been analyzed in this work. This was reasonable given the scenario of ultra-short range high data rate communication between printed circuit boards, which motivated this work. With communication moving to ever higher frequency bands with ever higher bandwidths, it is of interest if 1-bit quantization, FTN signaling and oversampling are also useful tools to achieve energy efficient communication in time or frequency selective channels.
- Designing an energy-efficient communication system requires an end-to-end approach. A first step has been taken in this work, e.g., by considering constant-envelope modulation schemes and aiming to simplify the decoder architecture. However, in order to fully consider the impact of modulation and coding on the energy consumption, the power consumption of encoder, decoder, power amplifier etc. would need to be included in further analysis. A simple modulation scheme does not help if it requires a complex decoder or complex channel codes in order to provide acceptable BERs. This holds especially true when considering time or frequency selective channels as suggested above such that co-design of modulation and coding is necessary.

- Modeling ISI as noise, as for example in Chapter 3, is clearly sub-optimal. Other imperfections which have not been considered include phase and timing offsets. Channel estimation and equalization under 1-bit quantization is already a topic under investigation, e.g., [SDF19a; SDF19b]. These insights need to be considered in order to quantify the impact of imperfect channel knowledge on the achievable rate and BER such that performance predictions of realistic systems are possible.
- Regarding the characterization of the mutual information rate of the continuous-time bandlimited 1-bit quantized channel, the proof of Shannon's coding theorem for insertion and deletion channels with continuous and potentially infinite random variables was beyond of the scope of this work and remains an open topic.
- The systematic design of codes for coded FTN for CPM has not yet been addressed and offers the potential of further performance improvements.

Appendix Chapter 3

A

A.1 Spectrum of the Continuous-Time Signal $\mathsf{x}(t)$

The definition of the power spectral density (PSD) of a random process is

$$S_{\mathsf{X}}(\omega) = \lim_{K\to\infty} \frac{\mathbb{E}\left[|\mathsf{X}(\omega)|^2\right]}{KT_{\text{avg}}} \tag{A.1}$$

where the spectrum $\mathsf{X}(\omega)$ of $\mathsf{x}(t)$ is given by

$$\mathsf{X}(\omega) = \sqrt{\hat{P}}\left(\sum_{k=1}^{K}(-1)^k\, G(\omega)\, e^{-j\omega \mathsf{T}_k}\right) + 2\pi\delta(\omega). \tag{A.2}$$

It holds that $G(\omega)$, cf. (3.13), can be expressed as

$$G(\omega) = -j\left[\frac{1+e^{-j\omega\beta}}{\omega} + e^{-j\omega\frac{\beta}{2}}a(\omega)\right] \tag{A.3}$$

where $a(\omega)$ is a real function in $\mathbb{R}$ given as

$$a(\omega) = -\frac{1}{j}\int_{-\frac{\beta}{2}}^{\frac{\beta}{2}} \zeta(t)e^{-j\omega t}dt. \tag{A.4}$$

The squared magnitude of (A.2) becomes

$$\frac{|\mathsf{X}(\omega)|^2}{\hat{P}} = 4\pi^2\delta^2(\omega) + 4\pi\delta(\omega)\Re\left\{G(\omega)\sum_{k=1}^{K}(-1)^k e^{-j\omega\mathsf{T}_k}\right\} + \left|G(\omega)\sum_{k=1}^{K}(-1)^k e^{-j\omega\mathsf{T}_k}\right|^2. \tag{A.5}$$

where third term on the right-hand side of (A.5) can be written as

$$|G(\omega)|^2\sum_{k=1}^{K}\sum_{v=1}^{K}(-1)^{k+v}\cos(\omega(\mathsf{T}_k - \mathsf{T}_v)). \tag{A.6}$$

Here, we have

$$|G(\omega)|^2 = \frac{2(1+\cos(\omega\beta))}{\frac{1}{2}\omega^2} + a^2(\omega) + \frac{4a(\omega)\cos\left(\frac{\omega\beta}{2}\right)}{\omega}. \tag{A.7}$$

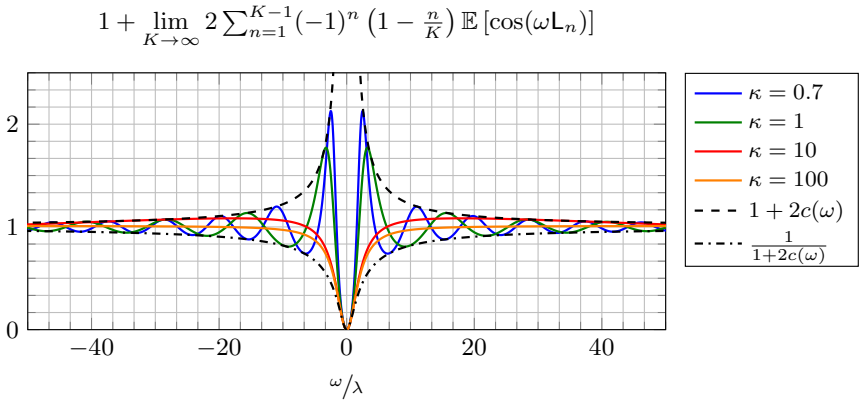

Fig. A.1. Numerical evaluation of the infinite sum in (A.8) for $K = 10^4$ as well as the bounds from (A.9) and (A.11)

In (A.5) the first two terms correspond to a DC-component, which is not relevant for further calculations. Since the cosine is an even function, the PSD of $\mathsf{x}(t)$ for $|\omega| > 0$ in (A.1) is given by

$$S_{\mathsf{X}}(\omega) = \frac{\hat{P}\,|G(\omega)|^2}{T_{\text{avg}}}\left(1 + \lim_{K\to\infty} 2\sum_{n=1}^{K-1}(-1)^n\left(1-\frac{n}{K}\right)\mathbb{E}\left[\cos(\omega \mathsf{L}_n)\right]\right). \tag{A.8}$$

Here, $n = k - v$ describes the distance between two arbitrary zero-crossing (ZC) instants in terms of the number of symbols A_k and $\mathsf{L}_n = \mathsf{T}_k - \mathsf{T}_v = \sum_{i=1}^{n}\mathsf{A}_{k+i}$ denotes the corresponding random variable. As given in (3.8), L_n follows an Erlang-distribution since it is a sum of exponentially distributed random variables. Therefore, we can write for the expectation in (A.8)

$$\mathbb{E}\left[\cos(\omega \mathsf{L}_n)\right] = \left(\frac{\lambda}{\sqrt{\lambda^2+\omega^2}}\right)^n \cos\left(n\left(\omega\beta + \arctan\left(\frac{\omega}{\lambda}\right)\right)\right) \leq \left(\frac{\lambda}{\sqrt{\lambda^2+\omega^2}}\right)^n \tag{A.9}$$

with $q = \frac{\lambda}{\sqrt{\lambda^2+\omega^2}}$. For the upper bound of the infinite sum in (A.8) we have

$$\lim_{K\to\infty}\sum_{n=1}^{K-1}\left(1-\frac{n}{K}\right)\left(\frac{\lambda}{\sqrt{\lambda^2+\omega^2}}\right)^n = \frac{\lambda}{\sqrt{\lambda^2+\omega^2}-\lambda} = c(\omega) \tag{A.10}$$

yielding the upper bound in (3.21). Furthermore, we find that the infinite sum in (A.8) has periodic minima, cf. Fig. A.1, which always occur if $\omega\beta + \arctan\left(\frac{\omega}{\lambda}\right) = 2m\pi$, $m \in \mathbb{Z}$. In this case the cosine in (A.8) is always one such that it remains

$$\lim_{K\to\infty}\sum_{n=1}^{K-1}(-1)^n\left(1-\frac{n}{K}\right)\left(\frac{\lambda}{\sqrt{\lambda^2+\omega^2}}\right)^n = -\frac{\lambda}{\sqrt{\lambda^2+\omega^2}+\lambda}. \tag{A.11}$$

With (A.11) we can lower-bound $S_{\mathsf{X}}(\omega)$ in (A.8) by (3.21). Here, we have used that $1 - 2\frac{\lambda}{\sqrt{\lambda^2+\omega^2}+\lambda} = \frac{1}{1+2c(\omega)}$ with $c(\omega)$ given in (A.10).

A.2 Occurrence of Zero-Crossing Deletions and Insertions

We assumed a fixed relation of bandwidth W and transition time β in (3.9), i.e., $W = 1/2\beta$. Removing this restriction and allowing β to take on any value, augments the design space and potentially increases the spectral efficiency of the communication system. However, this also means that the minimum distance between two ZCs β is no longer linked to the coherence time of the noise, which was the reason for excluding deletion errors in the analysis. In this section, we aim to verify this connection by simulation.

We generate $\mathsf{x}(t)$ for a given signal-to-noise ratio (SNR) ρ, λ, and β. We assume a long sequence of $K = 10^3$ symbols and a time resolution $\Delta t = 10^{-3}\lambda^{-1}$, for which we analyze the received signal $\mathsf{r}(t)$ after the receive filter. The locations of insertions and deletions can be identified by matching every received upcrossing in $\mathsf{r}(t)$ to the closest upcrossing in $\mathsf{x}(t)$, likewise for the downcrossings. Then, the inserted and deleted symbols are counted.

The number of deletions and insertions are depicted in Fig. A.2 and Fig. A.3 exemplarily for a sine-based and a linear transition waveform as well as for three different SNR values of 5 dB, 10 dB and 15 dB. Here, we defined $\tilde{\kappa} = 1/(2\beta\lambda)$. It can be seen in parts a) to c) of the figures, that the SNR has a rather small impact on the number of deletions occurring in the mid-to-high SNR regime. The black line represents the case $W = 1/2\beta$. We observe that the number of deletions is negligible for bandwidths $W \geq 1/2\beta$, i.e., above the black line. This is due to the fact that the dynamics of the noise are high compared to the minimum symbol duration β. However, for bandwidths $W < 1/2\beta$ deletions are possible and have to be included in the system model as otherwise the spectral efficiency of the system will be overestimated. The validity of this observations is not affected by the choice of the waveform. Given that this error event mainly occurs in the hold period (HP), i.e., when $\mathsf{x}(t) = \pm\sqrt{\hat{P}}$, this was to be expected.

On the other hand, we see in part d) to f) of Fig. A.2 and Fig. A.3 that insertions are to be considered as possible error event for small to medium SNR and that they occur much more frequently than deletions. Intuitively in contrast to the deletions, the occurrence of insertions is particularly frequent when the dynamics of the noise are high compared to the minimum symbol duration β, i.e., $W > \frac{1}{2\beta}$. Even for $W = 1/2\beta$ the number of insertions is non-zero for small to medium SNR and only in the high SNR domain they disappear. This observations coincide with the results of our analysis presented in Section 3.4.

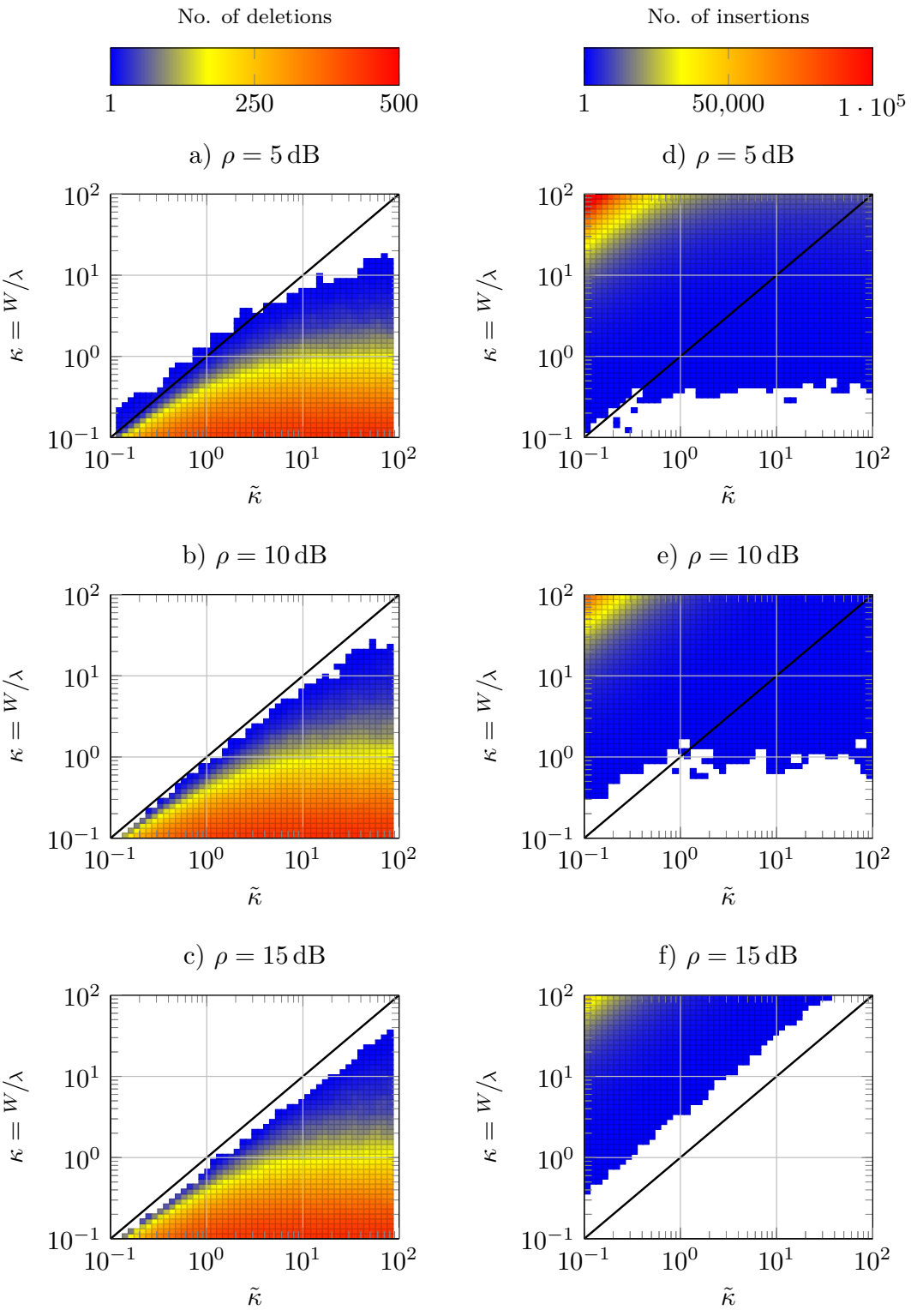

Fig. A.2. Number of deletions (part a) to c)) and insertions (part d) to f)) over normalized bandwidth $\kappa = {}^{W}/\lambda$ and over $\tilde{\kappa}$ for different ρ given a linear transition waveform

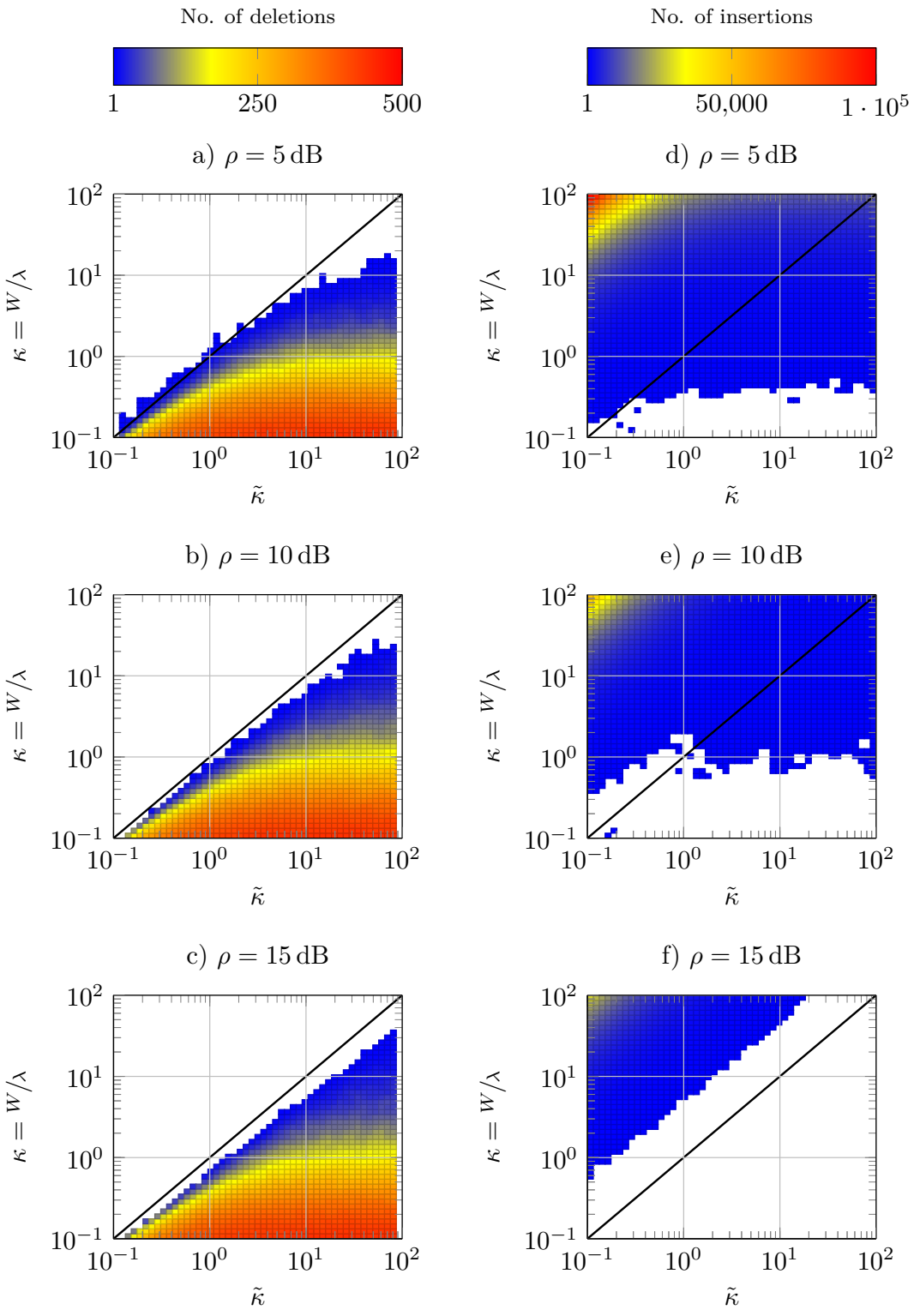

Fig. A.3. Number of deletions (part a) to c)) and insertions (part d) to f)) over normalized bandwidth $\kappa = {}^{W}/_{\lambda}$ and over $\tilde{\kappa}$ for different ρ given a sine-shaped transition waveform

A.3 Number of Zero-Crossings within a Transition Interval

Obtaining the expected number of ZCs in the transition interval (TI) $[\mathsf{T}_k, \mathsf{T}_k + \beta]$ is a curve crossing problem. It depends on the deterministic transmit waveform as well as the random process $\mathsf{z}(t)$, cf. (3.19). In Section 3.1.5 it was shown that $\tilde{\mathsf{x}}(t)$ and, thus, $\mathsf{z}(t)$ can be approximated to be Gaussian, see also (3.54). Therefore, this problem can be analyzed by studying the ZCs of a non-stationary Gaussian process $\mathsf{q}(t) = \mathsf{z}(t) - \psi(t)$. Here, $\psi(t)$ is the deterministic curve to be crossed by $\mathsf{z}(t)$. We thus define the TI $\mathbb{Y} = [0, \beta]$ with the time variable $y \in \mathbb{Y}$. Then $\psi(y)$ depends on the filtered transmit pulse $\hat{g}(y)$ as

$$\psi(y) = \hat{g}\,(y) - 1. \tag{A.12}$$

If $\mathsf{z}(y) = \psi(y)$, the process $\mathsf{q}(t)$ has a ZC in $\mathbb{Y}$. It is known that the number of crossings $\mathsf{N}_T(\psi)$ of a curve ψ by a stationary Gaussian process in a time interval of length T is given by [Kra06]

$$\mathbb{E}[\mathsf{N}_T(\psi)] = \sqrt{-s''(0)} \int_0^T \varphi(\psi(y)) \left[2\varphi\left(\frac{\psi'(y)}{\sqrt{-s''(0)}}\right) + \frac{\psi'(y)}{\sqrt{-s''(0)}} \left(2\Phi\left(\frac{\psi'(y)}{\sqrt{-s''(0)}}\right) - 1 \right) \right] dy. \tag{A.13}$$

Here, $s(\tau)$ denotes the auto-correlation function (ACF) of the Gaussian Process and $'$ is the derivative in time, i.e., w.r.t. y. Moreover, φ and Φ are the zero-mean Gaussian density and distribution functions with variance σ_{z}^2, respectively. It then results for the variance of the number of ZCs [Kra06]

$$\begin{aligned} \mathrm{Var}(\mathsf{N}_T(\psi)) = \mathbb{E}[\mathsf{N}_T(\psi)] - \mathbb{E}^2[\mathsf{N}_T(\psi)] + \int_0^T \int_0^T \int_{\mathbb{R}} |\mathsf{q}'_{t_1} - \psi'_{t_1}| \\ \cdot |\mathsf{q}'_{t_2} - \psi'_{t_2}| \phi_{t_1,t_2}(\psi_{t_1}, \mathsf{q}'_{t_1}, \psi_{t_2}, \mathsf{q}'_{t_2}) d\mathsf{q}'_{t_1} d\mathsf{q}'_{t_2} dt_1 dt_2 \end{aligned} \tag{A.14}$$

where the subscripts t_1 and t_2 denote the time instants. Furthermore, ϕ is the multivariate zero-mean normal distribution of $\mathsf{q}(t_1)$, $\mathsf{q}'(t_1)$, $\mathsf{q}(t_2)$, and $\mathsf{q}'(t_2)$ with covariance matrix

$$\Sigma = \begin{Bmatrix} s(0) & 0 & s(\tau) & s'(\tau) \\ 0 & -s''(0) & -s'(\tau) & -s''(\tau) \\ s(\tau) & -s'(\tau) & s(0) & 0 \\ s'(\tau) & -s''(\tau) & 0 & -s''(0) \end{Bmatrix}. \tag{A.15}$$

We evaluate the equations (A.13) and (A.14) exemplarily for a sine-shaped and a

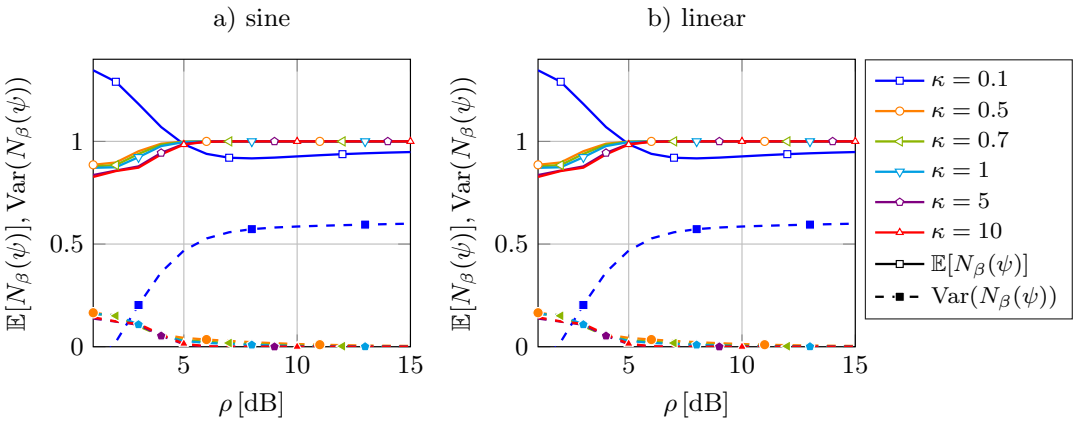

Fig. A.4. Expectation and variance of the number of ZCs in the TI $[T_k - \beta/2, T_k + \beta/2]$ for a) sine-shaped and b) linear transition

linear transition, which is depicted in Fig. A.4. We observe that the expectation of the number of ZCs converges to one for SNR $\gtrsim 5$ dB and $\kappa = W/\lambda \geq 0.5$. At the same time, the variance converges to zero. Thus, there exists with high probability only one ZC in every TI for an SNR $\gtrsim 5$ dB and, therefore, there are no insertions in the TI. For $\kappa \ll 1$ the validity of the assumption is not relevant since the lower bound on the mutual information rate in (3.85) becomes zero.

A.4 Mid-to-high SNR Assumption $\mathsf{S}_k \ll \beta$

We aim to quantify the SNR regions for which

(a) $\mathsf{S}_k < \left|\frac{\beta}{2}\right|$ with high probability, which is required for the temporal separation of the error events (ZC shifts and insertions) and

(b) the linearization in (3.53) is valid.

Therefore, the variances of both densities, (3.56) and (3.57), have been evaluated and compared numerically. Fig. A.5a) and b) depict the normalized variances $W^2\sigma_{\mathsf{S}}^2$ exemplarily for a linear and a sine-shaped transition waveform. Thereby, the variance of the original probability density function (pdf) in (3.56) is only plotted if

$$\Pr(|\mathsf{S}_k| < \beta/2) = \int_{-\beta/2}^{\beta/2} p_{\mathsf{S}}(s)ds \geq 0.95 \tag{A.16}$$

i.e., with large probability $|\mathsf{S}_k| < \beta/2$, cf. Fig. A.5c) and d). We observe that in the relevant regime of $\kappa = W/\lambda \geq 0.5$, the variances $\sigma_{\mathsf{S},\mathrm{orig}}^2$ and σ_{S}^2 are very close when $\rho \gtrsim 10$ dB, for which $\Pr(|\mathsf{S}_k| < \beta/2) > 0.99$. Again, due to filtering, the impact of the waveform is negligible, cf. Section 3.3. We thus conclude that for a sufficiently high SNR, for which the ZC is not shifted out of the TI with high probability, the

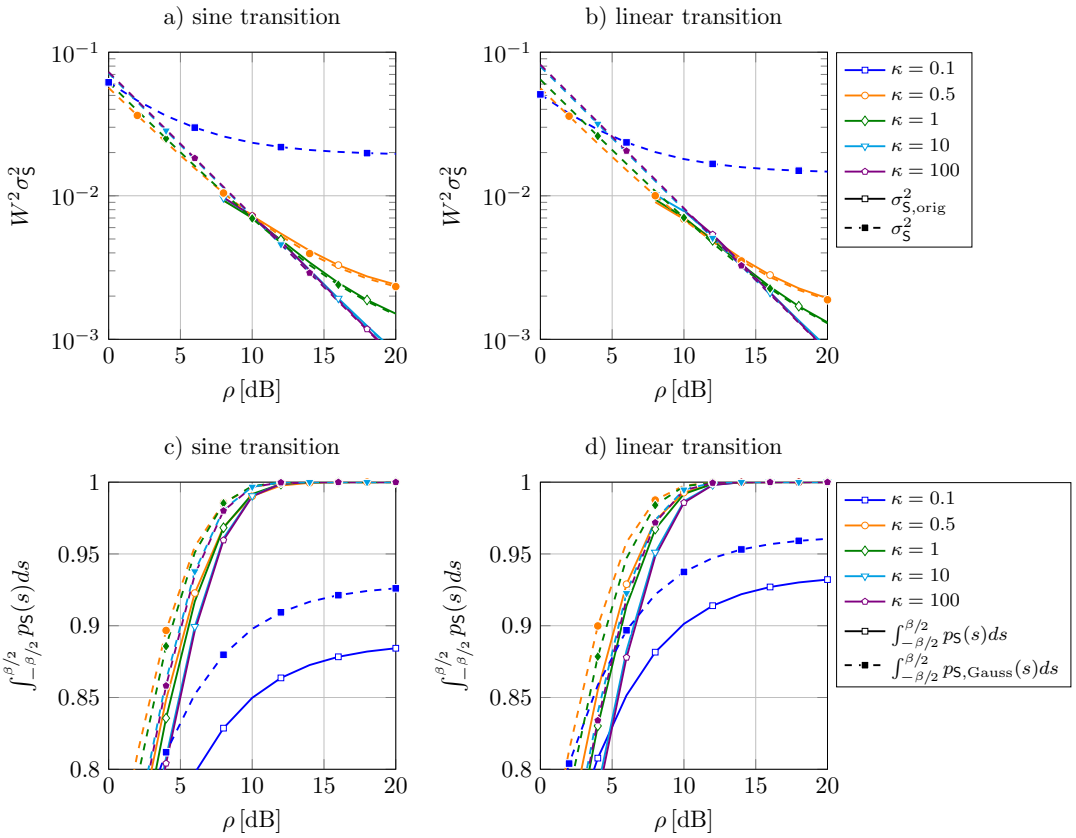

Fig. A.5. Properties of the original distribution $p_{\mathsf{S}}(s)$ (3.56) and the Gaussian approximation $p_{\mathsf{S},\text{Gauss}}(s)$ (3.57), with a), b) normalized variances $\sigma^2_{\mathsf{S},\text{orig}}$ of $p_{\mathsf{S}}(s)$ and σ^2_{S} of $p_{\mathsf{S},\text{Gauss}}(s)$ and c), d) $\Pr(|\mathsf{S}_k| < \beta/2) = \int_{-\beta/2}^{\beta/2} p_{\mathsf{S}}(s)ds$ for $p_{\mathsf{S}}(s)$ and $p_{\mathsf{S},\text{Gauss}}(s)$.

linearization holds as well. Again, for $\kappa \ll 1$ the validity of the assumption is not relevant since the lower bound on the mutual information rate in (3.85) becomes zero. It furthermore is sufficient to compare the variances since all further bounding steps of $I'(\mathbf{A}; \mathbf{D}, \mathbf{V})$ are solely based on the variance of a Gaussian random process with equal covariance matrix.

A.5 Assumption of Independence of $\mathsf{z}(t)$

In order to analyze the relation between $\mathbf{Q}_{\text{err}}$ and $\mathbf{R}_{\text{err}}$, cf. (3.68) and (3.71), we require an expression for $\mathbf{Q}_{\mathsf{A}\Delta}$ and $\mathbf{Q}_{\Delta}$. In what follows, we will use the short notation

Tab. A.1. Weights for computing the entries of cross-covariance matrix $\mathbf{Q}_{A\Delta}$

$\cdot(-1)^i$	...	ν_3	ν_2	ν_1	ξ_1	ξ_2	ξ_3	...
$j=i$	...	0	0	0	0	0	0	...
$j=i+1$	...	2	-2	1	-1	2	-2	...
$j=i+2$	...	-2	1	0	0	-1	2	...
$j=i+3$	...	1	0	0	0	0	-1	...

$\mathsf{z}_j = \mathsf{z}(\mathsf{T}_j)$ for all random processes at time T_j. Thus, we write for the respective entries $q_{\mathsf{A}\Delta,(i,j)} = \mathbb{E}[\mathsf{A}_i\Delta_j]$ and $q_{\Delta,(i,j)} = \mathbb{E}[\Delta_i\Delta_j]$

$$q_{\mathsf{A}\Delta,(i,j)} = \mathbb{E}\left[\mathsf{A}_i - \left(\sqrt{\hat{P}}\hat{\zeta}_0\right)^{-1}(\mathsf{z}_j - \mathsf{z}_{j-1})\right] = -\left(\sqrt{\hat{P}}\hat{\zeta}_0\right)^{-1}\mathbb{E}[\mathsf{A}_i\,(\tilde{\mathsf{x}}_j - \tilde{\mathsf{x}}_{j-1})] \tag{A.17}$$

$$\begin{aligned} q_{\Delta,(i,j)} &= \left(\sqrt{\hat{P}}\hat{\zeta}_0\right)^{-2}\mathbb{E}[(\mathsf{z}_i - \mathsf{z}_{i-1})(\mathsf{z}_j - \mathsf{z}_{j-1})] \\ &= \left(\sqrt{\hat{P}}\hat{\zeta}_0\right)^{-2}\{\mathbb{E}[(\hat{\mathsf{n}}_i - \hat{\mathsf{n}}_{i-1})(\hat{\mathsf{n}}_j - \hat{\mathsf{n}}_{j-1})] + \mathbb{E}[(\tilde{\mathsf{x}}_i - \tilde{\mathsf{x}}_{i-1})(\tilde{\mathsf{x}}_j - \tilde{\mathsf{x}}_{j-1})]\}. \end{aligned} \tag{A.18}$$

Here, (A.17) stems from the fact that the Gaussian noise $\hat{\mathsf{n}}(t)$ is independent of the signal such that only $\tilde{\mathsf{x}}(t)$ contributes to a correlation between **A** and **D**. We can write for the intersymbol interference (ISI) at time T_j

$$\tilde{\mathsf{x}}_j = (-1)^j\left[\ldots - \tilde{g}(\mathsf{L}_{j-2}^{j}) + \tilde{g}(\mathsf{L}_{j-1}^{j}) - \tilde{g}(\mathsf{L}_{j}^{j}) + \tilde{g}(\mathsf{L}_{j+1}^{j+1}) - \tilde{g}(\mathsf{L}_{j+1}^{j+2}) + \tilde{g}(\mathsf{L}_{j+1}^{j+3})\ldots\right] \tag{A.19}$$

where $\mathsf{L}_{j+l}^{j+m} = \sum_{k=j+l}^{j+m}\mathsf{A}_k$. For obtaining $\mathbf{Q}_{\mathsf{A}\Delta}$ we introduce with $n = m - l + 1$ and $m \geq l$

$$\xi_n = \mathbb{E}[\mathsf{A}_i]\,\mathbb{E}\left[\tilde{g}\left(\mathsf{L}_{j+l}^{j+m}\right)\right], i \notin [j+l, j+m] \tag{A.20}$$

$$\nu_n = \mathbb{E}\left[\mathsf{A}_i\tilde{g}\left(\mathsf{L}_{j+l}^{j+m}\right)\right],\ j+l \leq i \leq j+m \tag{A.21}$$

which we use to express the (i,j)th entries of $\mathbf{Q}_{\mathsf{A}\Delta}$ by

$$q_{\mathsf{A}\Delta,(i,i)} = (-1)^i\left[\ldots - \nu_3 - \xi_3 + \nu_2 + \xi_2 - \xi_1 + \xi_1 - \xi_2 - \nu_2 + \xi_3 + \nu_3\ldots\right] = q_1 = 0 \tag{A.22}$$

$$q_{\mathsf{A}\Delta,(i,i+1)} = (-1)^{i+1}\left[\ldots - 2\nu_3 + 2\nu_2 - \nu_1 + \xi_1 - 2\xi_2 + 2\xi_3\ldots\right] = (-1)^{i+1}q_2 \tag{A.23}$$

$$q_{\mathsf{A}\Delta,(i,i-1)} = (-1)^{i-1}\left[\ldots - 2\xi_3 + 2\xi_2 - \xi_1 + \nu_1 - 2\nu_2 + 2\nu_3\ldots\right] = (-1)^{i-1}(-q_2). \tag{A.24}$$

Table A.1 summarizes the weights of ν_n and ξ_n for some values of j. Exploiting the symmetry in (A.22) to (A.24) as well as in Table A.1, we obtain

$$\mathbf{Q}_{\mathsf{A}\Delta} = \begin{bmatrix} 0 & q_2 & q_3 & \ldots & q_K \\ q_2 & 0 & -q_2 & \ldots & -q_{K-1} \\ -q_3 & -q_2 & 0 & \ldots & q_{K-2} \\ & \vdots & & \ddots & \vdots \\ (-1)^K q_K & (-1)^K q_{K-1} & (-1)^K q_{K-2} & \ldots & 0 \end{bmatrix} \tag{A.25}$$

which can be evaluated by numerically solving the integrals in (A.20) and (A.21) using the probability distributions of A_i and L_{j+l}^{j+m}, cf. (3.5) and (3.8). From (A.18) we obtain $\mathbf{Q}_{\Delta}$ as

$$\mathbf{Q}_{\Delta} = a^2 \frac{\sigma_{\hat{\mathsf{n}}}^2}{\sigma_{\tilde{\mathsf{S}}}^2} \mathbf{R}_{\Delta} + \left(\sqrt{\hat{P}} \hat{\zeta}_0 \right)^{-2} \mathbf{R}_{\tilde{\mathsf{x}}} \tag{A.26}$$

where $\mathbf{R}_{\Delta}$ is given in (3.61). The elements of $\mathbf{R}_{\tilde{\mathsf{x}}}$ are given by

$$r_{\tilde{\mathsf{x}},(i,j)} = 2\,\mathbb{E}[\tilde{\mathsf{x}}_i \tilde{\mathsf{x}}_j] - \mathbb{E}[\tilde{\mathsf{x}}_i \tilde{\mathsf{x}}_{j-1}] - \mathbb{E}[\tilde{\mathsf{x}}_{i-1} \tilde{\mathsf{x}}_j]. \tag{A.27}$$

We see from (A.19) that $\mathbb{E}[\tilde{\mathsf{x}}_i \tilde{\mathsf{x}}_j]$ yields sums of expectations $\mathbb{E}\left[\tilde{g}\left(\mathsf{L}_{i+l_1}^{i+m_1}\right) \tilde{g}\left(\mathsf{L}_{j+l_2}^{j+m_2}\right)\right]$. Within this expression a number n of the A_k in the two sums $\mathsf{L}_{i+l_1}^{i+m_1}$ and $\mathsf{L}_{j+l_2}^{j+m_2}$ coincide whereas w and p summands A_k are unique to the first and the second sum, respectively. Which A_k coincide and which not depends on the choice of $i-j$, $m_{1/2}$, and $l_{1/2}$. Therefore, we define three random variables W_n, X_w, and Y_p. Each of these random variables is the sum over disjoint sets of summands A_k of cardinality n, w, and p, respectively. This yields for the expectation above $\mathbb{E}\left[\tilde{g}\left(\mathsf{W}_n + \mathsf{X}_w\right) \tilde{g}\left(\mathsf{W}_n + \mathsf{Y}_p\right)\right]$ where $p(\mathsf{W}_n, \mathsf{X}_w, \mathsf{Y}_p) = p(\mathsf{W}_n) p(\mathsf{X}_w) p(\mathsf{Y}_p)$ as the A_k are independent. Using (3.8), we now can numerically evaluate $\mathbb{E}\left[\tilde{g}\left(\mathsf{W}_n + \mathsf{X}_m\right) \tilde{g}\left(\mathsf{W}_n + \mathsf{Y}_p\right)\right]$, which yields $\mathbb{E}[\tilde{\mathsf{x}}_i \tilde{\mathsf{x}}_j]$, $r_{\tilde{\mathsf{x}},(i,j)}$, and eventually $\mathbf{Q}_{\Delta}$.

Finally, $\mathbf{Q}_{\text{err}}$ can be computed with $\mathbf{Q}_{\mathsf{A}\Delta}$, $\mathbf{Q}_{\Delta}$, and $\mathbf{Q}_{\mathsf{A}}$. Fig. A.6 depicts the difference between the two bounds in (3.69) and (3.70), i.e., $\frac{1}{2}(\log\det \mathbf{R}_{\text{err}} - \log\det \mathbf{Q}_{\text{err}})$. It can be seen that it is always positive, which indicates that the inequality in (3.70) holds. The bounds are tighter for low SNR values and larger κ. This can be attributed to the fact that setting $\mathbf{Q}_{\mathsf{A}\Delta} = 0$ as done for obtaining (3.70) is equivalent to neglecting the ISI, which is less dominant at low SNR and large κ. The non-smoothness of the curves in Fig. A.6b) can be attributed to inaccuracies in the numerical integration. In contrary to the sine-shaped transition waveform, the error pulse $\tilde{g}(t)$ for the linear transition is much less smooth, cf. Fig. 3.6a), leading to an increasing integration error, which dominates for high ρ and large κ.

A.6 Analytical Expression for the Signal Parameters Depending on the Transition Waveform

In the following, the full analytical expressions of the model parameters Γ_0, Γ_2, u, and $\hat{\zeta}_0$ for the sine, linear, and polynomial waveforms are listed. Where compactly expressable, also $G(\omega)$ and $\hat{g}(t)$ or $\tilde{g}(t)$ are given. Hereby, $\gamma \approx 0.5772$ is the Euler-Mascheroni constant.

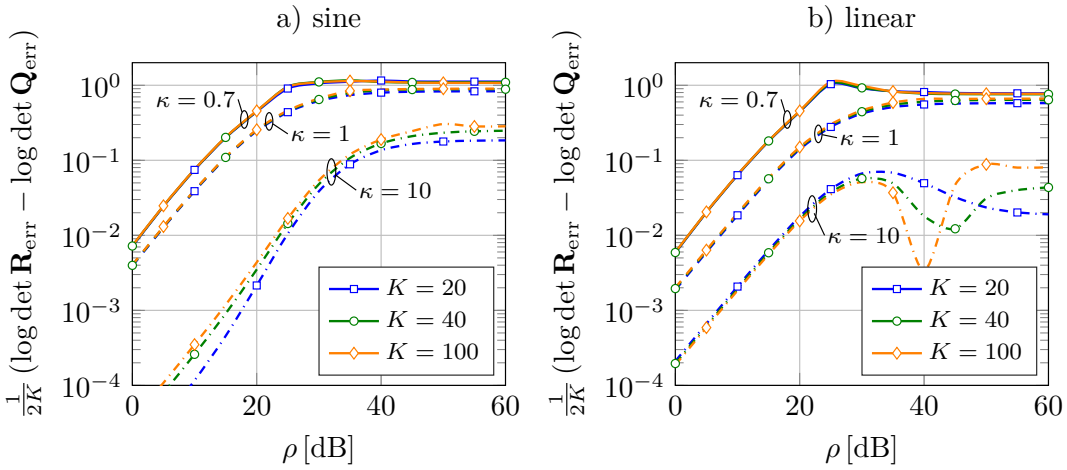

Fig. A.6. Difference between the two bounds in (3.69) and (3.70) normalized by the blocklength K for different $\kappa = {}^{W}/_{\lambda}$ and K for a) sine-shaped and b) linear transition waveforms

a) sine-shaped transition

$$G(\omega) = -j\frac{\pi^2\left(1+e^{-j\omega\beta}\right)}{\omega(\pi^2-\omega^2\beta^2)} \tag{A.28}$$

$$\hat{g}(t) = 1+\frac{1}{\pi}\left[\mathrm{Si}\left(\frac{\pi t}{\beta}\right)+\mathrm{Si}\left(\frac{\pi t}{\beta}-\pi\right)-\frac{\cos\left(\frac{\pi t}{\beta}\right)}{2}\left[\mathrm{Si}\left(\frac{2\pi t}{\beta}\right)-\mathrm{Si}\left(\frac{2\pi t}{\beta}-2\pi\right)\right)\right.$$
$$\left.+\frac{\sin\left(\frac{\pi t}{\beta}\right)}{2}\left[\mathrm{Ci}\left(\frac{2\pi t}{\beta}\right)-\mathrm{Ci}\left(\frac{2\pi t}{\beta}-2\pi\right)-\ln\left(\frac{\pi t}{\beta}\right)+\ln\left(\frac{\pi t}{\beta}-\pi\right)\right]\right] \tag{A.29}$$

$$\Gamma_0 = \frac{-3\gamma-3\log(2\pi)+3\,\mathrm{Ci}(2\pi)-\pi^2+4\pi\mathrm{Si}(\pi)-\pi\mathrm{Si}(2\pi)}{2\pi}\beta \approx 0.2601\beta \tag{A.30}$$

$$\Gamma_2 = \frac{\pi}{2\beta}\left(\pi^2-\gamma-\log(2\pi)-\pi\mathrm{Si}(2\pi)+\mathrm{Ci}(2\pi)\right) \approx \frac{1.1689}{\beta} \tag{A.31}$$

$$u = \frac{2\mathrm{Si}(\pi)+\mathrm{Si}(2\pi)}{2\pi}\sqrt{\hat{P}} \approx 0.81\sqrt{\hat{P}} \tag{A.32}$$

$$\hat{\zeta}_0 = \frac{\mathrm{Si}(\pi)}{\beta} \approx \frac{1.8519}{\beta} \tag{A.33}$$

b) linear transition

$$G(\omega) = \frac{2}{\omega^2\beta}\left(1-e^{-j\omega\beta}\right) \tag{A.34}$$

$$\tilde{g}(t) = \frac{2\left(t\mathrm{Si}\left(\frac{\pi t}{\beta}\right)-(t-\beta)\mathrm{Si}\left(\frac{\pi t}{\beta}-\pi\right)+\frac{\pi}{2}(|t-\beta|-|t|)+\frac{2\beta}{\pi}\cos\left(\frac{\pi t}{\beta}\right)\right)}{\beta\pi} \tag{A.35}$$

$$\Gamma_0 = \frac{2\beta}{3\pi^3}\left(8+2\pi^3\mathrm{Si}(\pi)-\pi^2(\pi^2+2)\right) \approx 0.1224\beta \tag{A.36}$$

$$\Gamma_2 = \frac{4}{\beta\pi}\left(4+\pi^2-2\pi\mathrm{Si}(\pi)\right) \approx \frac{2.8438}{\beta} \tag{A.37}$$

$$u = \frac{2}{\pi}\left(\mathrm{Si}(\pi) - \frac{2}{\pi}\right)\sqrt{\hat{P}} \approx 0.77\sqrt{\hat{P}} \tag{A.38}$$

$$\hat{\zeta}_0 = \frac{4}{\pi\beta}\mathrm{Si}\left(\frac{\pi}{2}\right) \approx \frac{1.7453}{\beta} \tag{A.39}$$

c) polynomial transition $2B+1=3$

$$\hat{g}(t) = \frac{2}{\pi^4}\left[\pi^3\left(\frac{t^2}{\beta^2}\left(3-\frac{2t}{\beta}\right)\mathrm{Si}\left(\frac{\pi t}{\beta}\right) - \left(\frac{t}{\beta}-1\right)^2\left(\frac{2t}{\beta}+1\right)\mathrm{Si}\left(\pi-\frac{\pi t}{\beta}\right)\right)\right.$$
$$\left.+2\pi\left(1-\frac{2t}{\beta}\right)\sin\left(\frac{\pi t}{\beta}\right) + \left(\pi^2\left(1-\frac{4t}{\beta}\left(\frac{t}{\beta}-1\right)\right)+8\right)\cos\left(\frac{\pi t}{\beta}\right)\right] \tag{A.40}$$

$$\Gamma_0 = -\frac{18\beta\left(-2\pi^7\mathrm{Si}(\pi)+\pi^8+2\pi^6-4\pi^4-64\pi^2-640\right)}{35\pi^7} \approx 0.2446\beta \tag{A.41}$$

$$\Gamma_2 = \frac{24\left(-2\pi^5\mathrm{Si}(\pi)+\pi^6+2\pi^4+16\pi^2+96\right)}{5\pi^5\beta} \approx \frac{4.3395}{\beta} \tag{A.42}$$

$$\frac{u}{\sqrt{\hat{P}}} = \frac{2\left(\pi^3\mathrm{Si}(\pi)-\pi^2-8\right)}{\pi^4} \approx 0.8121 \tag{A.43}$$

$$\hat{\zeta}_0 = \frac{1}{\beta}\frac{12\left(\frac{1}{2}\pi^2\mathrm{Si}\left(\frac{\pi}{2}\right)-2\right)}{\pi^3} \approx \frac{1.8439}{\beta} \tag{A.44}$$

d) polynomial transition $2B+1=5$

$$\Gamma_0 = -\frac{100\beta}{231\pi^{11}}(-2\pi^{11}\mathrm{Si}(\pi)+\pi^{12}+2\pi^{10}-4\pi^8+48\pi^6-10944\pi^4$$
$$+258048\pi^2-1741824) \approx 0.3399\beta \tag{A.45}$$

$$\Gamma_2 = \frac{40}{7\pi^9\beta}\left(-2\pi^9\mathrm{Si}(\pi)+\pi^{10}+2\pi^8-4\pi^6+1056\pi^4-23040\pi^2+161280\right) \approx \frac{6.7321}{\beta} \tag{A.46}$$

$$\frac{u}{\sqrt{\hat{P}}} = \frac{2\left(\pi^5\mathrm{Si}(\pi)-\pi^2\left(\pi^2-22\right)-288\right)}{\pi^6} \approx 0.8289 \tag{A.47}$$

$$\hat{\zeta}_0 = \frac{1}{\beta}\frac{60\left(\frac{1}{8}\pi^4\mathrm{Si}\left(\frac{\pi}{2}\right)+\frac{\pi^2}{2}-12\right)}{\pi^5} \approx \frac{1.88721}{\beta} \tag{A.48}$$

e) polynomial transition $2B+1=7$

$$\Gamma_0 = -\frac{490\beta}{1287\pi^{15}}(-2\pi^{15}\mathrm{Si}(\pi)+\pi^{16}+2\pi^{14}-4\pi^{12}+48\pi^{10}-1440\pi^8-1566720\pi^6$$
$$-247173120\pi^4+5474304000\pi^2-28466380800) \approx 0.4123\beta \approx \tag{A.49}$$

$$\Gamma_2 = \frac{2800}{429\pi^{13}\beta}(-2\pi^{13}\mathrm{Si}(\pi)+\pi^{14}+2\pi^{12}-4\pi^{10}+48\pi^8+122112\pi^6+16224768\pi^4$$
$$-365783040\pi^2+1916006400) \approx \frac{9.0052}{\beta} \tag{A.50}$$

$$\frac{u}{\sqrt{\hat{P}}} = \frac{2\pi^7\mathrm{Si}(\pi)-2\left(28800-2832\pi^2-2\pi^4+\pi^6\right)}{\pi^8} \approx 0.8384 \tag{A.51}$$

$$\hat{\zeta}_0 = \frac{1}{\beta} \frac{280\left(\frac{1}{32}\pi^6 \mathrm{Si}\left(\frac{\pi}{2}\right) + \frac{\pi^4}{8} + 21\pi^2 - 240\right)}{\pi^7} \approx \frac{1.9116}{\beta} \tag{A.52}$$

Appendix Chapter 4

B

B.1 Maximizing the Average Sample Flipping Probability

Our goal is to maximize (4.38) given the partly non-linear conditions in (4.35). In what follows, we will abbreviate the condition $\mathsf{x}_n = 1$ with $x_{1,n}$. Since we use a probability density function (pdf) with two mass points at z_1 and z_2 with weights p_{z_1} and p_{z_2}, we have

$$p_{z_1} + p_{z_2} = 1 \qquad \rightarrow\ p_{z_2} = 1 - p_{z_1} \tag{B.1}$$

$$p_{z_1} z_1 + p_{z_2} z_2 = \mathbb{E}[\mathsf{z}_{\mathrm{ISI}} | x_{1,n}] \qquad \rightarrow\ z_2 = \frac{\mathbb{E}[\mathsf{z}_{\mathrm{ISI}} | x_{1,n}] - p_{z_1} z_1}{1 - p_{z_1}} \tag{B.2}$$

$$p_{z_1} z_1^2 + p_{z_2} z_2^2 = \mathbb{E}[z_{\mathrm{ISI}}^2 | x_{1,n}] \qquad \rightarrow\ z_2^2 = \frac{\mathbb{E}[z_{\mathrm{ISI}}^2 | x_{1,n}] - p_{z_1} z_1^2}{1 - p_{z_1}}. \tag{B.3}$$

Combining (B.2) and (B.3) yields

$$z_1^2 - 2\mathbb{E}[\mathsf{z}_{\mathrm{ISI}} | x_{1,n}] z_1 + \mathbb{E}[z_{\mathrm{ISI}}^2 | x_{1,n}] - \frac{\mathrm{Var}(\mathsf{z}_{\mathrm{ISI}} | x_{1,n})}{p_{z_1}} = 0 \tag{B.4}$$

where $\mathrm{Var}(\mathsf{z}_{\mathrm{ISI}} | x_{1,n}) = \mathbb{E}[z_{\mathrm{ISI}}^2 | x_{1,n}] - \mathbb{E}[\mathsf{z}_{\mathrm{ISI}} | x_{1,n}]^2$. Solving (B.4) yields

$$z_1 = \mathbb{E}[\mathsf{z}_{\mathrm{ISI}} | x_{1,n}] \pm \sqrt{\mathrm{Var}(\mathsf{z}_{\mathrm{ISI}} | x_{1,n}) \left(\frac{1}{p_{z_1}} - 1 \right)} \tag{B.5}$$

$$z_2 = \mathbb{E}[\mathsf{z}_{\mathrm{ISI}} | x_{1,n}] \mp \frac{p_{z_1}}{1 - p_{z_1}} \sqrt{\mathrm{Var}(\mathsf{z}_{\mathrm{ISI}} | x_{1,n}) \left(\frac{1}{p_{z_1}} - 1 \right)}. \tag{B.6}$$

The resulting cost function becomes a function solely of p_{z_1}

$$p_{b,\mathrm{UB}}(p_{z_1}) = \frac{1}{2} \left(p_{z_1} \operatorname{erfc} \left(\frac{h_0 + \mathbb{E}[\mathsf{z}_{\mathrm{ISI}} | x_{1,n}] - \sqrt{\mathrm{Var}(\mathsf{z}_{\mathrm{ISI}} | x_{1,n}) \left(\frac{1}{p_{z_1}} - 1 \right)}}{\sqrt{2} \sigma_{\hat{\mathsf{n}}}} \right) \right.$$
$$\left. + (1 - p_{z_1}) \operatorname{erfc} \left(\frac{h_0 + \mathbb{E}[\mathsf{z}_{\mathrm{ISI}} | x_{1,n}] + \frac{p_{z_1}}{1 - p_{z_1}} \sqrt{\mathrm{Var}(\mathsf{z}_{\mathrm{ISI}} | x_{1,n}) \left(\frac{1}{p_{z_1}} - 1 \right)}}{\sqrt{2} \sigma_{\hat{\mathsf{n}}}} \right) \right) \tag{B.7}$$

where we assumed without implying an additional restriction the order $z_1 < z_2$. Then, $p^*_{z_1}$ can be determined by numerically searching for the maximum of (B.7) or by taking the derivative and solving $\frac{\partial p_{b,\mathrm{UB}}(p_{z_1})}{\partial p_{z_1}} = 0$ numerically. If the latter yields that $z_1^* < Z_1$, we have to se $z_1^* = Z_1$ in order to comply with (4.35) since $\mathrm{erfc}(x)$ is monotonically decreasing in x. From (B.4) we then have

$$p^*_{z_1} = \frac{\mathrm{Var}(\mathsf{z}_{\mathrm{ISI}}|x_{1,n})}{Z_1^2 - 2\mathbb{E}[\mathsf{z}_{\mathrm{ISI}}|x_{1,n}]Z_1 + \mathbb{E}[z^2_{\mathrm{ISI}}|x_{1,n}]}. \tag{B.8}$$

Appendix Chapter 5

C

C.1 Minimum Intermediate Frequency

From the observations made w.r.t. Fig. 5.5 it can be deduced that in order to resolve all ambiguities given a sufficiently high oversampling factor, it must be avoided that more than one transition starts and ends in the same quadrant. In what follows we consider phase changes counterclockwise, i.e., $n_{\mathrm{IF}} > 0$, since Δf is already positive. However, the setup is symmetric w.r.t. $-\Delta f$, i.e., the same results that are obtained for $f_{\mathrm{IF},1} = \Delta f + \frac{n_{\mathrm{IF}}}{T_{\mathrm{symb}}}$ occur for $f_{\mathrm{IF},2} = -2\Delta f - \frac{n_{\mathrm{IF}}}{T_{\mathrm{symb}}}$.

According to the definitions in (5.2), (5.3), (5.6), (5.23), and (5.24) the phase changes introduced by the first and second symbol in $\mathcal{A}$, $\alpha_0 = -(M_{\mathrm{cpm}} - 1)$ and $\alpha_1 = \alpha_0 + 2$, are

$$\Delta\psi_0 = \pi h_{\mathrm{cpm}}\alpha_0 + 2\pi f_{\mathrm{IF}} T_{\mathrm{symb}} = 2\pi h_{\mathrm{cpm}} c_2 \tag{C.1}$$

$$\Delta\psi_1 = \pi h_{\mathrm{cpm}}\alpha_1 + 2\pi f_{\mathrm{IF}} T_{\mathrm{symb}} = 2\pi h_{\mathrm{cpm}}(1 + c_2). \tag{C.2}$$

Given that $0 \leq \phi_0 \leq \pi/4$, cf. (5.19), and with $n_{\mathrm{IF}} > 0$ the constellation point $\phi_0 + \Delta\psi_0$ can remain in the first quadrant while $\phi_0 + \Delta\psi_1$ has to be in the second quadrant. Thus

$$\frac{\pi}{2} < \phi_0 + 2\pi h_{\mathrm{cpm}}(1 + c_2) \tag{C.3}$$

$$= \frac{\pi}{4\lceil 2^{c_1-2}\rceil} + 2\pi 2^{-c_1}(1 + c_2) \tag{C.4}$$

$$c_2 > \frac{1 - \frac{1}{2\lceil 2^{c_1-2}\rceil}}{2^{2-c_1}} - 1 = \begin{cases} \frac{1}{2^{3-c_1}} - 1, & c_1 = \{0, 1\} \\ 2^{c_1-2} - \frac{3}{2} & c_1 \geq 2 \end{cases}. \tag{C.5}$$

Since c_2 is an integer number

$$c_{2,\min} = \begin{cases} \left\lceil \frac{1}{2\,2^{2-c_1}} \right\rceil - 1, & c_1 = \{0, 1\} \\ \left\lceil 2^{c_1-2} - \frac{1}{2} \right\rceil - 1, & c_1 \geq 2 \end{cases} \tag{C.6}$$

where 2^{c_1-2} is already an integer number for $c_1 \geq 2$, such that

$$\left\lceil 2^{c_1-2} - \frac{1}{2} \right\rceil = 2^{c_1-2} = \frac{1}{4h_{\mathrm{cpm}}}. \tag{C.7}$$

For $c_1 < 2$, we have

$$\left\lceil \frac{1}{2\,2^{2-c_1}} \right\rceil = \left\lceil \frac{1}{2} \frac{1}{4h_{\mathrm{cpm}}} \right\rceil = \left\lceil \frac{1}{4h_{\mathrm{cpm}}} \right\rceil \tag{C.8}$$

since for $c_1 = \{0, 1\}$ we have $\frac{1}{4h_{\mathrm{cpm}}} < 1$ such that the factor $1/2$ does not change the result of the ceiling function. Therefore, (C.6) can be simplified to

$$c_{2,\min} = \left\lceil \frac{1}{4h_{\mathrm{cpm}}} \right\rceil - 1 \tag{C.9}$$

which yields

$$n_{\mathrm{IF},\min} = h_{\mathrm{cpm}} \left(\left\lceil \frac{1}{4h_{\mathrm{cpm}}} \right\rceil - 1 \right). \tag{C.10}$$

C.2 Maximum Intermediate Frequency

In order to determine when the phase rotation induced by the intermediate frequency (IF) is too fast for a given oversampling ratio (OSR) assuming a 1-bit quantizer and continuous phase frequency shift keying (CPFSK), we consider the frequency associated with the ith symbol in $\mathcal{A}$

$$f_i = \frac{h_{\mathrm{cpm}} \alpha_i}{2T_{\mathrm{symb}}} + \Delta f + \frac{n_{\mathrm{IF}}}{T_{\mathrm{symb}}} = \frac{h_{\mathrm{cpm}}}{T_{\mathrm{symb}}} \left(\frac{\alpha_i + M_{\mathrm{cpm}} - 1}{2} + c_2 \right). \tag{C.11}$$

Since failing to generate a sample while traversing a quadrant represents a loss of information, it is clearly undesirable if a quadrant is passed without sampling. For the minimum and maximum values of α_i in $\mathcal{A}$, cf. (5.3), this yields upper bounds on n_{IF} as

$$\Delta\psi = 2\pi f_i \frac{T_{\mathrm{symb}}}{M_{\mathrm{osr}}} < \frac{\pi}{2} \tag{C.12}$$

$$\begin{aligned} n_{\mathrm{IF}} &< \frac{M_{\mathrm{osr}}}{4} - h_{\mathrm{cpm}} \left(\frac{\alpha_i + M_{\mathrm{cpm}} - 1}{2} \right) \\ &= \begin{cases} \frac{M_{\mathrm{osr}}}{4} = n_{\mathrm{IF},\max,1}, & \alpha_i = -(M_{\mathrm{cpm}} - 1) \\ \frac{M_{\mathrm{osr}}}{4} - \frac{h_{\mathrm{cpm}}(M_{\mathrm{cpm}}-1)}{2} = n_{\mathrm{IF},\max,2}, & \alpha_i = M_{\mathrm{cpm}} - 1 \end{cases}. \end{aligned} \tag{C.13}$$

We evaluated the above expression for different OSRs and compared the results to the number of resolvable bits $\log_2 N_d$ for different M_{osr} and n_{IF} for $M_{\mathrm{cpm}} = \{4, 8, 16\}$. The results are depicted in Fig. C.1 to C.3. Hereby, the cases $M_{\mathrm{cpm}} = 4$, $M_{\mathrm{cpm}} h_{\mathrm{cpm}} = \{1, 2\}$ and $M_{\mathrm{cpm}} = 8$, $M_{\mathrm{cpm}} h_{\mathrm{cpm}} = 2$ are somewhat trivial. We will see in Fig. 5.7 that in these cases the minimum required OSR is constant and as soon as it is reached, all symbols can be always resolved independently of n_{IF}.

For all other cases we observe a periodic behavior over n_{IF}, which is not surprising since the phase is periodic w.r.t. modulo two. The period increases with the OSR. Inter-

estingly, we observe that the IFs between $n_{\mathrm{IF,max,2}}$ and $n_{\mathrm{IF,max,1}}$ perform particularly bad in these cases. We therefore conclude that our initial assumption was correct and that passing a quadrant without sampling – even for just one out of the M_{cpm} symbols, i.e., $\max_{\alpha_i}(2\pi f_i \frac{T_{\mathrm{symb}}}{M_{\mathrm{osr}}}) \geq \frac{\pi}{2}$ – represents a significant loss of information. However, if this is true for all symbols, i.e., $\min_{\alpha_i}(2\pi f_i \frac{T_{\mathrm{symb}}}{M_{\mathrm{osr}}}) \geq \frac{\pi}{2}$, the IF is again large enough such that distinction of all symbols can be possible.

We thus conclude that for CPFSK $n_{\mathrm{IF,max,2}}$ is a valid description for the IF beyond which increasing n_{IF} is not beneficial anymore – given that $n_{\mathrm{IF,max,2}}$ is larger or equal $n_{\mathrm{IF,min}}$. Therefore, the domain to search in order to find a suitable n_{IF} for CPFSK is

$$n_{\mathrm{IF,min}} \leq n_{\mathrm{IF,CPFSK}} < n_{\mathrm{IF,max,2}}. \tag{C.14}$$

Note that this does not mean that below $n_{\mathrm{IF,max,2}}$ all symbols can be resolved but that there is an IF below n_{IF} that achieves at least the same $\log_2(N_d)$ bits as any IF above.

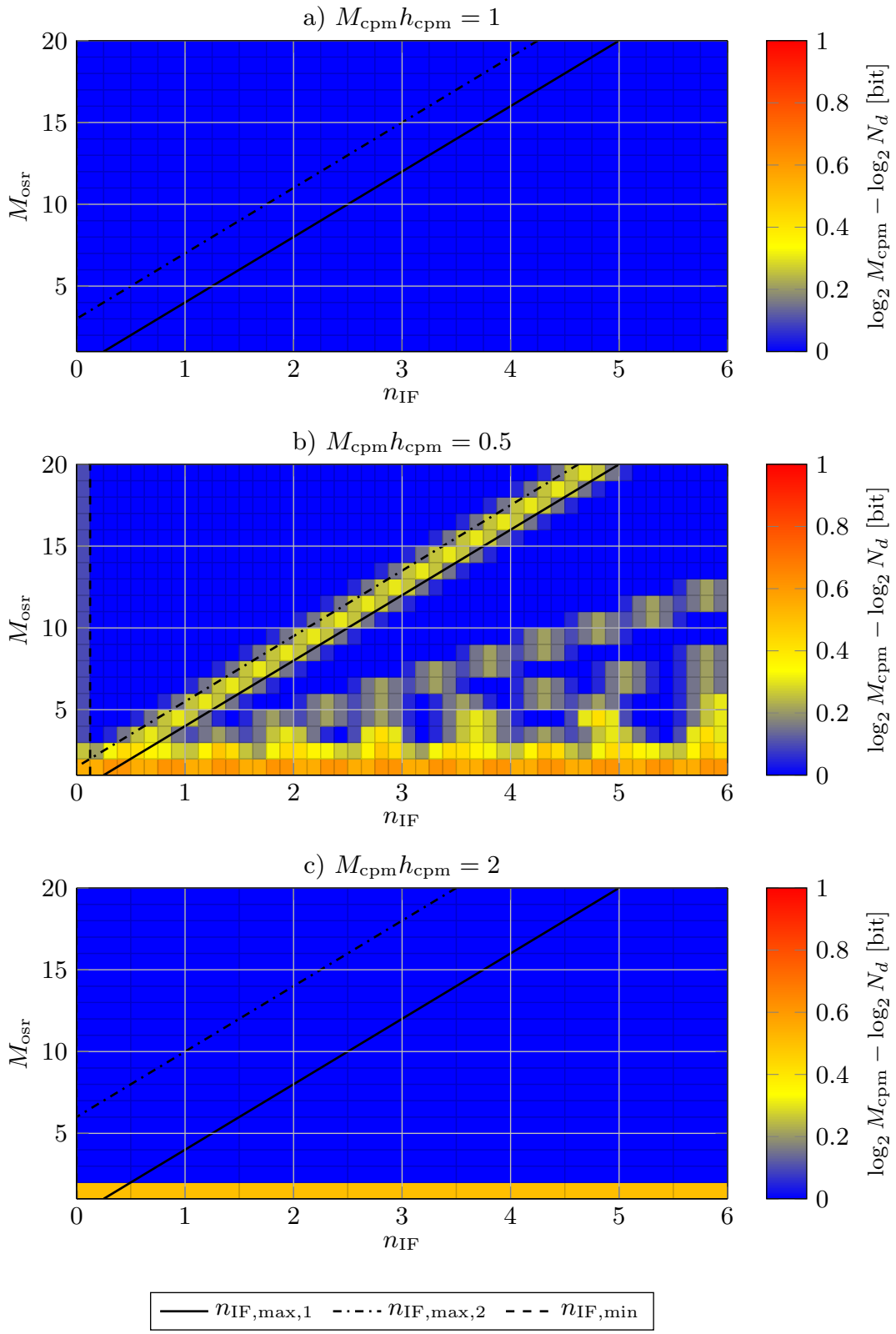

Fig. C.1. Difference between $\log_2(M_{\mathrm{cpm}})$ and $\log_2(N_d)$ for $M_{\mathrm{cpm}} = 4$ and different $M_{\mathrm{cpm}} h_{\mathrm{cpm}}$

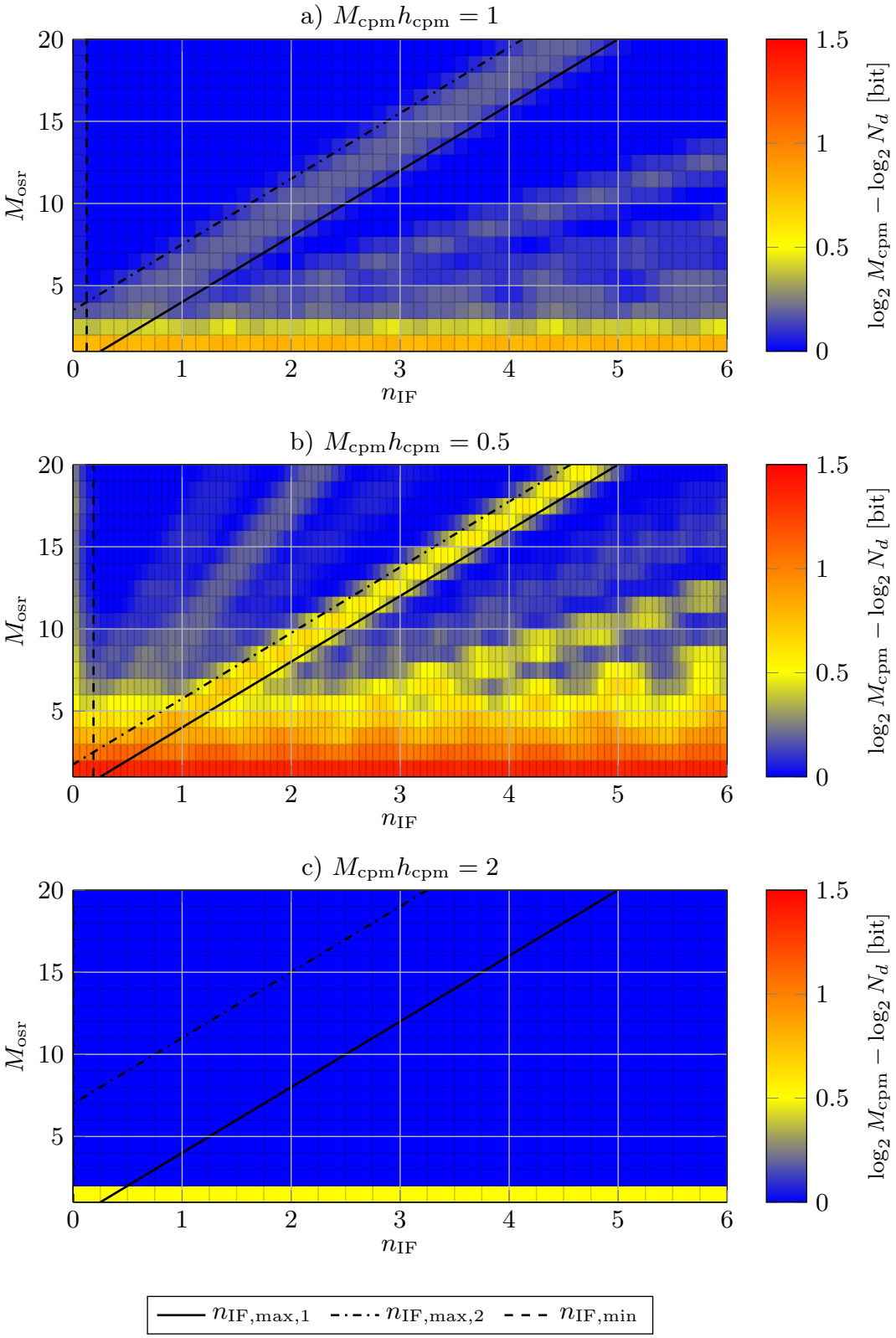

Fig. C.2. Difference between $\log_2(M_{\mathrm{cpm}})$ and $\log_2(N_d)$ for $M_{\mathrm{cpm}} = 8$ and different $M_{\mathrm{cpm}} h_{\mathrm{cpm}}$

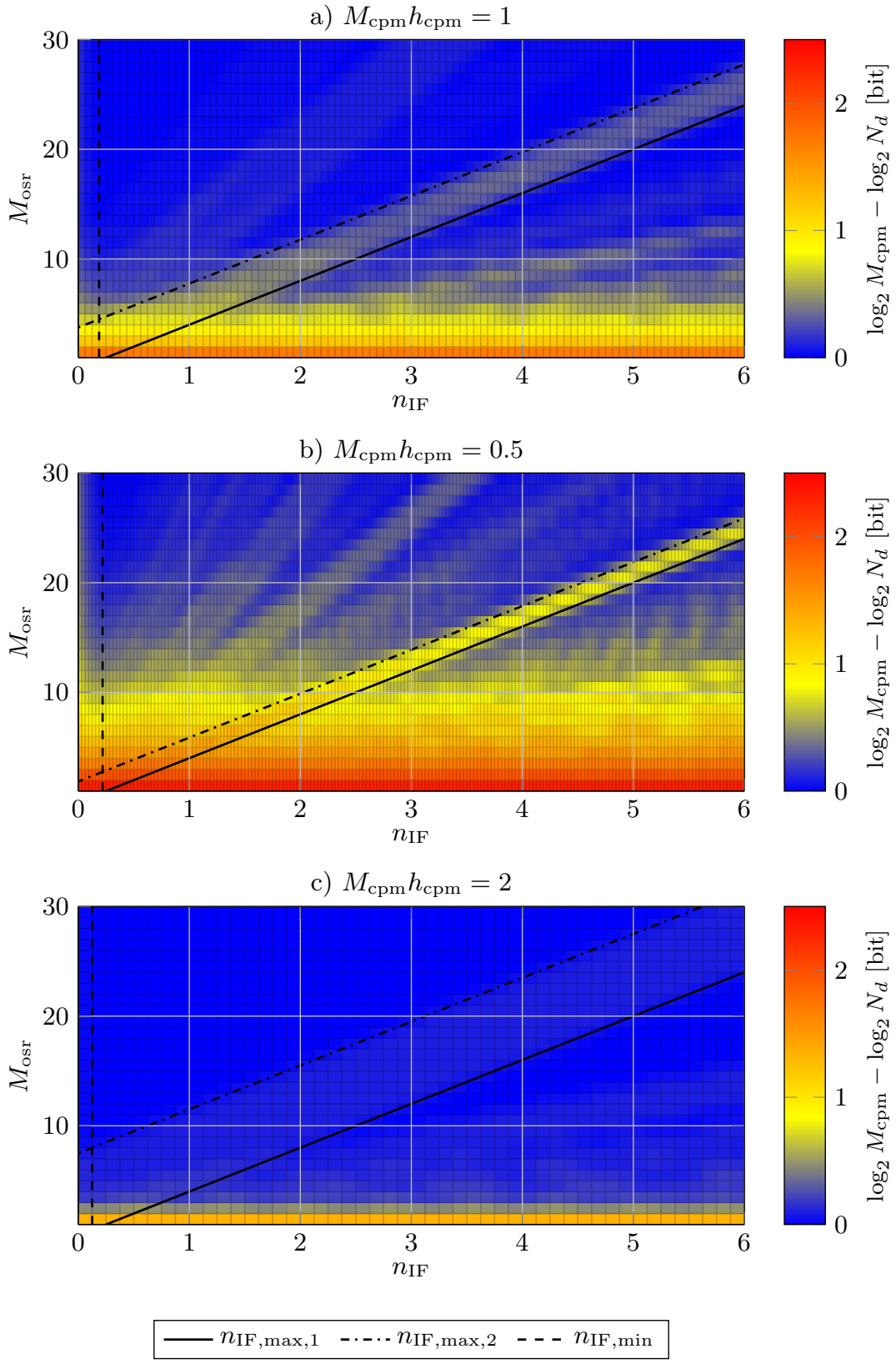

Fig. C.3. Difference between $\log_2(M_{\mathrm{cpm}})$ and $\log_2(N_d)$ for $M_{\mathrm{cpm}} = 16$ and different $M_{\mathrm{cpm}} h_{\mathrm{cpm}}$

C.3 Minimum Oversampling Ratio

For CPFSK, the phase changes linearly during a symbol, i.e., the instantaneous frequency per symbol is constant. This can be used to obtain an expression for the minimum OSR to distinguish all symbols. The incremental phase change $\Delta\psi_{\text{incr},i}$ during the sampling period $T_s = {}^{T_{\text{symb}}}/_{M_{\text{osr}}}$ caused by the ith symbol in $\mathcal{A}$ is

$$\Delta\psi_{\text{incr},i} = 2\pi f_i \frac{T_{\text{symb}}}{M_{\text{osr}}} = 2\pi \frac{h_{\text{cpm}}}{M_{\text{osr}}} \left(\frac{\alpha_i + M_{\text{cpm}} - 1}{2} + c_2 \right) \tag{C.15}$$

where the normalized frequency associated with the ith symbol is

$$\tilde{f}_i = f_i T_{\text{symb}} = h_{\text{cpm}} \left(\frac{\alpha_i + M_{\text{cpm}} - 1}{2} + c_2 \right). \tag{C.16}$$

There are $\lceil {}^{n_{\text{const}}}/_4 \rceil$ constellation points in the first quadrant, cf. (5.18), resulting in that many possible starting points[1]

$$\phi_{0,p} = \phi_0 + 2\pi(p-1)h_{\text{cpm}},\ p = 1, 2, ..., \lceil {}^{n_{\text{const}}}/_4 \rceil \tag{C.17}$$

for a phase trajectory. The time instant $t_{p,k,n}$ at which the phase trajectory corresponding to the ith symbol in $\mathcal{A}$ starting at the pth constellation point equals $n\frac{\pi}{2}$ is

$$t_{p,k,n} = \frac{n\frac{\pi}{2} - \phi_{0,p}}{\Delta\psi_{\text{incr},i}} = M_{\text{osr}} \frac{{}^{n}/_2 - {}^{\phi_{0,p}}/_{\pi}}{2\tilde{f}_i} \tag{C.18}$$

where the $t_{p,k,n}$ represent non-integer multiples of T_s. Thus, assuming the ith symbol started at t_0, $n\frac{\pi}{2}$ is reached at time $t_{p,k,n}T_s + t_0$. Consequently, if the OSR is smaller/larger than

$$\tilde{M}_{\text{osr}} = t_{p,k,n} \frac{2\tilde{f}_i}{{}^{n}/_2 - {}^{\phi_{0,p}}/_{\pi}} \tag{C.19}$$

the phase trajectory crosses the nth threshold before/after $t_{p,k,n}$. We thus write for the *minimum* OSR, for which a phase trajectory of symbol i starting at $\phi_{0,p}$ crosses $n\frac{\pi}{2}$ *after* the sth sample, $s \in \mathbb{N}$

$$\left\lceil s \frac{2\tilde{f}_i}{{}^{n}/_2 - {}^{\phi_{0,p}}/_{\pi}} + \epsilon \right\rceil \tag{C.20}$$

while the *maximum* OSR for symbol j to cross $n\frac{\pi}{2}$ *before* the sth sample is

$$\left\lfloor \frac{2s\tilde{f}_j}{{}^{n}/_2 - {}^{\phi_{0,p}}/_{\pi}} \right\rfloor. \tag{C.21}$$

[1] Considering starting points in the first quadrant is sufficient due to the assumption of a symmetric constellation diagram.

We hereby assume that a threshold belongs to the quadrant, which follows the threshold counterclockwise. The arbitrary small number ϵ in (C.20) avoids counting the thresholds as also belonging to the previous quadrant. All integer numbers within these boundaries are valid oversampling factors to distinguish the ith and the jth symbol at the sth sample and the nth threshold when starting at $\phi_{0,p}$. They are contained in the set

$$\mathcal{B}_{p,s,n,(i,j)} = \left\{ m \,\middle|\, m \in \mathbb{N} \wedge m \in \left[\left\lceil \frac{2s\tilde{f}_i}{\frac{n}{2} - \frac{\phi_{0,p}}{\pi}} + \epsilon \right\rceil, \left\lfloor \frac{2s\tilde{f}_j}{\frac{n}{2} - \frac{\phi_{0,p}}{\pi}} \right\rfloor \right] \wedge m \geq s \right\}. \tag{C.22}$$

Hereby, $i < j$ such that the symbols are ordered w.r.t. increasing frequency. Furthermore, $m \geq s$ since two symbols can only be distinguished by the sth sample if the temporal OSR is at least s.

Now the union over all thresholds n has to be taken. The largest threshold n_{max} that can be crossed is given by

$$n_{\text{max}} = \left\lfloor \frac{1}{\pi/2} \left(\phi_0 + 2\pi \left(\lceil n_{\text{const}}/4 \rceil - 1 \right) h_{\text{cpm}} + 2\pi f_{M_{\text{cpm}}} T_{\text{symb}} \right) \right\rfloor \tag{C.23}$$

using the frequency $f_{M_{\text{cpm}}}$ associated with the largest symbol in $\mathcal{A}$. Thus,

$$\mathcal{B}_{p,s,(i,j)} = \bigcup_{n=1}^{n_{\text{max}}} \mathcal{B}_{p,s,n,(i,j)}. \tag{C.24}$$

However, OSRs that distinguish at the sth sample at four (or 8, 12,...) consecutive thresholds need to be excluded since this represents a phase change of multiples of 2π, which cannot be resolved. Fortunately, one OSR m can only distinguish at consecutive thresholds. Denoting as $\text{count}_m(\cdot)$ the number of occurrences of element m in a set, we thus can write

$$\mathcal{E} = \left\{ m \,\middle|\, \text{count}_m \left(\mathcal{B}_{p,s,(i,j)} \right) \bmod 4 = 0 \right\} \tag{C.25}$$

for the set of OSRs that has to be excluded. Then the set

$$\mathcal{B}_{p,(i,j)} = \bigcup_{s=1}^{\infty} \left(\mathcal{B}_{p,s,(i,j)} \setminus \mathcal{E} \right) \tag{C.26}$$

contains all OSRs that can distinguish between the ith and the jth symbol when starting at $\phi_{0,p}$. The intersection over all symbols and all $\phi_{0,p}$ then yields the set

$$\mathcal{B} = \bigcap_{p=1}^{n_{\text{const}}/4} \left(\bigcap_{(i,j),i<j} \mathcal{B}_{p,(i,j)} \right) \tag{C.27}$$

of all OSRs that allow distinction between all symbols starting from any constellation point and

$$M_{\mathrm{osr,min}} = \min \mathcal{B} \tag{C.28}$$

is the corresponding minimum OSR. Note that $n_{\mathrm{IF}} \geq n_{\mathrm{IF,min}}$ is required for $\mathcal{B}$ to be non-empty. Then for numerical evaluation of (C.28), the union over s in (C.26) has not to be carried out until $s \to \infty$ but only until $\mathcal{B}$ is non-empty.

Bibliography

[AFAS16] F. Ahmed, M. Furqan, K. Aufinger, and A. Stelzer, "A SiGe-based broadband 100–180-GHz differential power amplifier with 11 dBm peak output power and >1.3THz GBW", in *Proc. Europ. Microwave Integrated Circuits Conf. (EuMIC)*, London, U.K., Oct. 2016, pp. 257–260 (cit. on p. 7).

[AV96] V. Anantharam and S. Verdú, "Bits through queues", *IEEE Trans. Inf. Theory*, vol. 42, no. 1, pp. 4–18, 1996 (cit. on pp. 25, 49).

[AZ12] J. B. Anderson and M. Zeinali, "Best rate 1/2 convolutional codes for turbo equalization with severe isi", in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, 2012, pp. 2366–2370 (cit. on pp. 89, 90).

[AAS86] J. Anderson, T. Aulin, and C. Sundberg, *Digital Phase Modulation*, ser. Applications of Communications Theory. Springer US, 1986 (cit. on pp. 71, 73, 74).

[ARO13] J. B. Anderson, F. Rusek, and V. Owall, "Faster-than-Nyquist signaling", *Proc. of the IEEE*, vol. 101, no. 8, pp. 1817–1830, Aug. 2013 (cit. on pp. 11, 13, 84, 86).

[ALV+06] D. Arnold, H.-A. Loeliger, P. Vontobel, a. Kavcic, and W. Z. W. Zeng, "Simulation-Based Computation of Information Rates for Channels With Memory", *IEEE Trans. Inf. Theory*, vol. 52, no. 8, pp. 3498–3508, 2006 (cit. on pp. 19, 20).

[ABG+10] G. Auer, O. Blume, V. Giannini, *et al.*, "Energy efficiency analysis of the reference systems, areas of improvements and target breakdown", EARTH Project, FP7-ICT, Project Deliverable 2.3, Nov. 2010 (cit. on p. 1).

[BCJR74] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate", *IEEE Trans. Inf. Theory*, vol. 20, no. 2, pp. 284–287, Mar. 1974 (cit. on pp. 10, 20).

[BHZ17] M. Bao, Z. S. He, and H. Zirath, "A 100–145 GHz area-efficient power amplifier in a 130 nm SiGe technology", in *Proc. Europ. Microwave Conf. (EuMC)*, Nuremberg, Germany, Oct. 2017, pp. 1017–1020 (cit. on p. 7).

[Big74] E. Biglieri, "On Glave's upper bound to the error probability for correlated digital signals (corresp.)", *IEEE Trans. Inf. Theory*, vol. 20, no. 1, pp. 115–119, Jan. 1974 (cit. on p. 58).

[Bus52] J. J. Bussgang, "Crosscorrelation functions of amplitude-distorted Gaussian signals", 1952 (cit. on p. 23).

[CT06] T. Cover and J. Thomas, *Elements of Information Theory, 2nd edition*. New York, U.S.A: Wiley & Sons, 2006 (cit. on pp. 41, 42).

[CL67] H. Cramer and M. R. Leadbetter, *Stationary and Related Stochastic Processes*. New York: Wiley, 1967 (cit. on pp. 43, 44).

[DJM+17] O. De Candido, H. Jedda, A. Mezghani, A. L. Swindlehurst, and J. A. Nossek, "Are traditional signal processing techniques rate maximizing in quantized SU-MISO systems?", in *Proc. IEEE Global Communications Conf. (Globecom)*, Singapore, Dec. 2017, pp. 1–6 (cit. on p. 23).

[dAL19] R. R. M. de Alencar and L. T. N. Landau, *Continuous phase modulation with faster-than-Nyquist signaling for channels with 1-bit quantization and oversampling at the receiver*, 2019. arXiv: `1910.06465 [cs.IT]` (cit. on pp. 85, 86, 91).

[DZZ19] R. Deng, J. Zhou, and W. Zhang, *Bandlimited communication with one-bit quantization and oversampling: Transceiver design and performance evaluation*, 2019. arXiv: `1912.11232 [cs.IT]` (cit. on pp. 18, 21).

[DMP07] S. Diggavi, M. Mitzenmacher, and H. Pfister, "Capacity upper bounds for deletion channels", in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Nice, France, Jun. 2007, pp. 1716–1720 (cit. on p. 30).

[Dob67] R. L. Dobrushin, "Shannon's theorems for channels with synchronization errors", *Probl. Peredachi Inf.*, vol. 3, no. 4, pp. 18–36, 1967 (cit. on pp. 30, 55).

[ET13] M. El Hefnawy and H. Taoka, "Overview of faster-than-Nyquist for future mobile communication systems", in *Proc. IEEE Vehicular Tech. Conf. (VTC Spring)*, Dresden, Germany, Jun. 2013, pp. 1–5 (cit. on p. 11).

[FRCE19] A. Ferchichi, S. U. Rehman, C. Carta, and F. Ellinger, "60-GHz SiGe-BiCMOS power amplifier with 14.7 dBm output power and 18 dB power gain", in *Proc. German Microwave Conf. (GeMiC)*, Stuttgart, Germany, Mar. 2019, pp. 229–231 (cit. on p. 7).

[FDE11] D. Fertonani, T. Duman, and M. Erden, "Bounds on the capacity of channels with insertions, deletions and substitutions", *IEEE Trans. Commun.*, vol. 59, no. 1, pp. 2–6, Jan. 2011 (cit. on pp. 30, 37, 54).

[FDC+19] G. P. Fettweis, M. Dörpinghaus, J. Castrillon, *et al.*, "Architecture and advanced electronics pathways toward highly adaptive energy-efficient computing", *Proc. of the IEEE*, vol. 107, no. 1, pp. 204–231, Jan. 2019 (cit. on pp. 1, 10).

[FuHLF13] G. P. Fettweis, N. ul Hassan, L. Landau, and E. Fischer, "Wireless interconnect for board and chip level", in *Proc. Design, Automation and Test in Europe (DATE)*, EDA Consortium, Grenoble, France, Mar. 2013, pp. 958–963 (cit. on pp. 1, 67).

[GBSS08] M. Gabrowska, M. Bossert, S. Shavgulidze, and S. Schober, "Serially concatenated space time convolutional codes and continuous phase modulation", *IEEE Trans. Commun.*, vol. 56, no. 9, pp. 1442–1450, 2008 (cit. on p. 90).

[Gal61] R. G. Gallager, "Sequential decoding for binary channels with noise and synchronization errors", Massachusetts Institute of Technology: Lincoln Laboratory, Tech. Rep., 1961 (cit. on p. 30).

[GLT00] A. Ganti, A. Lapidoth, and I. E. Telatar, "Mismatched decoding revisited: General alphabets, channels with memory, and the wide-band limit", *IEEE Trans. Inf. Theory*, vol. 46, no. 7, pp. 2315–2328, Nov. 2000 (cit. on p. 18).

[Gil93] E. N. Gilbert, "Increased information rate by oversampling", *IEEE Trans. Inf. Theory*, vol. 39, no. 6, pp. 1973–1976, 1993 (cit. on pp. 2, 15, 17).

[Gla72] F. Glave, “An upper bound on the probability of error due to intersymbol interference for correlated digital signals”, *IEEE Trans. Inf. Theory*, vol. 18, no. 3, pp. 356–363, May 1972 (cit. on pp. 56, 58).

[GBLV17] A. Gokceoglu, E. Björnson, E. G. Larsson, and M. Valkama, “Spatio-temporal waveform design for multiuser massive MIMO downlink with 1-bit receivers”, *IEEE J. Sel. Topics Signal Process.*, vol. 11, no. 2, pp. 347–362, Mar. 2017 (cit. on p. 23).

[Gra06] R. M. Gray, “Toeplitz and circulant matrices: A review”, *Found. Trends Commun. Inf. Theory*, vol. 2, no. 3, pp. 155–239, 2006 (cit. on p. 42).

[GS58] U. Grenander and G. Szegö, *Toeplitz Forms and Their Applications*. Berkeley, CA, U.S.A.: Univ. Calif. Press, 1958 (cit. on p. 42).

[GURP16] Z. Griffith, M. Urteaga, P. Rowell, and R. Pierson, “71–95 GHz (23–40% PAE) and 96–120 GHz (19–22% PAE) high efficiency 100–130 mW power amplifiers in InP HBT”, in *Proc. IEEE MTT-S International Microwave Symposium (IMS)*, San Francisco, USA, May 2016, pp. 1–4 (cit. on pp. 7, 93).

[HLF14a] T. Hälsig, L. Landau, and G. Fettweis, “Information rates for faster-than-Nyquist signaling with 1-bit quantization and oversampling at the receiver”, in *Proc. IEEE Vehicular Tech. Conf. (VTC Spring)*, IEEE, 2014 (cit. on p. 21).

[HLF14b] ——, “Spectral efficient communications employing 1-bit quantization and oversampling at the receiver”, in *Proc. IEEE Vehicular Tech. Conf. (VTC Fall)*, IEEE, Vancouver, Canada, 2014 (cit. on p. 20).

[HL17] T. Hälsig and B. Lankl, “Spatial oversampling in LOS MIMO systems with 1-bit quantization at the receiver”, in *Proc. Int. ITG Conf. on Systems, Communications and Coding (SCC)*, Hamburg, Germany, Feb. 2017 (cit. on p. 23).

[HBSU98] A. Häutle, M. Bossert, S. Shavgulidze, and N. Ugrelidze, “Generalized concatenated cpfsk modulation”, in *1998 Information Theory Workshop (Cat. No.98EX131)*, 1998, pp. 102–103 (cit. on p. 86).

[Imm90] K. Immink, “Runlength-limited sequences”, *Proc. of the IEEE*, vol. 78, no. 11, pp. 1745–1759, 1990 (cit. on pp. 12, 52).

[IR17] P. Indirayanti and P. Reynaert, “A 32 GHz 20 dBm-PSAT transformer-based Doherty power amplifier for multi-Gb/s 5G applications in 28 nm bulk CMOS”, in *Proc. IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*, Hawaii, USA, Jun. 2017, pp. 45–48 (cit. on pp. 6, 7).

[IN07] M. T. Ivrlac and J. A. Nossek, “On MIMO channel estimation with single-bit signal-quantization”, in *Proc. Int. ITG Workshop on Smart Antennas (WSA)*, Vienna, Austria, Feb. 2007 (cit. on p. 67).

[JDC+17] S. Jacobsson, G. Durisi, M. Coldrey, U. Gustavsson, and C. Studer, “Throughput analysis of massive MIMO uplink with low-resolution ADCs”, *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 4038–4051, Jun. 2017 (cit. on pp. 22, 23).

[JHAS15] J. Joung, C. K. Ho, K. Adachi, and S. Sun, “A survey on power-amplifier-centric techniques for spectrum- and energy-efficient wireless communications”, *IEEE Commun. Surveys Tuts.*, vol. 17, no. 1, pp. 315–333, Firstquarter 2015 (cit. on p. 6).

[JHS12] J. Joung, C. K. Ho, and S. Sun, “Green wireless communications: A power amplifier perspective”, in *Proc. Asia Pacific Signal and Information Processing Association Annual Summit and Conference*, Hollywood, California, USA, Dec. 2012, pp. 1–8 (cit. on pp. 6, 7).

[Kac43] M. Kac, “On the average number of real roots of a random algebraic equation”, *Bull. Amer. Math. Soc.*, vol. 49, no. 4, pp. 314–320, Apr. 1943 (cit. on p. 43).

[Kap93] J. Kapur, *Maximum-entropy Models in Science and Engineering*. New Dehli: Wiley Eastern, 1993 (cit. on pp. 12, 44).

[KGE15] A. Kipnis, A. J. Goldsmith, and Y. C. Eldar, “Sub-Nyquist sampling achieves optimal rate-distortion”, in *Proc. IEEE Inform. Theory Workshop (ITW)*, Jerusalem, Israel, Apr. 2015, pp. 1–5 (cit. on p. 15).

[KL10] T. Koch and A. Lapidoth, “Increased capacity per unit-cost by oversampling”, in *Proc. IEEE Conv. of Elect. and Electron. Engineers in Israel (IEEEI)*, Eilat, Israel, Nov. 2010, pp. 684–688 (cit. on pp. 2, 17).

[Kra06] M. Kratz, “Level crossings and other level functionals of stationary Gaussian processes”, *Probab. Surveys*, vol. 3, pp. 230–288, 2006 (cit. on p. 104).

[KF12] S. Krone and G. Fettweis, “Capacity of communications channels with 1-bit quantization and oversampling at the receiver”, in *Proc. IEEE Sarnoff Symposium*, Newark, USA, May 2012 (cit. on p. 17).

[KF10] ——, “Fading channels with 1-bit output quantization: Optimal modulation, ergodic capacity and outage probability”, in *Proc. IEEE Inform. Theory Workshop (ITW)*, Dublin, Ireland, Aug. 2010, pp. 1–5 (cit. on p. 16).

[LDdF18] L. T. N. Landau, M. Dörpinghaus, R. C. de Lamare, and G. P. Fettweis, “Achievable rate with 1-bit quantization and oversampling using continuous phase modulation-based sequences”, *IEEE Trans. Wireless Commun.*, vol. 17, no. 10, pp. 7080–7095, Oct. 2018 (cit. on pp. 21, 79).

[LDF15] L. Landau, M. Dörpinghaus, and G. Fettweis, “Communications employing 1-bit quantization and oversampling at the receiver: Faster-than-Nyquist signaling and sequence design”, in *Proc. IEEE Int. Conf. on Ubiquitous Wireless Broadband (ICUWB)*, Montreal, Canada, Oct. 2015 (cit. on p. 21).

[LF14a] L. Landau and G. Fettweis, “Information rates employing 1-bit quantization and oversampling at the receiver”, in *Proc. IEEE Int. Workshop on Signal Processing Advances in Wireless Commun. (SPAWC)*, Toronto, Canada, Jun. 2014, pp. 219–223 (cit. on p. 20).

[Lan17] L. T. Landau, *1-Bit quantization and oversampling at the receiver - how to benefit in terms of achievable rate*. Dresden: Vogt, 2017 (cit. on pp. 5, 10, 14).

[LDF18] L. T. Landau, M. Dörpinghaus, and G. P. Fettweis, “1-bit quantization and oversampling at the receiver: Sequence-based communication”, *EURASIP Journal on Wireless Communications and Networking*, vol. 2018, no. 1, p. 83, 2018 (cit. on pp. 2, 14, 21).

[LDF17] L. Landau, M. Dörpinghaus, and G. P. Fettweis, “1-bit quantization and oversampling at the receiver: Communication over bandlimited channels with noise”, *IEEE Commun. Lett.*, vol. 21, no. 5, pp. 1007–1010, 2017 (cit. on pp. 2, 21).

[LF14b] L. Landau and G. Fettweis, “On reconstructable ASK-sequences for receivers employing 1-bit quantization and oversampling”, in *Proc. IEEE Int. Conf. on Ultra-Wide-Band (ICUWB)*, Paris, France, Sep. 2014, pp. 180–184 (cit. on p. 21).

[Lap17] A. Lapidoth, *A Foundation in Digital Communication*, 2nd ed. Cambridge University Press, 2017 (cit. on p. 76).

[LNG+16] Y. Lin, V. K. Nguyen, J. Gao, *et al.*, “A wideband power amplifier with 13.2 dBm Psat and 19.5% PAE for 6094 GHz wireless communication systems in 90 nm CMOS”, in *Proc. IEEE Radio and Wireless Symposium (RWS)*, Austin, USA, Jan. 2016, pp. 95–98 (cit. on p. 7).

[Lug69] R. Lugannani, “Intersymbol interference and probability of error in digital systems”, *IEEE Trans. Inf. Theory*, vol. 15, no. 6, pp. 682–688, Nov. 1969 (cit. on p. 56).

[Man18] G. Manganaro, “Emerging data converter architectures and techniques”, in *Proc. IEEE Custom Integrated Circuits Conference (CICC)*, San Diego, USA, Apr. 2018, pp. 1–8 (cit. on pp. 8, 9).

[Mat73] J. Matthews, “Sharp error bounds for intersymbol interference”, *IEEE Trans. Inf. Theory*, vol. 19, no. 4, pp. 440–447, Jul. 1973 (cit. on pp. 56, 57).

[MLL18] J. C. Mayeda, D. Y. C. Lie, and J. Lopez, “A 24-28GHz reconfigurable CMOS power amplifier in 22nm FD-SOI for intelligent SoC applications”, in *Proc. International SoC Design Conference (ISOCC)*, Daegu, Korea, Nov. 2018, pp. 111–112 (cit. on p. 7).

[MLL17] ——, “A highly efficient and linear 15 GHz GaN power amplifier design for 5G communications”, in *Proc. Texas Symposium on Wireless and Microwave Circuits and Systems (WMCS)*, Waco, USA, Mar. 2017, pp. 1–4 (cit. on pp. 6, 7).

[Maz75] J. Mazo, “Faster-than-Nyquist signaling”, *Bell System Technical Journal*, vol. 54, pp. 1451–1462, 8 1975 (cit. on p. 11).

[McC13] E. W. McCune, “pPSK for bandwidth and energy efficiency”, in *Proc. Europ. Microwave Conf. (EuMC)*, Nuremberg, Germany, Oct. 2013, pp. 569–572 (cit. on p. 6).

[MAN10] A. Mezghani, F. Antreich, and J. A. Nossek, “Multiple parameter estimation with quantized channel output”, in *Proc. Int. ITG Workshop on Smart Antennas (WSA)*, Bremen, Germany, Feb. 2010, pp. 143–150 (cit. on p. 67).

[MN07] A. Mezghani and J. A. Nossek, “On ultra-wideband MIMO systems with 1-bit quantized outputs: Performance analysis and input optimization”, in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Nice, France, Jun. 2007, pp. 1286–1289 (cit. on p. 23).

[MN12] A. Mezghani and J. A. Nossek, “Capacity lower bound of MIMO channels with output quantization and correlated noise”, in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Boston, USA, Jul. 2012 (cit. on pp. 23, 69).

[MSPH14] J. Mo, P. Schniter, N. G. Prelcic, and R. W. Heath, “Channel estimation in millimeter wave MIMO systems with one-bit quantization”, in *Proc. Annual Asilomar Conf. on Signals, Systems, and Computers (ASILOMAR)*, Pacific Grove, USA, Nov. 2014, pp. 957–961 (cit. on p. 67).

[MH15] J. Mo and R. W. Heath, “Capacity analysis of one-bit quantized MIMO systems with transmitter channel state information”, *IEEE Trans. on Signal Process.*, vol. 63, no. 20, pp. 5498–5512, 2015 (cit. on pp. 23, 69).

[MCLH17] C. Mollén, J. Choi, E. G. Larsson, and R. W. Heath, “Uplink performance of wideband massive MIMO with one-bit ADCs”, *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 87–100, Jan. 2017 (cit. on p. 22).

[Mur18] B. Murmann, "*ADC performance survey 1997-2018*", `http://web.stanford.edu/~murmann/adcsurvey.html`, 2018 (cit. on pp. 2, 8, 9).

[Mur08] ——, “A/D converter trends: Power dissipation, scaling and digitally assisted architectures”, in *Proc. IEEE Custom Integrated Circuits Conference (CICC)*, San Jose, USA, Sep. 2008, pp. 105–112 (cit. on p. 8).

[Nar09] S. Naraghi, “Time-based analog to digital converters”, PhD thesis, University of Michigan, 2009 (cit. on p. 9).

[NAU19] A. A. Nawaz, J. D. Albrecht, and A. C. Ulusoy, “A 28/60 GHz dual-band power amplifier”, in *Proc. IEEE MTT-S International Microwave Symposium (IMS)*, Boston, USA, Jun. 2019, pp. 630–633 (cit. on p. 7).

[Nus15] H. Nuszkowski, *Digitale Signalübertragung: Grundlagen der digitalen Nachrichtenübertragungssysteme*. Jörg Vogt Verlag, Oct. 2015 (cit. on p. 10).

[PJJ+16] B. Park, S. Jin, D. Jeong, *et al.*, “Highly linear mm-wave CMOS power amplifier”, *IEEE Trans. Microw. Theory Tech.*, vol. 64, no. 12, pp. 4535–4544, Dec. 2016 (cit. on pp. 6, 7).

[PS17] H. Pirzadeh and A. L. Swindlehurst, “Analysis of MRC for mixed-ADC massive MIMO”, in *Proc. IEEE Int. Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, Curacao, Dutch Antilles, Dec. 2017, pp. 1–5 (cit. on p. 22).

[PS08] J. Proakis and M. Salehi, *Digital Communications, 5th edition*. McGraw-Hill Higher Education, 2008 (cit. on p. 90).

[RKD12] A. Rajaie, A. Khorshid, and A. Darwish, “Recent implementations of class-E power amplifiers”, in *Proc. Japan-Egypt Conf. on Electronics, Communications and Computers*, Alexandria, Egypt, Mar. 2012, pp. 26–30 (cit. on p. 6).

[Ric44] S. O. Rice, “Mathematical analysis of random noise”, *Bell System Technical Journal*, vol. 23, no. 3, pp. 282–332, 1944 (cit. on p. 43).

[RK18] T. Richardson and S. Kudekar, “Design of low-density parity check codes for 5G new radio”, *IEEE Commun. Mag.*, vol. 56, no. 3, pp. 28–34, Mar. 2018 (cit. on p. 10).

[Rim88] B. E. Rimoldi, “A decomposition approach to CPM”, *IEEE Trans. Inf. Theory*, vol. 34, no. 2, pp. 260–270, Mar. 1988 (cit. on p. 72).

[RPL14] C. Risi, D. Persson, and E. G. Larsson, "Massive MIMO with 1-bit ADC", *arXiv pre-print*, 2014 (cit. on p. 23).

[RA10] G. W. Roberts and M. Ali-Bakhshian, "A brief introduction to time-to-digital and digital-to-time converters", *IEEE Trans. Circuits Syst., II, Exp. Briefs*, vol. 57, no. 3, pp. 153–157, Mar. 2010 (cit. on p. 9).

[RPSN18] K. Roth, H. Pirzadeh, A. L. Swindlehurst, and J. A. Nossek, "A comparison of hybrid beamforming and digital beamforming with low-resolution ADCs for multiple users and imperfect CSI", *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 3, pp. 484–498, Jun. 2018 (cit. on p. 22).

[Rud87] W. Rudin, *Real and Complex Analysis*, Third. New York: McGraw-Hill Book Co., 1987, pp. xiv+416 (cit. on p. 31).

[RHS+19] J. Rusanen, M. Hietanen, A. Sethi, *et al.*, "Ka-band stacked power amplifier on 22 nm CMOS FDSOI technology utilizing back-gate bias for linearity improvement", in *Proc. IEEE Nordic Circuits and Systems Conference (NORCAS): NORCHIP and International Symposium of System-on-Chip (SoC)*, Helsinki, Finland, Oct. 2019, pp. 1–4 (cit. on p. 7).

[RA09] F. Rusek and J. B. Anderson, "Constrained capacities for faster-than-Nyquist signaling", *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 764–775, Feb. 2009 (cit. on p. 12).

[Sal68] B. Saltzberg, "Intersymbol interference error bounds with application to ideal bandlimited signaling", *IEEE Trans. Inf. Theory*, vol. 14, no. 4, pp. 563–568, Jul. 1968 (cit. on p. 56).

[SALP16] N. Sarmah, K. Aufinger, R. Lachner, and U. R. Pfeiffer, "A 200–225 GHz SiGe power amplifier with peak Psat of 9.6 dBm using wideband power combination", in *Proc. Europ. Conf. on Solid-State Circuits (ESSCIRC)*, Lausanne, Switzerland, Sep. 2016, pp. 193–196 (cit. on p. 7).

[SDF19a] M. Schlüter, M. Dörpinghaus, and G. Fettweis, "Bounds on phase and frequency estimation from 1-bit quantized signals with phase dithering", in *Proc. IEEE Int. Conf. on Communications (ICC)*, Shanghai, China, May 2019 (cit. on pp. 17, 98).

[SDF19b] ——, "Least squares phase estimation of 1-bit quantized signals with phase dithering", in *Proc. IEEE Int. Workshop on Signal Processing Advances in Wireless Commun. (SPAWC)*, Cannes, France, Jul. 2019 (cit. on pp. 17, 98).

[Sha94] S. Shamai, "Information rates by oversampling the sign of a bandlimited process", *IEEE Trans. Inf. Theory*, vol. 40, no. 4, pp. 1230–1236, 1994 (cit. on pp. 2, 15, 17, 18).

[Sha49] C. E. Shannon, "Communication in the presence of noise", *Proc. of the IRE*, vol. 37, no. 1, pp. 10–21, Jan. 1949 (cit. on p. 15).

[Sha48] C. E. Shannon, "A mathematical theory of communication", *Bell System Technical Journal*, vol. 27, pp. 623–656, 1948 (cit. on p. 25).

[SLd19] Z. Shao, L. T. N. Landau, and R. C. de Lamare, “Channel estimation using 1-bit quantization and oversampling for large-scale multiple-antenna systems”, in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Brighton, UK, May 2019, pp. 4669–4673 (cit. on p. 67).

[SLd18] ——, “Sliding window based linear signal detection using 1-bit quantization and oversampling for large-scale multiple-antenna systems”, in *Proc. IEEE Statistical Signal Processing Workshop (SSP)*, Freiburg, Germany, Jun. 2018, pp. 183–187 (cit. on p. 23).

[SLF15] G. Singh, L. Landau, and G. Fettweis, “Finite length reconstructible ASK-sequences received with 1-bit quantization and oversampling”, in *Proc. Int. ITG Conf. on Systems, Communications and Coding (SCC)*, Hamburg, Germany, Feb. 2015, pp. 1–6 (cit. on p. 21).

[SDM09] J. Singh, O. Dabeer, and U. Madhow, “On the limits of communication with low-precision analog-to-digital conversion at the receiver”, *IEEE Trans. Inf. Theory*, vol. 57, no. 12, pp. 3629–3639, Dec. 2009 (cit. on p. 16).

[SPM09] J. Singh, S. Ponnuru, and U. Madhow, “Multi-gigabit communication: The ADC bottleneck”, in *Proc. IEEE Int. Conf. on Ultra-Wide-Band (ICUWB)*, Vancouver, Canada, Sep. 2009, pp. 22–27 (cit. on p. 2).

[Sta12] R. B. Staszewski, “Digitally intensive wireless transceivers”, *IEEE Design & Test of Computers*, vol. 29, no. 6, 2012 (cit. on pp. 1, 2, 8).

[SMS09] T. Sundstrom, B. Murmann, and C. Svensson, “Power dissipation bounds for high-speed Nyquist analog-to-digital converters”, *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 56, no. 3, pp. 509–518, Mar. 2009 (cit. on pp. 2, 8).

[TNB19] M. Tanio, K. Ning, and J. F. Buckwalter, “A 60-GHz symmetric Doherty power amplifier with 20.4% 6-dB back-off efficiency”, in *Proc. Europ. Microwave Conf. (EuMC)*, Paris, France, Oct. 2019, pp. 559–562 (cit. on p. 7).

[ÜY18] A. B. Üçüncü and A. Ö. Yilmaz, “Oversampling in one-bit quantized massive MIMO systems and performance analysis”, *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 7952–7964, Dec. 2018 (cit. on pp. 22, 23).

[UKCE18] S. Ur Rehman, M. M. Khafaji, C. Carta, and F. Ellinger, “A 25-Gb/s 270-mW time-to-digital converter-based $8\times$oversampling input-delayed data-receiver in 45-nm SOI CMOS”, *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 65, no. 11, pp. 3720–3733, Nov. 2018 (cit. on p. 9).

[Vit67] A. Viterbi, “Error bounds for convolutional codes and an asymptotically optimum decoding algorithm”, *IEEE Trans. Inf. Theory*, vol. 13, no. 2, pp. 260–269, Apr. 1967 (cit. on p. 10).

[Wal99] R. Walden, “Analog-to-digital converter survey and analysis”, *IEEE J. Sel. Areas Commun.*, 1999 (cit. on p. 8).

[ZW88] E. Zehavi and J. Wolf, “On runlength codes”, *IEEE Trans. Inf. Theory*, vol. 34, no. 1, 1988 (cit. on p. 13).

[ZSK12] G. Zeitler, A. C. Singer, and G. Kramer, “Low-precision A/D conversion for maximum information rate in channels with memory”, *IEEE Trans. Commun.*, vol. 60, no. 9, pp. 2511–2521, Sep. 2012 (cit. on pp. 16, 67).

[ZDL+18] J. Zhang, L. Dai, X. Li, Y. Liu, and L. Hanzo, “On low-resolution ADCs in practical 5G millimeter-wave massive MIMO systems”, *IEEE Commun. Mag.*, vol. 56, no. 7, pp. 205–211, Jul. 2018 (cit. on p. 22).

[Zha12] W. Zhang, “A general framework for transmission with transceiver distortion and some applications”, *IEEE Trans. Commun.*, vol. 60, no. 2, pp. 384–399, 2012 (cit. on pp. 2, 18).

[ZGQ+19] J. Zhou, M. Guo, Y. Qiao, *et al.*, “Digital signal processing for faster-than-Nyquist non-orthogonal systems: An overview”, in *Proc. IEEE Int. Conf. on Telecommunications (ICT)*, Hanoi, Vietnam, Apr. 2019, pp. 295–299 (cit. on p. 11).

[Zig69] K. S. Zigangirov, “Sequential decoding for a binary channel with drop-outs and insertions”, *Probl. Peredachi Inf.*, vol. 5, no. 2, pp. 23–30, 1969 (cit. on p. 30).

[ZSB99] V. Zyablov, S. Shavgulidze, and M. Bossert, “An introduction to generalized concatenated codes”, *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 609–622, 1999 (cit. on p. 86).

Own Publications

[BDF17a] **S. Bender**, M. Dörpinghaus, and G. Fettweis, "On the achievable rate of bandlimited continuous-time 1-bit quantized AWGN channels", in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Aachen, Germany, Jun. 2017 (cit. on pp. 25, 59).

[BDFa] ——, "On the achievable rate of bandlimited continuous-time AWGN channels with 1-bit output quantization", *EURASIP Journal on Wireless Communications and Networking*, submitted Aug. 2019 (cit. on pp. 25, 45, 48, 59).

[BDF17b] ——, "On the achievable rate of multi-antenna receivers with oversampled 1-bit quantization", in *Proc. IEEE Int. Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, Curacao, Dutch Antilles, Dec. 2017 (cit. on pp. 65, 66, 68, 69).

[BDFb] ——, "On the spectral efficiency of bandlimited 1-bit quantized AWGN channels with runlength-coding", *IEEE Commun. Lett.*, submitted Apr. 2020 (cit. on p. 51).

[BDF19] ——, "The potential of continuous phase modulation for oversampled 1-bit quantized channels", in *Proc. IEEE Int. Workshop on Signal Processing Advances in Wireless Commun. (SPAWC)*, Cannes, France, Jul. 2019 (cit. on pp. 71, 84, 86).

[BLDF16] **S. Bender**, L. Landau, M. Dörpinghaus, and G. Fettweis, "Communication with 1-bit quantization and oversampling at the receiver: Spectral constrained waveform optimization", in *Proc. IEEE Int. Workshop on Signal Processing Advances in Wireless Commun. (SPAWC)*, Edinburgh, U.K., Jul. 2016 (cit. on pp. 13, 21).

[BSK+17] **S. Bender**, P. Seiler, B. Klein, *et al.*, "Pathways towards Tb/s wireless", in *Proc. IEEE Int. Conf. on Ubiquitous Wireless Broadband (ICUWB)*, Salamanca, Spain, Sep. 2017 (cit. on pp. 51, 59).

[FDB+19] G. Fettweis, M. Dörpinghaus, **S. Bender**, *et al.*, "Zero crossing modulation for communication with temporally oversampled 1-bit quantization", in *Proc. Annual Asilomar Conf. on Signals, Systems, and Computers (ASILOMAR)*, Pacific Grove, California, Nov. 2019 (cit. on p. 71).

List of Symbols

Constants and sets of numbers

$\mathbf{0}_a$	all-zero row vector of length a
e	Euler's number
γ	Euler-Mascheroni constant
$\boldsymbol{I}^{(K)}$	identity matrix of size $K \times K$
j	imaginary unit
k_B	Boltzmann constant
$\mathbb{N}$	space of natural numbers
$\mathbb{N}_0$	space of natural numbers including zero
π	ratio of the circumference of a circle to the diameter
$\mathbb{R}$	space of real numbers
$\mathbb{Z}$	space of integer numbers

Functions and Operators

$*$	convolution
$\lfloor \cdot \rfloor$	floor function
$\mathbb{1}_{[u,v]}(x)$	indicator function, one in the interval $[u, v]$ and zero otherwise
$\lceil \cdot \rceil$	ceil function
$\mathrm{Ci}(\cdot)$	cosine integral function
$\mathrm{count}_x(\mathcal{X})$	number of occurences of x in the set $\mathcal{X}$
$\delta(\cdot)$	Dirac delta function
$d(\boldsymbol{x}, \boldsymbol{y})$	distance metric between two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$
$d_H(\boldsymbol{x}, \boldsymbol{y})$	hamming distance of two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$
$\mathbb{E}$	expectation
$H_b(\cdot)$	binary entropy function
$\Im(\cdot)$	imaginary part of a complex number
$\min_a b(a)$	minimum of b over all a
mod	modulo operator
$Q_1(x)$	binary quantization function with threshold zero
$\Re(\cdot)$	real part of a complex number
$\mathrm{rect}(t)$	rectangular function, $\mathrm{rect}(t) = \mathbb{1}_{[-\frac{1}{2}, \frac{1}{2}]}(t)$
$\mathrm{Si}(\cdot)$	sine integral function
$\mathrm{tri}(t)$	triangular function, $\mathrm{tri}(t) = (1 - \lvert t \rvert)\mathbb{1}_{[-1,1]}(t)$
$\mathrm{Var}(\cdot)$	variance

Units of measure

GHz	Gigahertz, 10^9 Hz
Hz	Hertz, unit of frequency
bit	unit of information w.r.t base 2
dB	Decibel
dBm	Decibel w.r.t. 10^{-3} W
J	Joule, unit of energy
s	Second, unit of time
W	Watt, unit of power

Variables, Signals, Sets and Parameters

$\mathbf{A}$	random process at the channel input
A_k	runlength at input, i.e., distance between kth and $(k-1)$th zero-crossing of $\mathsf{x}(t)$
α	ratio of clippped energy contributing to ISI
$a_{\min}$	minimum runlength of runlength-limited sequences
α_{RC}	rolloff of the raised cosine filter
a_k	runlength of the kth run of a runlength-limited sequence
α	CPM-symbol
β	transition duration
b	quantizer resolution
B_x	bandwidth w.r.t. to a bandwidth measure x, e.g., 90 % power containment bandwidth
C	channel capacity
$\mathbf{C}$	precoding operation for 1-bit MIMO channels
c_1	parameter of $S_{\mathsf{X}}(\omega)$, $c_1 = c(2\pi W)$, cf. A.10
Δ_k	magnitude error of kth transmit symbol A_k
D_m	distance between the mth and the $(m-1)$th zero-crossing of $\mathsf{r}(t)$
$\mathbf{D}$	random process at the channel output
$\hat{\mathsf{D}}_k$	distance between zero-crossings after applying the auxiliary information
$d_{\min}$	minimum distance
d_{rll}	d-constraint of runlength-limited sequences
E_b	bit energy
E_s	symbol energy
Δf	intermediate frequency to obtain a time invariant trellis for CPM modulation
f	frequency
f_0	carrier frequency
f_{IF}	intermediate frequency
f_s	sampling rate, $f_s = \frac{1}{T_s}$

f_{sig}	signaling rate at the transmitter
$g(t)$	transmit pulse
$\hat{g}(t)$	filtered transmit pulse
$\hat{g}_{\mathrm{appr}}(t)$	approximation of the filtered transmit pulse
$\tilde{g}(t)$	deviation between original and filtered transmit pulse, $\tilde{g}(t) = \hat{g}(t) - g(t)$
$G(\omega)$	spectrum of the transmit pulse $g(t)$
Γ_i	parameter related to the transmit pulse spectrum, $\Gamma_i = \int_{2\pi W}^{\infty} \omega^i \left\lvert G(\omega)\right\rvert^2 d\omega$
G_{PA}	gain of the power amplifier
$\mathbf{H}$	channel matrix
$H(\omega)$	channel transfer function
$\boldsymbol{H}$	Teoplitz matrix containing the sampled channel impulse response $\boldsymbol{h}$
h_0	maximum of the channel impulse response, ususally $h_0 = h(0)$
$h(t)$	channel impulse response
H	entropy
h_{cpm}	modulation index fo the CPM-signal
h	differential entropy
$\boldsymbol{h}$	sampled channel impulse response $h(t)$
H'	entropy rate
I'	mutual information rate
I	mutual information
K	number of input symbols
k	discrete time index
κ	normalized bandwidth $\kappa = {}^{W}/_{\lambda}$
k_{rll}	k-constraint of runlength-limited sequences
L_n	sum of n consecutive A_k
λ	rate parameter of the exponential distribution
λ_{rll}	parameter of the max-entropic RLL-sequences
$(\cdot)_{\mathrm{LB}}$	lower bound
L_{cpm}	memory of the CPM-symbol in terms of the symbol duration
L_h	length of the channel impulse response in multiples of T_{symb}
L_v	length of the receive filter in multiples of T_{symb}
M	number of output symbols
m	contribution of the transition waveform to the average signal power, $m = 2W \int_0^{\beta} \zeta^2 \left(t - \frac{\beta}{2}\right) dt$, cf. (3.15)
M_{cpm}	Order of the CPM modulation alphabet
M_{FTN}	faster-than-Nyquist factor
M_{osr}	oversampling ratio w.r.t. bandwidth or symbol duration
μ	expectation of V_k related to the number of inserted symbols
$\mathcal{N}(\mu, \sigma^2)$	normal distribution with mean μ and variance σ^2
N_r	number of receive antennas

N_t	number of transmit antennas
$\mathsf{n}(t)$	noise signal
N_0	noise power spectral density
n_{const}	number of constellation points of the CPM-signal
N_{d}	number of distinguishable CPM-symbols
$\hat{\mathsf{n}}(t)$	filtered noise signal
n_{IF}	normalized additional intermediate frequency
N_{samp}	number of samples
$P.$	power consumed by, emitted by or fed to a device
$p_{\mathsf{X}}(\cdot)$	probability density function of random variable X
P	average signal power
p_b	sample or bit flipping probability
$\phi(t)$	phase trajectory of the CPM signal corresponding to a baseband signal
$\psi(t)$	phase trajectory of the CPM signal including a frequency offset
ϕ_0	initial signal phase, $\phi_0 = \phi(0)$
$\hat{P}$	peak power constraint
$\mathbf{Q}_{\mathsf{XY}}$	cross-covariance matrix of X and Y (or covariance matrix $\mathbf{Q}_{\mathsf{X}}$ if $\mathsf{X} = \mathsf{Y}$)
$q(t)$	phase smoothing response
$\mathbf{R}_{\mathsf{XY}}$	cross-covariance matrix of X and Y (or covariance matrix $\mathbf{R}_{\mathsf{X}}$ if $\mathsf{X} = \mathsf{Y}$)
R	data rate
$\mathsf{r}(t)$	analog output signal
ρ	signal-to-noise ratio
$S_{\mathsf{x}}(\omega)$	power spectral density of $\mathsf{x}(t)$
S_k	shift of kth zero-crossing due to distortion
$s_{\mathsf{xx}}(\tau)$	autocorrelation function of $\mathsf{x}(t)$
$\hat{\mathsf{s}}(t)$	filtered bandpass signal
$\mathsf{s}(t)$	bandpass signal
σ_{x}^2	variance of the random variable or process x
T	absolute temperature
T'_k	time instant of the kth zero-crossing
T_k	time instant of the kth transition
τ	delay or time difference
T_{avg}	average symbol duration
θ	phase of the sampled and 1-bit quantized received signal $\mathbf{y}$
T_s	sampling interval, $T_s = \frac{1}{f_s}$
T_{Nyq}	time duration of a Nyquist interval
T'_{symb}	reduced symbol duration due to FTN transmission
T_{symb}	symbol duration
u	lowest value of $\hat{g}(t) - 1$ for $t > \beta$, cf. (3.29)
$(\cdot)_{\mathrm{UB}}$	upper bound

$\boldsymbol{V}$	Teoplitz matrix containing the sampled IF receive filter $\boldsymbol{v}_{\mathrm{rx}}(t)$
V_k	auxiliary process for kth input symbol
$\boldsymbol{v}_{\mathrm{rx}}$	sampled receive filter
$v_{\mathrm{rx}}(t)$	receive filter
$v_{\mathrm{tx}}(t)$	transmit filter
W	one-sided channel bandwdith
ω	angular frequency
$\mathsf{X}(\omega)$	spectrum of the signal $\mathsf{x}(t)$
$\hat{\mathsf{x}}(t)$	filtered transmit signal
$\tilde{\mathsf{x}}(t)$	filter distortion signal
$\mathsf{x}(t)$	analog input signal
$\mathsf{y}(t)$	1-bit quantized received signal
$\mathbf{y}$	(over)sampled 1-bit quantized received sequence
$\mathsf{z}_{\mathrm{ISI}}$	intersymbol interference
$\mathsf{z}(t)$	overall distortion
$\hat{\zeta}_0$	slope of the filtered transition waveform $\hat{\zeta}(t)$, $\hat{\zeta}_0 = \frac{d\hat{\zeta}(t)}{dt}$
$\hat{\zeta}(t)$	filtered transition waveform
$\zeta(t)$	transition waveform

List of Abbreviations

ACF	auto-correlation function
ADC	analog-to-digital converter
AQNM	additive quantization noise model
ASK	amplitude shift keying
AGC	automatic gain control
AWGN	additive white Gaussian noise
BER	bit error rate
BCJR	Bahl-Cocke-Jelinek-Raviv-algorithm
bpcu	bits per channel use
bpni	bits per Nyquist interval
CI	channel inversion
CMOS	complementary metal–oxide–semiconductor
CPFSK	continuous phase frequency shift keying
CPM	continuous phase modulation
CSI	channel state information
CSIT	channel state information at the transmitter
DMC	discrete memoryless channel
DTC	digital-to-time converter
EHF	extremely high frequency
ENOB	effective number of bits
EP	equal power allocation
FIR	finite impulse response
FoM	figure of merit
FSM	finite state machine
FTN	faster-than-Nyquist
GCC	generalized concatenated code
GMI	generalized mutual information
HP	hold period
IF	intermediate frequency
i.i.d.	independent and identically distributed
IoT	Internet of Things
ISI	intersymbol interference
ISSCC	International Solid-State Circuits Conference
i.u.d.	independent and uniformly distributed

LB lower bound
LDPC low density parity check
LMMSE linear minimum mean-squared error
LP lowpass
LOS line-of-sight
MIMO multiple-input multiple-output
ML maximum likelihood
mmWave millimeter wave
MSK minimum shift keying
NN neural network
NRZI non-return-to-zero inverse
OFDM orthogonal frequency-division multiplexing
OSR oversampling ratio
PA power amplifier
PAE power added efficiency
PAPR peak-to-average power ratio
PCBW power containment bandwidth
pdf probability density function
poly polynomial transition
PSD power spectral density
PSK phase shift keying
QAM quadrature amplitude modulation
QoS quality of service
QPSK quadrature phase-shift keying
RC raised cosine
RF radio frequency
RLL runlength limited
RRC root raised cosine
SE spectral efficiency
SHF super high frequency
SNDR signal-to-noise-and-distortion ratio
SNR signal-to-noise ratio
SRC squared raised cosine
SVD singular value decomposition
TI transition interval
TDC time domain converter
UB upper bound
UHF ultra high frequency
VLSI Symposium on VLSI Circuits
WF waterfilling
ZC zero-crossing

List of Figures

List of Tables