Advanced Satellite Monitoring Using Blind Demodulation Techniques

Markus Flohberger

Institute of Communication Networks and Satellite Communications
Graz University of Technology, Austria

Advisors

Prof. Otto Koudelka
Prof. Alessandro Vanelli-Coralli

A thesis submitted for the degree of
Dr. techn.
December 2010

Statutory Declaration

I declare that I have authored this thesis independently, that I have not used other than the declared sources / resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

Graz, December 16, 2010 Markus Flohberger

I would like to dedicate this thesis to my loving parents.

Acknowledgements

I would like to thank my supervisor, Professor Otto Koudelka, for his guidance, support and the opportunity to work in a professional field. Furthermore, many thanks must be given to my second examiner, Professor Alessandro Vanelli-Coralli from University of Bologna (Italy), for his comments and time for discussion.

Moreover, I would like to express my gratitude to all my colleagues at the Institute of Communication Networks and Satellite Communications (Graz, University of Technology) and the Institute of Space Technology and Acoustics (Joanneum Research), for their teamwork, help and interesting discussions. In this context, special thanks must be given to Wilfried Gappmair for proofreading and support in mathematical concerns.

Part of this work was carried out within the Advanced Research on Telecommunication Satellite Systems (ARTES) program provided by the European Space Agency (ESA), the Austrian National Space Program (ASAP) and the Satellite Communications Network of Excellence (SatNEx-II) launched by the European Commission.

Abstract

With the growing customer demands the amount of satellite services and traffic increases as well. For this reason an efficient usage of the existing resources, i.e. geostationary arc and available frequency bands, is mandatory. Thus, it is obvious that appropriate monitoring is of paramount importance for flawless operation of today's and future satellite networks.

With the rapid increase of available computational power and the benefits of digital signal processing, it is now possible to perform tasks that were not even thinkable a decade ago. For successful execution of the required monitoring tasks, a blind demodulation framework is proposed which is suited for implementation on software-defined radio (SDR) platforms.

A comprehensive overview of the developed framework is given, including design of specific algorithms, performance assessment of the required stages and remarks on efficient implementation. The suggested blind demodulation structure was integrated and tested on an in-house developed SDR platform.

Moreover, it is shown on the basis of two examples, how blind demodulation techniques can be used in applications which are beyond the capabilities of state-of-the-art monitoring equipment. The first example demonstrates an effective way for online estimation of the cross-polarization discrimination (XPD) and the second application presents an enhancement of the geolocation accuracy using time difference of arrival (TDOA) and frequency difference of arrival (FDOA) techniques.

Contents

List of Figures

List of Tables

Acronyms

AAF anti-aliasing filter

ACM adaptive coding and modulation

ADC analog-to-digital converter

APSK amplitude phase shift keying

ASIC application-specific integrated circuit

AWGN additive white Gaussian noise

BER bit error rate

BPSK binary phase shift keying

CAF cross-ambiguity function

CIC cascaded integrator comb

CCIR international radio consultative commitee

CRLB Cramer-Rao lower bound

DA data-aided

DDC digital down-conversion

DD decision-directed

DFT discrete Fourier transform

DSR differential slant range

DSRR differential slant range rate

DVB-S2 digital video broadcasting - satellite 2

EMC electromagnetic compatibility

EVM error vector magnitude

FB feedback

FDOA frequency difference of arrival

FEC forward error correction

FF feedforward

FFT fast Fourier transform

FPGA field-programmable gate array

GA Gardner detector

GEO geostationary earth orbit

GPS global positioning system

HPA high power amplifier

IFFT inverse fast Fourier transform

IF intermediate frequency

IPP Intel performance primitives

IRGN iterative regularized Gauss-Newton

ISI inter-symbol interference

LOP line of position

LTI linear time-invariant

LUT look up table

LVDS low voltage differential signal

MB Moeneclaey and Batsele

MCRLB modified Cramer-Rao lower bound

MF matched filter

ML maximum likelihood

MSE mean square error

NDA non-data-aided

NMSE normalized mean square error

OM Oerder and Meyr

OMT orthogonal mode transducer

PCI peripheral component interconnect

PC personal computer

PSD power spectral density

PSK phase shift keying

QAM quadrature amplitude modulation

QOS quality of service

QPSK quadrature phase shift keying

RAM random-access memory

RCOS raised cosine

RCVO receive only

RRCOS root-raised cosine

SDR software-defined radio

SNR signal-to-noise ratio

SVD singular value decomposition

TDMA time-division multiple access

TDOA time difference of arrival

TSVD truncated singular value decomposition

TWT traveling wave tube

VSAT very small aperture terminal

xGA extended Gardner detector

XPD cross-polarization discrimination

CHAPTER 1

Introduction

1.1 Background

Since the launch of INTELSAT I as the first commercial geostationary earth orbit (GEO) satellite in 1965, satellite communication has experienced an enormous evolution [1, 2]. The increase of transmission power available onboard satellites and advances in physical layer techniques lead to an evolvement of ground terminals as well. In particular, a reduction of terminal sizes away from huge earth stations (30 m antenna diameter) to very small aperture terminals (VSATs) (< 3 m antenna diameter) and further on to hand-held devices was possible. A benefit resulting from this trend was a significant cut of user terminal costs. Together with the aforementioned cost reduction, the advantage of wide coverage and high bandwidth, satellite communication is well established for fixed services. In this context, digital television broadcast with an enormous number of customers is the most prominent example.

Nowadays, customer terminals are not exclusively restricted to receive only (RCVO) stations, but feature transmission capabilities as well. This allows two-way communication services such as interactive television, speech or broadband Internet services. It turned out that satellite communications can not compete the terrestrial business in urban and suburban areas. However, rural areas as well as the maritime and aeronautical market are dominated by satellite services. With the advances in technology, the development of new services and the increasing demands of customers, a significant traffic growth can be expected in the future. To be capable to serve future user needs, the advantages of satellite communications, i.e. wide area coverage, speed and rapid deployment, may be used to establish networks, consisting of satellite and terrestrial networks [3].

1.2 Motivation and Objectives

Since the satellite communication resources, i.e. geostationary arc and available frequency bands, are limited, an efficient usage of the latter is mandatory. For this reason, appropriate

monitoring tools are necessary to guarantee flawless operation of satellite networks. The availability and the quality of service (QOS) are key requirements to hold the position on a competitive market. Thus, satellite operators are urged to identify and fix problems quickly to satisfy the demands of their customers. Advances in technology supporting the daily required monitoring tasks in an efficient way would be a great favor for satellite operators.

With the rapid increase of available computational power, it is now possible to perform operations that were not possible a decade ago. In this work it will be shown that a versatile blind demodulation framework could be an attractive tool for carrying out state-of-the-art monitoring tasks and aiding measurement of enhanced parameters as well. To achieve a flexible and low-cost solution the developed framework is running on a software-defined radio (SDR) platform [4].

Two main objectives of the enhanced monitoring approach are: (i) it relies only on passive measurements, i.e. no active transmission of signals is necessary which may affect the active traffic; (ii) monitoring is performed in an online manner and results are provided close to real-time. In this context, the definition of close to real time is elastic and related to the specific task to solve and the according complexity.

1.3 Outline of the Thesis

In Chapter 2, the used signal model for generation of multi-carrier signals and the supported modulation schemes are introduced briefly. It is common practice to assume that, apart from attenuation aspects, additive white Gaussian noise (AWGN) is the dominating impairment on GEO satellite channels [5]. For this reason, propagation effects or system-induced effects, e.g. phase noise, I/Q imbalance or nonlinearities, are not considered.

Usually, monitoring is performed in the frequency domain by investigation of the satellite traffic using a spectrum analyzer. This measurement procedure delivers standard parameters such as the presence of traffic, center frequency, symbol rate and signal-to-noise ratio (SNR). The proposed framework provides these parameters at increased accuracy in an automated way. Additionally, information on the filter characteristics and the used modulation scheme becomes available. In case carrier specifications are violated, alarms could be triggered to inform the ground personnel about the situation. A comprehensive overview of the developed blind demodulation framework is given in Chapter 3. At first, appropriate detection of carriers in a noisy observation is discussed, followed by the successive blind demodulation steps of the identified signals. Finally, the performance of the blind demodulation framework is assessed in terms of the error vector magnitude (EVM) for the remodulation stage.

In Chapter 4, two applications are presented in which the benefits of the proposed blind demodulation framework are emphasized. Interference, caused by insufficient cross-polarization discrimination (XPD), malfunctioning system components, unauthorized

access of transponders or saturation of the high power amplifier (HPA), may affect the active traffic severely. Hence, monitoring is necessary to prevent the issue at all or identify the cause to remove it as quickly as possible.

For surveillance of the XPD it is state-of-the-art to stop services on the transponder of interest and perform active measurements using a continuous wave signal. Since these measurements are carried out on a weekly basis, both the interruption of services and the required on-site ground personnel are problematic for satellite operators due to the additional costs. It will be shown that an SDR platform acquiring co- and cross-polar signals simultaneously can perform estimation of the XPD online and in a passive manner. The application of blind demodulation techniques may additionally enhance the accuracy of the result with moderate complexity.

If the interference stems from a malfunctioning terminal, the transmitting station can be identified by using the satellite operator database. In case no information is available or the interference is caused by satellite piracy, other measures have to be taken to determine the source of the perturbing signal. Using time difference of arrival (TDOA) and frequency difference of arrival (FDOA) techniques, a solution of this geolocation task is obtained. Since adequate geolocation accuracy requires the processing of a huge amount of data, blind demodulation techniques may be used to enhance the transmitter location significantly.

1.4 Contributions

To the best of the author's knowledge no such detailed investigation of a comparable blind demodulation framework is available from the open literature. Most publications dedicated to blind demodulation deal solely on single stages, e.g. modulation classification or symbol timing recovery, and do not span the whole spectrum from carrier detection to symbol decision. In the presented work it is tried to assess the performance of the overall framework by determining the EVM of the remodulation stage as well. Furthermore the proposed algorithms were integrated and tested on a versatile in-house developed SDR platform. So it was possible to give several remarks on implementation details and verify that the applications in Chapter 4, which are beyond the state-of-the-art, can be performed on a low-cost platform by using sophisticated digital signal processing techniques. In the following the major contributions and related published work are summarized.

- Design, implementation and testing of an innovative blind demodulation framework suited for SDR platforms. The latter can alleviate the daily monitoring tasks of satellite operators considerably. A future publication presenting the overall blind demodulation performance is under progress [6].

- Hardly any estimators for assessing the symbol rate of baseband filtered signals are available from the open literature. Moreover no performance results on non-integer oversampling ratios were published. Motivated by this, in Section 3.3, a symbol rate estimation algorithm which significantly enhances the method by

Kueckenwaitz [7] is introduced. The performance of the proposed scheme is verified by numerous simulation runs, including non-integer oversampling ratios. If applicable, the refinement step published by the author in [8] can be employed as well.

- For solving parts of the blind demodulation problem, the usage of non-data-aided (NDA) synchronization algorithms suggests itself. For this reason an in-depth roundup on NDA synchronization of the supported modulation schemes was carried out, including adaptation of existing and design of new algorithms. Particular attention was paid to the analysis of efficient algorithms for symbol timing recovery. In Section 3.8, a very detailed discussion on powerful NDA timing recovery schemes is presented as it is not available in the open literature. Additionally, the acquisition behavior of the Gardner detector (GA) was examined thoroughly in [9]. The detector characteristic of the Moeneclaey and Batsele (MB) detector and the feedforward counterpart are derived in [10] and [11]. Moreover it was shown in [12] that Nyquist pulses apart from raised cosines can have attractive properties for feedback timing recovery.

- No publications have been found on a modulation classifier, capable of estimating all the signal constellations suggested in the DVB-S2 standard [13]. Since such a classifier is of paramount importance in future applications, in Section 3.9, an estimation scheme is proposed for the aforementioned constellations. The introduced method additionally supports quadrature amplitude modulation (QAM) signals as representatives of bandwidth efficient modulation schemes. The introduced modulation classifier was published recently in [14].

- The need for powerful SNR estimation schemes arises in satellite networks, e.g. adaptive coding and modulation (ACM) or satellite monitoring. To achieve acceptable SNR estimation performance for signal constellations with constant as well as non-constant envelope, the combination of three moment-based methods operating on baud-rate is suggested in Section 3.10. Furthermore, a moment-based scheme for oversampled signals has been explored by the author in [15] and [16].

- In Section 3.11, a roll-off estimation scheme relying on SNR estimation at the matched filter (MF) output is introduced and compared to the only available scheme from the open literature.

- Finally, in Chapter 4, the advantages of the presented framework are demonstrated by two important applications. Firstly, an online XPD measurement is proposed instead of the cumbersome state-of-the-art method. This new innovative approach would not be possible without the blind demodulation framework. Secondly, it is demonstrated how blind demodulation techniques can be used to enhance the geolocation accuracy of TDOA/TDOA or TDOA/FDOA schemes.

The complete publication list as first and co-author can be found in the Appendix.

CHAPTER 2

Equivalent Baseband Model

For GEO communication satellites, attenuation and AWGN can be considered as the main impairments of the transmitted signal [5]. If the nonlinear effect of the HPA, commonly of traveling wave tube (TWT) type, is neglected, the signal transmission can be formulated as linear time-invariant (LTI) system. In the following an overview of the used signal model and modulation schemes is provided.

2.1 Signals and Modulation Schemes

The received multi-carrier signal $r(t)$ can be stated as follows:

$$r(t) = \sum_{q} u_q(t) + \sqrt{E_n} w(t) \tag{2.1}$$

where each single carrier $u_q(t)$ can be formulated as

$$u_q(t) = \sqrt{E_{s,q}}\, e^{\jmath(2\pi f_{c,q} t + \theta_q)} \sum_{i} c_{i,q}\, h_q(t - iT_q - \tau_q) \tag{2.2}$$

Herein q is the index of the carrier, such that $f_{c,q}$ denotes the center frequency of the q-th carrier. Let $w(t)$ be the complex zero-mean AWGN with independent real and imaginary part. The noise power corresponds to E_n and the signal power for each carrier to $E_{s,q}$ such that the SNR of the specific carrier is defined by $\gamma_{s,q} = E_{s,q}/E_n$. Furthermore, $c_{i,q}$ are the transmitted data symbols, which are assumed to be independent and identically distributed (i.i.d.) and normalized to unit energy such that $E[|c_{i,q}|^2] = 1$; T_q denotes the symbol period, which is the reciprocal value of the symbol rate, i.e. $T_q = 1/f_{d,q}$; θ_q expresses the phase offset with $\theta_q \in [-\pi,\pi)$ and τ_q, $|\tau_q| \leq T_q/2$, the symbol timing offset of the specific carrier; $h_q(t)$ describes the baseband pulse shape of the q-th carrier and is a function of the roll-off factor α_q, with $0 \leq \alpha_q \leq 1$. The impulse response exhibits unit

energy such that

$$\int_{-\infty}^{\infty} |h(t)|^2 \, dt = 1 \tag{2.3}$$

It is assumed that the pulse shape coincides with the de facto standard in satellite communications, namely the root-raised cosine (RRCOS) [13]. If this assumption does not hold, an additional module for estimating the baseband pulse shape has to be added.

The received signal is assumed to be sampled appropriately for subsequent processing. The carrier of interest is converted down to baseband and applied to the MF. Thus, the sampled signal at the filter output can be formulated as

$$x_{k,q} \triangleq r(t) \otimes h_q^*(-t)\big|_{t=kT_s} \approx e^{\jmath(2\pi k \Delta f_q T_q / N_{s,q} + \theta_{k,q})} s_{k,q} + \sqrt{E_n} n_{k,q} \tag{2.4}$$

with the signal component

$$s_{k,q} = \sqrt{E_{s,q}} \sum_i c_{i,q} \, g_q((k/N_{s,q} - i - \epsilon_{k,q}) T_q) \tag{2.5}$$

In this context, Δf_q expresses the residual frequency offset which is present after down conversion; T_s denotes the sampling period, such that the sampling rate can be stated as $f_s = 1/T_s$. Thus the number of samples per symbol, or the so called oversampling factor, for the q-th carrier is given by $N_{s,q} = T_q/T_s = f_s/f_{d,q}$; $n_{k,q}$ is the (non-white) Gaussian noise sample, which is shaped by the receiving filter $h_q^*(-t)$. The overall baseband pulse shape $g_q(t) = h_q(t) \otimes h_q^*(-t)$ results in a raised cosine (RCOS) to guarantee absence of inter-symbol interference (ISI), where $\otimes$ denotes convolution. It is common practice to normalize the timing offset τ_q by the symbol period T_q, expressed by ϵ_q.

In the following the generation of the data symbols c_n is briefly described for modulation schemes considered as most important in satellite communications. For the time being, seven signal constellations are supported by the demodulation framework.

2.1.1 M-PSK

For the most commonly used modulation scheme in satellite communications, i.e. M-ary phase shift keying (PSK), 2-PSK, 4-PSK and 8-PSK are supported. The symbols c_n for M-PSK constellations can be generated according to

$$c_n = \begin{cases} e^{\jmath\pi n} = \pm 1, & M = 2 \\ e^{\jmath\pi(2n+1)/M}, & M \neq 2 \end{cases} \tag{2.6}$$

Herein, M is the order of modulation given by $M = 2^\nu$ with $\nu \in \mathbb{N}$ and $n = 0,1,\ldots,M-1$. Figure 2.1 shows the supported constellations: 2-PSK (red crosses), 4-PSK (blue plus signs) and 8-PSK (green circles).

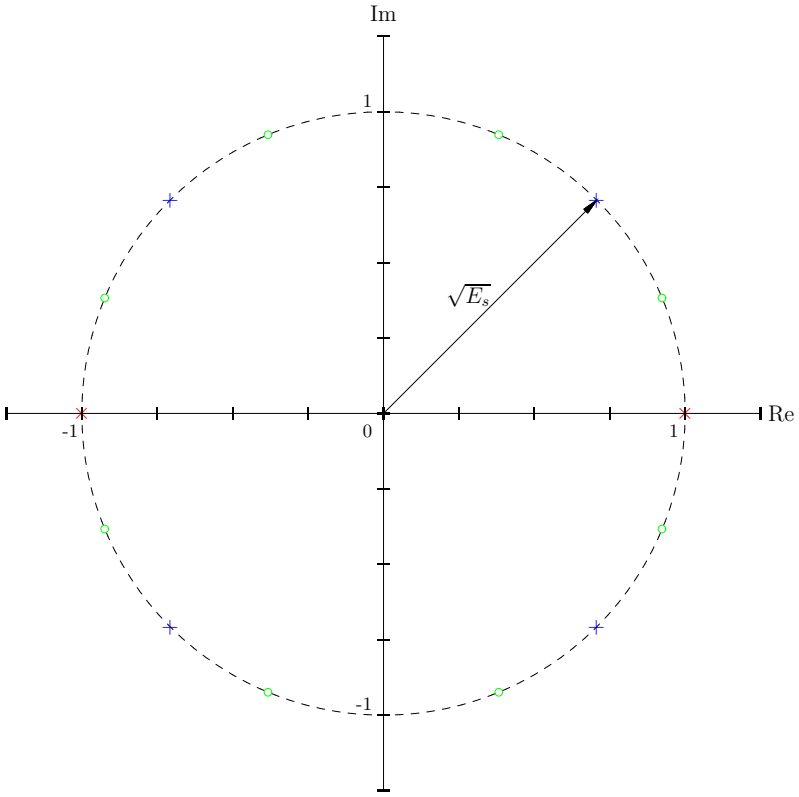


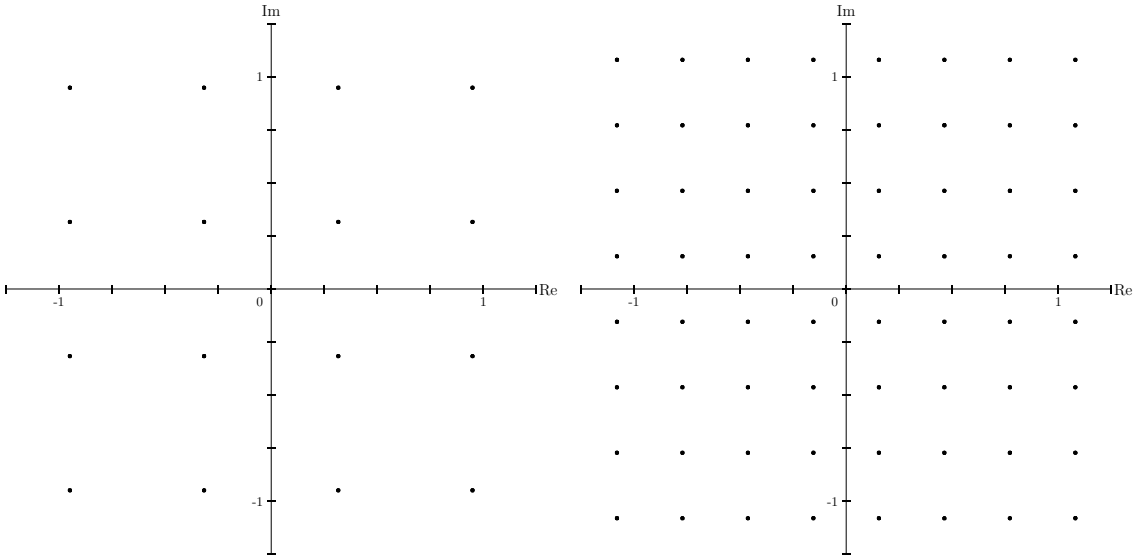
Figure 2.1: 2/4/8-PSK constellations

2.1.2 M-QAM

Since the requirement for more bandwidth efficient systems arises, higher order modulation schemes have to be investigated as well. Rectangular M-ary QAM is not frequently applied in satellite communications due to the sensitivity to nonlinear distortions. Nevertheless, 16-QAM and 64-QAM will be considered in the context of this thesis. For M-QAM the order of modulation is calculated by $M = 2^{2\nu}$ with $\nu \in \mathbb{N}$ and the symbols are defined as

$$c_n = \sqrt{\frac{3}{2(M-1)}}[(2n_1 + 1 - \sqrt{M}) + \jmath(2n_2 + 1 - \sqrt{M})] \tag{2.7}$$

with $n_1, n_2 = 0{,}1, \ldots \sqrt{M} - 1$. The constellation is normalized to arrive at unit energy. In Figure 2.2 both supported QAM constellations are shown.



(a) 16-QAM **(b)** 64-QAM

Figure 2.2: M-QAM constellations

2.1.3 M-APSK

Finally, 16-APSK and 32-APSK are implemented as members of amplitude phase shift keying (APSK). They have the advantage of less sensitivity to nonlinear distortions compared to QAM [17], due to the restricted number of signal amplitude levels. Moreover, APSK signals exhibit a smaller peak-to-average power ratio than square QAM so that they can be operated at less backoff of the HPA onboard the satellite. For higher order constellations, APSK performs better than PSK in presence of noise. Despite the mentioned advantages of APSK and the recommendation of 16-APSK and 32-APSK in the DVB-S2 standard [13], surprisingly few publications are available dealing on synchronization of the latter.

16-APSK 16-APSK consists of a 4-PSK constellation located on the inner circle with radius r_1 and a 12-PSK scheme on the outer circle with radius r_2. Using the above assumption of unit energy yields $1/4\, r_1^2 + 3/4\, r_2^2 = 1$ such that the corresponding radii can be calculated as

$$r_1 = \sqrt{\frac{4}{1+3\beta_r^2}} \tag{2.8}$$

$$r_2 = \sqrt{\frac{4\beta_r^2}{1+3\beta_r^2}} \tag{2.9}$$

The ring ratio $\beta_r = r_2/r_1$ depends on the selected code rate r; for example, with $r = 4/5$, $\beta_r = 2.75$, as suggested in the DVB-S2 standard [13]. Thus the symbols can be constructed by

$$c_n = \begin{cases} r_1 e^{\jmath\pi(2n+1)/4}, & n = 0,1,\ldots,3 \\ r_2 e^{\jmath\pi(2n+1)/12}, & n = 4,5,\ldots,15 \end{cases} \tag{2.10}$$

The red crosses in Figure 2.3 show the 16-APSK constellation.

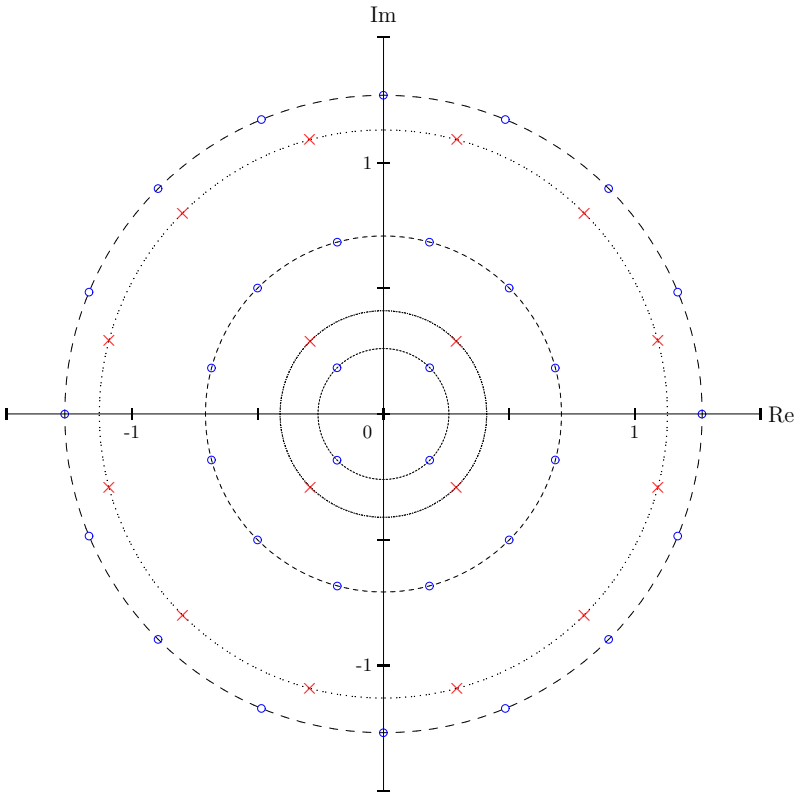


Figure 2.3: 16/32-APSK constellations

32-APSK This modulation scheme extends 16-APSK by an additional circle that holds a 16-PSK constellation. Since three amplitude levels exist, two ring ratios $\beta_{r1} = r_2/r_1 = 2.72$ and $\beta_{r2} = r_3/r_1 = 4.87$ have to be formed when assuming a code rate of $r = 4/5$. The

ring radii can be determined according to

$$r_1 = \sqrt{\frac{8}{1+3\beta_{r1}^2+4\beta_{r2}^2}} \tag{2.11}$$

$$r_2 = \sqrt{\frac{8\beta_{r1}^2}{1+3\beta_{r1}^2+4\beta_{r2}^2}} \tag{2.12}$$

$$r_3 = \sqrt{\frac{8\beta_{r2}^2}{1+3\beta_{r1}^2+4\beta_{r2}^2}} \tag{2.13}$$

Thus the symbols can be constructed by

$$c_n = \begin{cases} r_1 e^{\jmath\pi(2n+1)/4}, & n = 0,1,\ldots,3 \\ r_2 e^{\jmath\pi(2n+1)/12}, & n = 4,5,\ldots,15 \\ r_3 e^{\jmath\pi(2n)/16}, & n = 16,17,\ldots,31 \end{cases} \tag{2.14}$$

It should be pointed out that the outermost PSK constellation is not offset to the real axis like the inner and middle ones. Figure 2.3 depicts the 32-APSK constellation using blue circles.

CHAPTER 3

Blind Demodulation Framework

For reliable transmission of data in communication systems, it is inevitable to determine the most important parameters of the receiving signal. The evaluation of these parameters is termed parameter estimation and synchronization. The unknowns to be computed include carrier frequency offset, carrier phase error, optimum timing instant and, of course, the transmitted data symbols [18, 19]. For successful exchange of data the general framework of the transmission has to be known at both transmitter and receiver side. Parameters such as nominal carrier frequency, modulation scheme or symbol rate, are normally available to solve the task of synchronization at the receiver side. However, this condition does not hold for monitoring systems or non-cooperative demodulation where no information about modulation parameters and even basic specifications are available. The challenge in this field of work is the estimation of as many parameters as possible out of the received signal, with minimum deviation from the true values.

The signal model introduced in the previous section will be used as a basis throughout the rest of this work. The automatic blind demodulation of arbitrary carriers inside a received signal impaired by noise is a very challenging task, since the number of possible input signals is infinite. So it should be obvious that limitations have to be introduced to arrive at a tractable problem. These limitations could contain the type of modulation, the baseband pulse shape and allowed ranges of parameters. Throughout this work the linear digital modulation schemes, presented in Section 2.1, are considered. As already mentioned, RRCOSs are assumed for baseband pulse shaping which seems to be justified since the latter are established as de facto standard in satellite communications. Modularity of the developed framework is a key requirement, thus the addition of new features, e.g. modulation schemes or synchronization algorithms, can be performed in an easy manner.

For the blind demodulation process, available synchronization algorithms that do not rely on any prior knowledge, e.g. preambles or unique words, can be used. Such algorithms are termed as NDA in contrast to their data-aided (DA) counterparts. Another possibility is using so called decision-directed (DD) schemes, which incorporate the decisions of the transmitted data symbols into the estimation process. If the probability for decision errors

is very low, the performance of DD algorithms is comparable to that of DA methods. Thus, it should be clear that part of the blind demodulation process is closely related to NDA synchronization. For the framework, algorithms are selected which have satisfactory performance and require only knowledge of the received signal that is available at the related blind demodulation stage.

In addition to standard synchronization, accompanying algorithms have to be applied that provide the required parameters for synchronization at the successive demodulation stages [20]. These accompanying algorithms include schemes for estimation of the symbol rate, classification of the modulation constellation, determination of the baseband pulse shape and estimation of the SNR. Some of these methods are treated intensively throughout the open literature, but some are hardly mentioned at all. Throughout this work available algorithms will be investigated and compared to proposed schemes. If possible, modifications are applied to improve performance of the existing methods.

An overview of the blind demodulation framework is sketched in Figure 3.1. Before any demodulation algorithm can be started, it is crucial to detect appropriate carriers. Thus, the received noisy observation is investigated in the frequency domain to identify possible candidates for carriers. Pre-estimates, such as center frequency and occupied bandwidth, are evaluated for each carrier. During this process the noise level is estimated which is required for calculation of the SNR of each carrier. Later it will be seen that this estimate is of paramount importance for several blocks in the demodulation procedure.

After successful carrier detection, the procedure of demodulation can be summarized briefly as follows. Each of the carrier candidates is down-converted to baseband, filtered appropriately and decimated according to the requirements of the subsequent stages. At first the symbol rate has to be determined for further processing of the received signal. The carrier is resampled to a fixed rate to fulfill the requirements of following algorithms and to reduce computational complexity. Before applying the signal to the MF, the excess bandwidth of the filter has to be determined. Afterwards, the signal is rescaled to suit the subsequent stages. To recover the symbol timing, standard algorithms can be applied which are independent of the modulation scheme. The latter has to be determined in the following, since all successive stages require this knowledge. An additional SNR estimator may be used to verify the success of the preceding procedure and to refine the available pre-estimate. Apart from a possible roll-off refinement for the remaining stages, i.e. frequency and phase offset removal, algorithms like in common systems can be applied. Finally, based on the recovered signals, it is decided which data symbol has been transmitted. Alternatively, the framework should be capable to decide, whether the candidate signal was incorrectly identified as a carrier or the type of signal is not supported at all. The estimated data symbols might be remodulated and used for the applications detailed in Chapter 4.

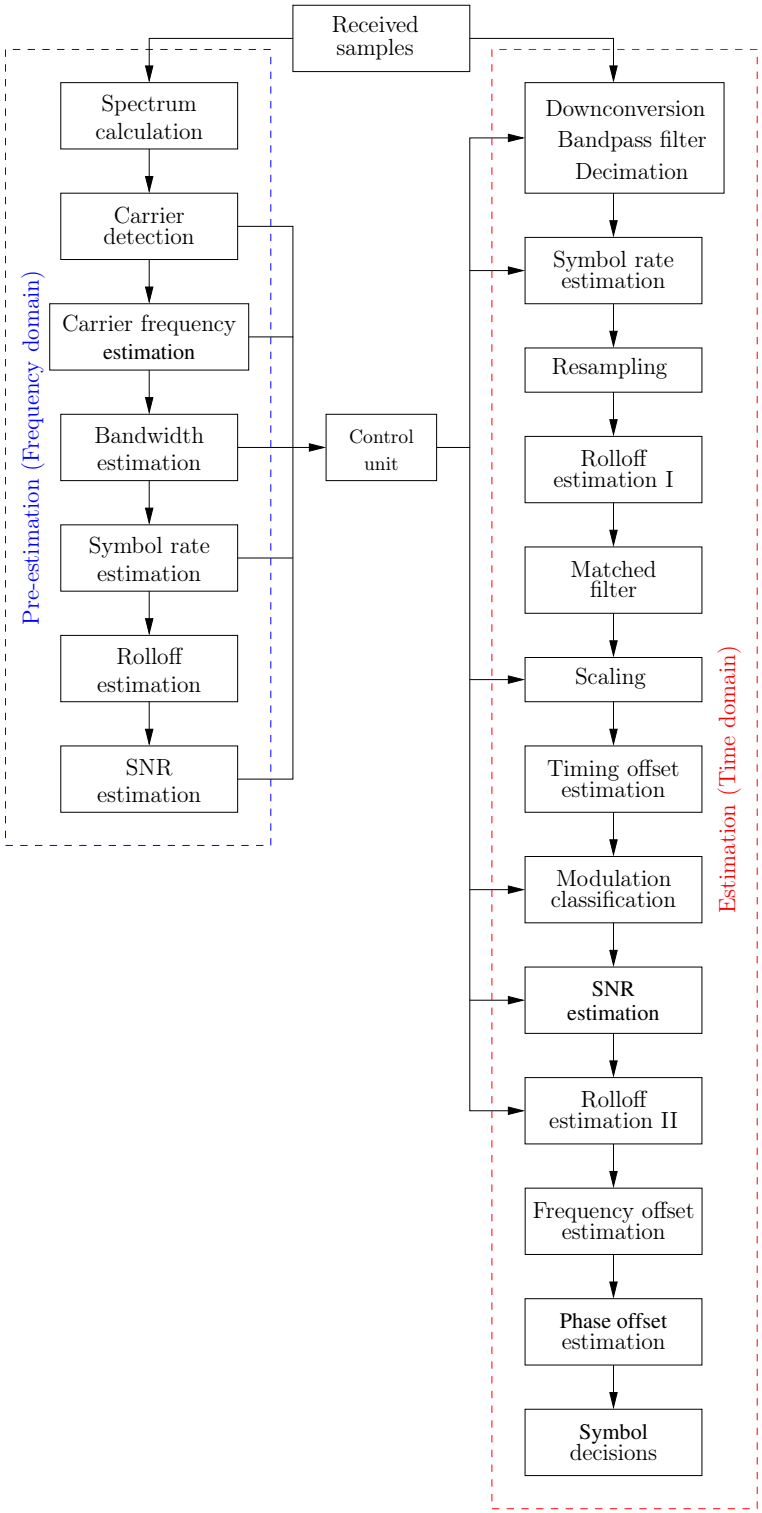


Figure 3.1: Overview of the blind demodulation framework

3.1 Carrier Detection

The detection of possible carriers is performed in the frequency domain by applying a method similar to the periodogram by Welch [21] for estimation of the power spectral density (PSD). The task of finding valid carriers is a trade-off between false detection and missing a carrier. A well-averaged periodogram of the received signal is the input to the detection stage. In Figure 3.2 the power spectrum R_n of the received signal r_k, a discretized version of Equation 2.1, is depicted for an example scenario. The latter uses the three supported PSK signal constellations. Increasing the number of averages reduces

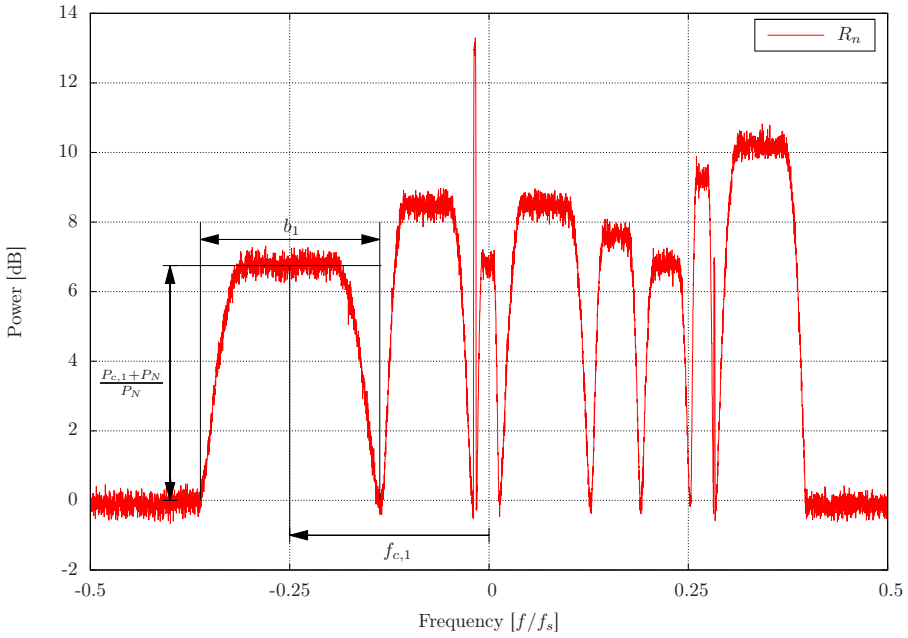


Figure 3.2: Power spectrum R_n of the example scenario using M-PSK

the jitter inherent to the spectral representation. For further reduction of fluctuations, smoothing is applied similar to video filtering in an analog spectrum analyzer.

As mentioned above, restrictions have to be applied on the range of specific parameters to arrive at a tractable problem. A first reasonable directive is the limitation of the range of symbol rates. Thus the degree of smoothing can be adjusted in such a way that carriers with the smallest symbol rate of interest are still not filtered out and fluctuations are reduced to avoid false detections. This choice is of course dependent on the fast Fourier transform (FFT) resolution and thus by the sampling rate and the length of the used FFT. A second restriction is related to the signal power of the carrier candidates. If the latter is below a pre-defined threshold, the carrier is ignored and considered as noise. The value of this threshold is set by the capabilities of the spectral detection algorithm and

the performance of the subsequent synchronization algorithms. If reliable detection is still requested, smaller threshold values increase the computational effort by augmentation of the number of averages used for power spectrum calculation. Moreover, with decreasing SNR the effort in the following blind demodulation procedure raises as well, since increasing of estimator lengths may be required. It can even happen that a complete exchange of synchronization algorithms is required which are better suited to smaller SNR values.

3.1.1 Detection Procedure

The spectral detection procedure starts with the estimation of the noise floor P_N of the received signal r_k, which can be estimated by evaluating the histogram of the power spectrum R_n. The first concentration of values in the low power range of the histogram is an easy and robust estimate $\hat{P}_N^*$ for the noise floor; the asterisk indicates that it is a pre-estimate. In Figure 3.3 the histogram of the previously introduced example scenario is illustrated. Afterwards, the complete spectrum is scanned and it is recorded whether

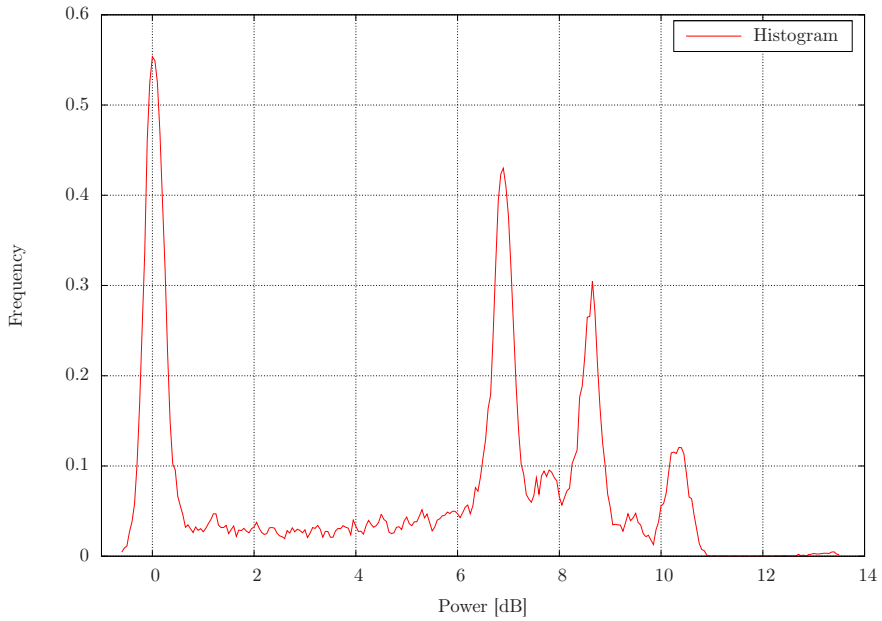


Figure 3.3: Histogram of the power spectrum R_n

the distance $R_n - \hat{P}_N$ is below or above the defined threshold. Additionally, a kind of hysteresis is applied such that fluctuations close to the threshold are mitigated. Finally, several criteria, e.g. symbol rate limitations, carrier spacing or SNR restrictions, are applied on the detected candidates to knock out possible invalid carriers. The number and position of the detected carrier candidates are stored and can be accessed by the subsequent pre-estimation stage.

3.2 Pre-Estimation

Using the carrier candidates of the previous stage, a rough pre-estimation of relevant parameters can be performed. The considered modulation types feature symmetric spectra around the center frequency, such that following spectral methods are valid.

A power-level $P_{c,q}$ can be computed for each carrier when averaging spectrum values R_n inside the passband of the specific carrier. The number of values that can be used for averaging is determined by the roll-off factor α_q. Communication systems are getting more and more bandwidth efficient, so the assumption of $\alpha_q \leq 0.5$ seems to be appropriate. The obtained power-level $\hat{P}^*_{c,q}$ represents signal-plus-noise power of the q-th carrier. Now, together with the previously computed noise level $\hat{P}^*_N$, an estimate of the SNR $\gamma_{s,q}$ in units of dB can be evaluated by

$$\hat{\gamma}^*_{s,q} = 10 \log_{10}(10^{(\hat{P}^*_{c,q} - \hat{P}^*_N)/10} - 1) \tag{3.1}$$

As already mentioned, the pre-estimate of the SNR turns out to be a very important value for the blind demodulation process. It is necessary for scaling the input signal, required for timing, frequency and phase recovery, for modulation classification and finally for symbol detection. When using sufficient averages for the periodogram calculation and appropriate smoothing techniques, the error of $\hat{\gamma}^*_{s,q}$ tends to be quite small, as will be verified by simulations in Section 3.15.

A guess for the occupied bandwidth b_q of each carrier can be made by searching the frequencies where the values R_n of the transition band enter the noise level P_N. Similarly, a rough estimate for the symbol rate $f_{d,q}$ can be computed by investigating the drop of the transition band values of the carrier spectrum as well. The frequencies have to be found, where the values R_n of the transition band excess a 3 dB difference with respect to the power level $P_{c,q}$ of the pass band. Due to the symmetry property, subtracting the higher frequency value from the lower one delivers a rough estimate for the symbol rate. And finally the calculation of the arithmetic mean of these two values roughly determines the center frequency $f_{c,q}$ of the specific carrier. Using occupied bandwidth and symbol rate, a pre-estimate for the excess bandwidth can be determined by

$$\hat{\alpha}^*_q = \hat{b}^*_q / \hat{f}^*_{d,q} - 1 \tag{3.2}$$

The estimated bandwidth $\hat{b}^*_q$ and the center frequency $\hat{f}^*_{c,q}$ are used as input for the digital down-conversion (DDC) and decimation stage to produce a single-carrier signal, denoted by $\tilde{r}_{k,q}$. Besides spectral shifting and filtering, cascaded integrator comb (CIC) filters [22] could be an attractive possibility for filtering out the carrier of interest and decimating it appropriately. Since, after DDC and decimation, only one carrier is further processed, the carrier index q and the tilde will be dropped for the sake of readability. Some of the above pre-estimates are indicated in the multi-carrier spectrum depicted in Figure 3.2.

3.3 Symbol Rate Estimation

Before any further synchronization algorithms can be applied to the single-carrier signal, the symbol rate f_d has to be determined with sufficient precision. Many existing symbol rate estimation algorithms do not work properly, if baseband pulse shaping is applied as it is common practice in satellite communications [23]. Moreover, in the open literature hardly any simulation results are available on algorithms working with non-integer oversampling rates. However, for monitoring systems it can not be assumed that the sampling rate is a multiple integer of the symbol rate. An algorithm which overcomes the above mentioned problems with acceptable computational burden consists of a modification of the scheme explained in [7] based on a rough estimate of the symbol rate from the pre-estimation stage. An optional refinement step using a symbol synchronizer whose output is fed back to a resampling block can be used for fine-tuning as detailed in [8]. Both methods are introduced in the following.

3.3.1 Estimation Stage

Standard Approach The algorithm in [7] consists of three stages, i.e. coarse, medium and fine. A nonlinearity $f(\cdot)$ is applied to the received signal r_k to generate a spectral line at the symbol rate f_d. Thus the periodogram of length N can be computed as

$$\tilde{S}_1[n] = \frac{1}{2\pi} \frac{|\sum_{k=0}^{N-1} f(r_k) w_k e^{-\jmath 2\pi k n/N}|^2}{\sum_{k=0}^{N-1} |w_k|^2} \tag{3.3}$$

Herein w denotes the windowing function and is assumed to be a Hanning window and the absolute value is used as nonlinearity $f(\cdot)$. For reduction of fluctuations, M sample blocks overlapping by 50 % or 75 % are used for calculation of $\tilde{S}_1[n]$. The resulting periodogram contains a DC component, a contour of the processed signal and a spectral line at the symbol rate. To remove the contour, a median filter (ordered statistic) is applied that filters out high frequency components. The filtered result is then subtracted from the original periodogram to arrive at a flat spectrum S_1. The coarse estimate can be extracted from the periodogram by a simple maximum search:

$$\hat{f}_{d1} = \frac{f_s}{N} \arg\max_n S_1[n] \tag{3.4}$$

The medium estimate $\hat{f}_{d2}$ is computed by performing a parabolic interpolation of the form $y = cx^2 + bx + a$ using the points adjacent to the maximum (see Appendix A.1).

The interpolated value is used as input for the final estimate $\hat{f}_d$. For this purpose, the available sample set is zero-padded by a factor k_{zp}, e.g. $k_{zp} = 64$, to increase the resolution of the spectrum. To save computation time and memory, only spectral values in the vicinity of the result from the second medium estimation step are evaluated. Again, the maximum of the calculated values is retrieved and interpolated to enhance the resolution.

The above described procedure corresponds almost to the algorithm in [7]. From the simulation runs it is evident that the estimate is either very accurate or completely wrong, when identifying a spectral line as maximum. The problem is that several parameters have an impact on the significance of the spectral line at the symbol rate. Thus, it may not be guaranteed that the spectral line at symbol rate is the global maximum. Degrading factors are decreasing SNR γ_s and roll-off α, small observation intervals and non-integer oversampling ratios N_s.

Enhanced Approach It can be summarized that the performance of the algorithm in a standard mode is not satisfactory for monitoring systems. However, since the pre-estimation stage extracts additional information, the latter might be used for enhancing this stage. It can be shown that, when applying the symbol rate value (used for DDC) from the pre-estimation stage, this effect can be mitigated and the performance of the algorithm improves significantly. Basically, the idea is to use the symbol rate hint for prefiltering and for limitation of the search range in the spectrum. For prefiltering a narrowband filter is applied to the received signal to remove unwanted noise. In case of multi-carrier scenarios such a filter has to be applied anyway to avoid interference by adjacent carriers. It was validated in numerous simulation runs that the error of the pre-estimated carrier bandwidth is well below 10 %. Thus a filter with a bandwidth slightly larger than the required bandwidth of the carrier can be designed. The additional limitation of the spectral search range will speed up calculation and decrease the probability of identifying a wrong peak as the assumed spectral line.

In the sequel, simulation results are presented to demonstrate the problems of the standard algorithm and show the benefits of the enhanced version. When not stated explicitly, the default parameters for all simulations are set as follows. As modulation scheme quadrature phase shift keying (QPSK) is used, the oversampling ratio is assumed as $N_s = 4.5$, the roll-off factor $\alpha = 0.2$ and the prefiltering bandwidth is set to $b_f = 1.2b$. The FFT length for periodogram estimation is set to $N = 2^n = 4096$, the number of sampling blocks to $M = 50$ which are overlapping by $3/4N$ and the zero-padding factor to $k_{zp} = 64$. For each simulation point 100 iterations are performed.

In the first test run the SNR is set to $\gamma_s = 0\,\mathrm{dB}$ and the roll-off factor to $\alpha = 0.35$. Figure 3.4 depicts the post-processed spectrum S_1 for both the standard and enhanced approach of a single simulation run. There should be a spectral line at the symbol rate, i.e. $f_d = 0.\dot{2}$. It can be seen that this is not true for the standard approach at such a low SNR. On the other hand, the spectral line is considerably pronounced when prefiltering is applied.

In Figure 3.5 the evolution of the normalized mean square error (NMSE) over the SNR is shown for different roll-off factors, i.e. $\alpha = 0.35$ and $\alpha = 0.2$. It can be seen that for $\alpha = 0.35$ the NMSE of the standard approach increases significantly below $\gamma_s = 8\,\mathrm{dB}$, but remains low over the whole SNR range when applying prefiltering and search range limitation. When reducing the roll-off factor to $\alpha = 0.2$, the smallest value recommended

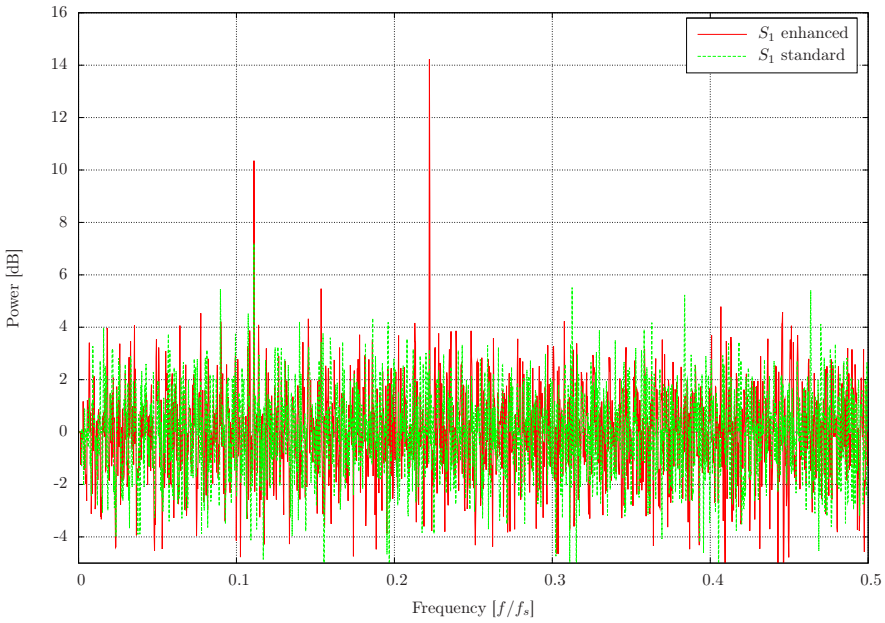


Figure 3.4: Peak detection in post-processed spectrum S_1

in [13], the degradation for the standard scheme occurs already below $\gamma_s = 12\,\mathrm{dB}$ and does not appear for the enhanced scheme at all. The rapid increase in NMSE below a specific SNR bound is due to wrong identification of a spectral line as the supposed symbol rate in the post-processed spectrum S_1.

The effect of changing the number of averages used for calculation of the periodogram is illustrated in Figure 3.6. It is obvious that increasing the number of processed samples increases the accuracy of the algorithm by reducing fluctuations inside the spectrum and thus shifting the degradation effect to lower SNR values. When using only $M = 10$ sample blocks, the performance of the enhanced scheme degrades below $\gamma_s = 2\,\mathrm{dB}$.

In the open literature there is hardly any information available on performance results of symbol rate estimation algorithms working with non-integer oversampling ratios. Since monitoring systems normally operate in a non-cooperative scenario, the assumption that symbol rates are multiple integers of the sampling frequency does not hold. The dependence of the introduced algorithms on the oversampling ratio N_s is shown in Figure 3.7. It can be seen that the algorithm favors integer ratios and the worst result is produced for $N_s = 4.5$. According to the chosen parameters the degradation effect is not visible for the enhanced algorithm at all.

Finally, the success rate is compared to the methods proposed in [7] (standard), [23] (filter bank) and [24] (cyclic correlation). An estimation is termed successful, if the estimation error is smaller than the spectral resolution. The amount of available symbols

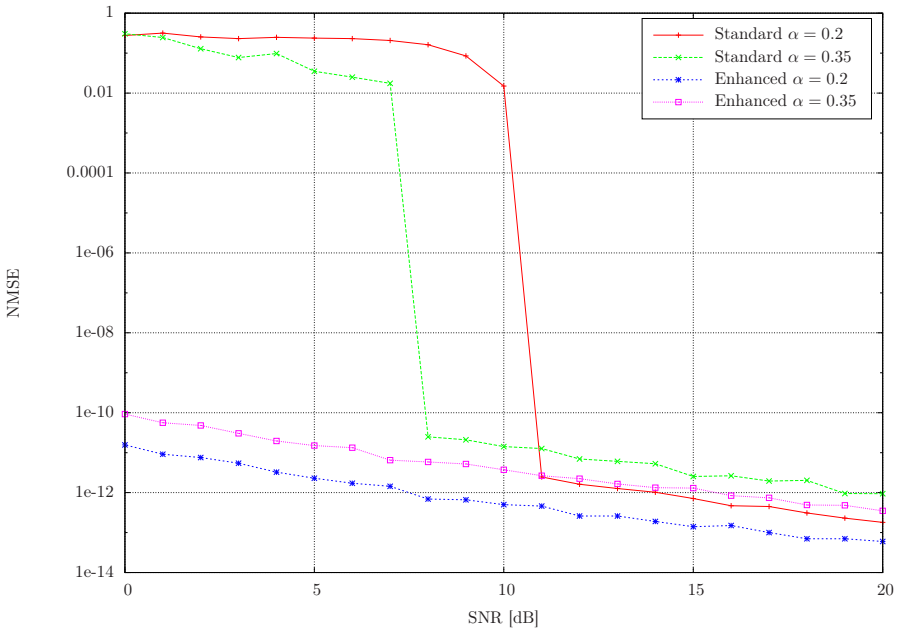


Figure 3.5: Evolution of the NMSE of $\hat{f}_d$ for different roll-offs α ($N = 2^{12}$, $M = 50$, $N_s = 4.5$)

is 4096; oversampling at $N_s = 4$ results in 16384 samples. The residual frequency offset is set to $\Delta fT = 0.001$. The length of the FFT is assumed to be $N = 2^{12}$ such that $M = 13$ overlapping sample blocks can be processed. The evolution of the success rate is illustrated in Figure 3.8. It can be seen that the enhanced algorithm is considerably superior to the standard approach and only outperformed by the filter bank method. For the sake of completeness, it is shown that increasing the observation interval to $M = 50$ overlapping blocks leads to a significant improvement of the success rate.

3.3.2 Refinement Stage

In the refinement stage a residual timing drift on the signal may be used for enhancing the precision of the symbol rate estimate $\hat{f}_d$ [8]. This is mainly due to the fact that any deviation from the true value f_d causes a timing drift and, hence, a timing error accumulated in a timing tracker. The block diagram of this refinement stage is sketched in Figure 3.9. According to the ratio of the sample rate conversion, appropriate filtering has to be applied to avoid aliasing effects. The resampling unit converts the filtered signal from $\hat{N}_s$ to a lower rate, which reduces the processing complexity in the following stages. The minimum required value is determined by the used estimator algorithms, e. g., $N_s' = 4$ for the Oerder and Meyr (OM) device [25]. If possible, the resampling module

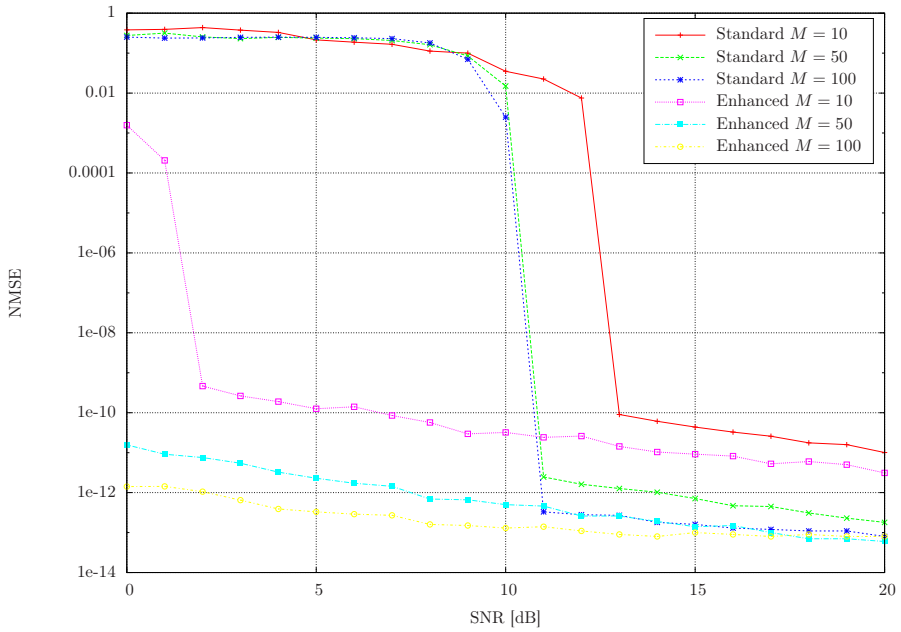


Figure 3.6: Evolution of the NMSE of $\hat{f}_d$ for different average numbers M ($N = 2^{12}$, $\alpha = 0.2$, $N_s = 4.5$)

could be moved in front of the MF to reduce computational complexity. Resampling techniques are presented in Section 3.4. For initial timing acquisition the OM method and for timing tracking a modification of the prominent Gardner detector (GA) [26], the extended Gardner detector (xGA) [27], are used. Successful tracking of signals drifting in time necessitates a second-order loop with an appropriate filter design in terms of damping ζ and normalized equivalent noise bandwidth $B_L T$. The design of second-order tracking loops is discussed in Section 3.13.3. At this point, it should be emphasized that the procedure is only applicable for SNR values, which can be handled by the selected acquisition and tracking algorithms. Furthermore, a trade-off between the ability to track a drifting signal and the accuracy of the accumulated timing error τ_L has to be made. Assuming an accumulated timing offset τ_L, given in symbol units and accumulated over an observation interval of L samples, the refined estimate of N_s appears as

$$\hat{N}_s' = \hat{N}_s \left(1 + \frac{\tau_L}{L} \hat{N}_s\right) \tag{3.5}$$

Figure 3.10 shows the performance of the refinement stage for different initial errors of the oversampling ratio $p_e = |\hat{N}_s / N_s - 1|$ and normalized equivalent noise bandwidths $B_L T$. The number of processed samples is $L = 2^{16}$ and 1000 iterations are performed for each

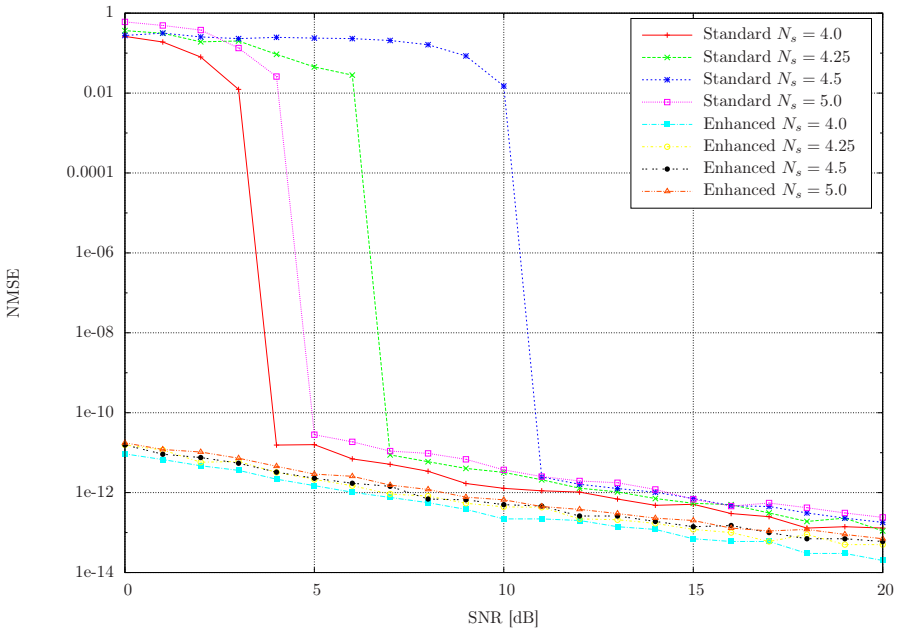


Figure 3.7: Evolution of the NMSE of $\hat{f}_d$ for different oversampling ratios N_s ($N = 2^{12}$, $M = 50$, $\alpha = 0.2$)

simulation point. It can be seen that for an initial error of $p_e = 1\,\%$ and a loop bandwidth of $B_L T = 10^{-3}$, the NMSE can be decreased by $3 - 4$ order of magnitudes down to an SNR of $\gamma_s = 9\,\mathrm{dB}$. For $B_L T = 10^{-4}$ the tracking loop is not capable to handle the timing drift produced by $p_e = 1\,\%$ of estimation error and thus the NMSE is large over the whole SNR range. On the other hand, when decreasing the error to $p_e = 0.1\,\%$ the tracking loop can handle the drift and a significant refinement can be achieved.

To summarize, the necessity of this refinement arises mainly, if the accuracy of the symbol rate estimation stage is not sufficient. In general, a reasonable design of the proposed symbol rate estimator in Section 3.3.1 allows omitting the refinement step. If the accumulated timing error in the symbol timing tracker is small enough, then the refinement step may be skipped at all. A possible limit could be 0.001 correcting percentage.

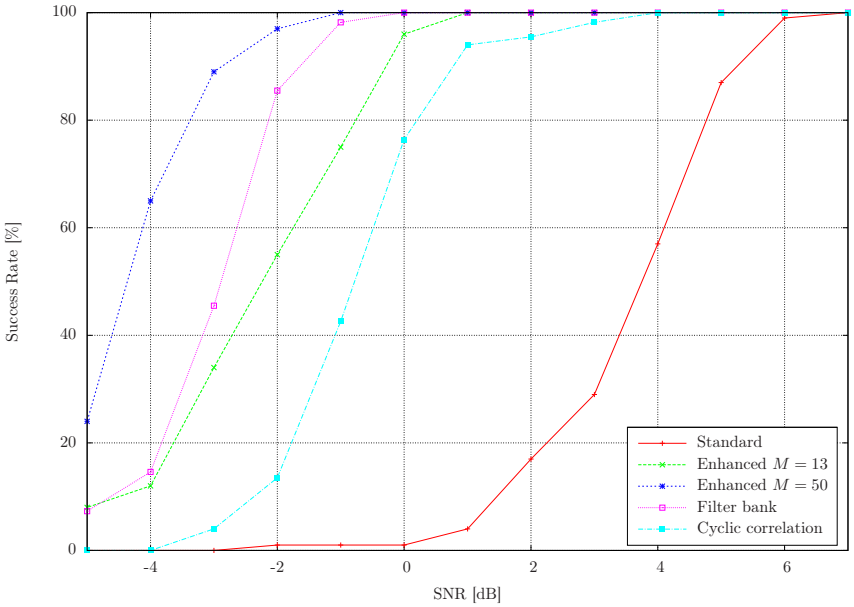


Figure 3.8: Evolution of the $\hat{f}_d$ success rate ($N = 2^{12}$, $M = 13$, $\alpha = 0.2$)

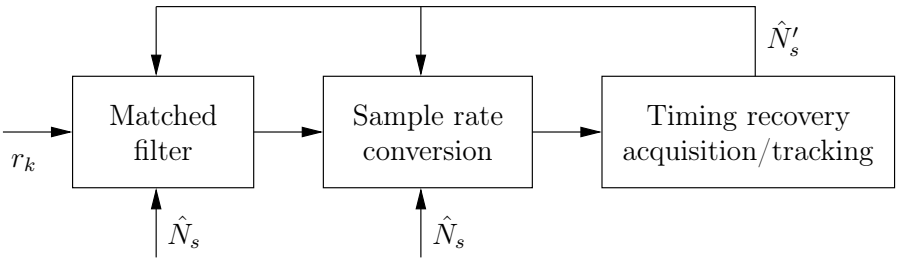


Figure 3.9: Block diagram for the symbol rate refinement stage

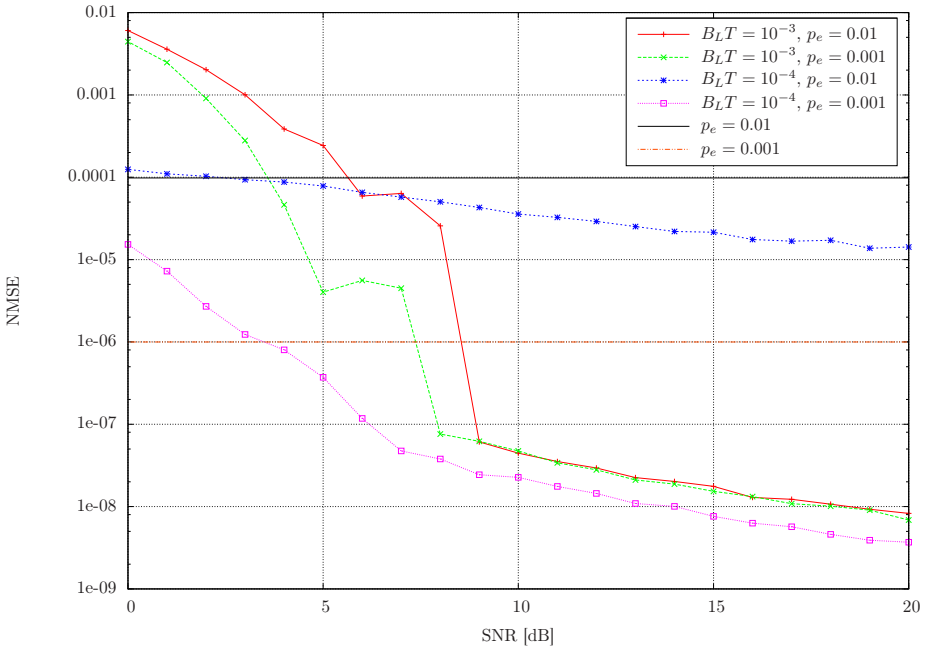


Figure 3.10: Evolution of the $\hat{f}'_d$ NMSE for different normalized equivalent noise bandwidths B_LT ($L = 2^{16}$, $\alpha = 0.2$, $N_s = 4.5$)

3.4 Resampling

To lower the computational complexity in the subsequent stages, the oversampling ratio N_s is reduced to the minimum required value N_s'. In the current implementation the minimum value of $N_s' = 4$ is given by the OM device [25]. Basically, during resampling the underlying continuous-time signal $r(t)$ has to be determined for values that do not coincide with the available samples at $r(kT_s)$. Due to the vast range of possible symbol rates, the resampler has to be capable of processing arbitrary oversampling ratios. Standard techniques that first interpolate by an integer factor L and afterwards decimate by an integer M can change the sampling rate by a rational factor [28]. For non-rational rate changes, as they arise in practical situations where symbol and sampling period are derived from different clocks [29], or varying interpolation indices as needed for correction of the estimated timing error, different approaches have to be selected.

3.4.1 Ideal Interpolation

If the received signal satisfies the sampling theorem

$$f_s = \frac{1}{T_s} \geq 2\, B_{AAF} \tag{3.6}$$

where B_{AAF} denotes the bandwidth of the anti-aliasing filter (AAF), then the underlying continuous-time signal $r(t)$ can be recovered perfectly by the interpolation formula [19]

$$r(t) = \sum_{k=-\infty}^{\infty} h_I(t - k \cdot T_s)\, r\,(k \cdot T_s) \tag{3.7}$$

where impulse response

$$h_I(t) = \frac{\sin(\pi\, t/T_s)}{\pi\, t/T_s} \tag{3.8}$$

is of infinite duration and any truncation would lead to a distortion of the interpolated signal. Alternatives are to apply a window on the ideal impulse response $h_I(t)$ to allow truncation [30] or to replace $h_I(t)$ by piecewise polynomial functions of short duration [31]. For resampling and timing correction frequently a cubic Lagrange interpolator is used, since accuracy is sufficient for synchronization purposes.

3.4.2 Lagrange Interpolation

To determine the value of $r(t)$ at time t_n, the value t_n has to be split up into the preceding sample index l_n and the remaining fractional part μ_n, such that t_n can be written as

$$t_n = (l_n + \mu_n)T_s \tag{3.9}$$

with

$$l_n = \mathrm{int}\left(\frac{t_n}{T_s}\right), \quad \mu_n = \mathrm{frc}\left(\frac{t_n}{T_s}\right) \tag{3.10}$$

where $\mathrm{frc}(x) = x - \mathrm{int}(x)$. The illustration in Figure 3.11 explains the introduced relationships. Substitution of the above formulas into Equation 3.7 and assuming that the

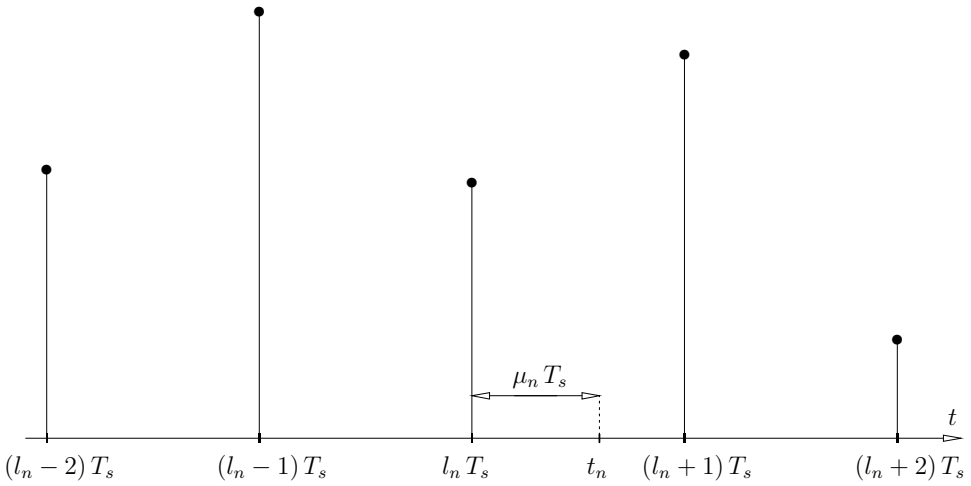


Figure 3.11: Explanation of basepoint and fractional index for Lagrange interpolation

impulse response $h_I(t)$ is limited to the range $-I_1\,T_s \leq t \leq (I_2 + 1)\;T_s$ yields

$$r(t_n) = \sum_{k=l_n-I_2}^{l_n+I_1} h_I((l_n - k)\cdot T_s + \mu_n\,T_s)\,r\,(k\cdot T_s) \tag{3.11}$$

Finally changing the summation index to $i = l_n - k$ leads to the interpolator formula

$$r(t_n) = \sum_{i=-I_1}^{I_2} h_I(i\cdot T_s + \mu_n\,T_s)\,r\,((l_n - i)\cdot T_s) \tag{3.12}$$

Additionally to the supporting points $r(kT_s)$ in the range of $k = l_n - I_2, \ldots, l_n + I_1$, only the fractional index μ_n has to be provided. Thus the coefficients $c_i(\mu_n) = h_I(i\cdot T_s + \mu_n\,T_s)$ of a $(n-1)$-order Lagrange interpolator can be computed as

$$c_i = \prod_{\substack{k\,=\,0 \\ k\,\neq\,i}}^{n-1} \frac{\mu_n - k}{i - k} \tag{3.13}$$

The block diagram of a cubic interpolator is depicted in Figure 3.12. The coefficients can either be calculated online or be stored in a look up table (LUT), when discretizing

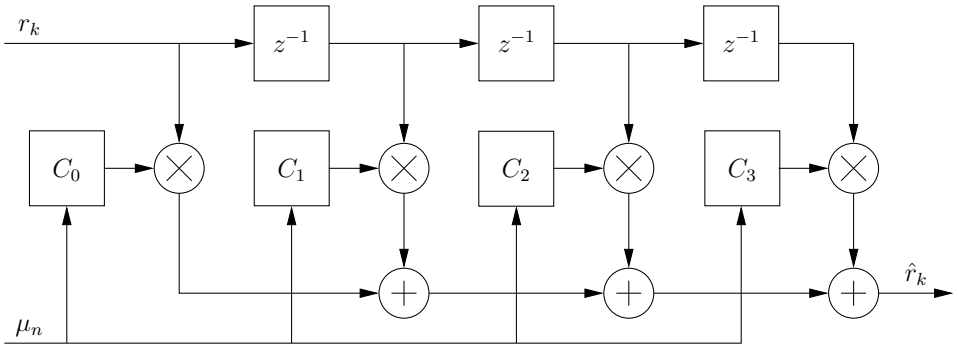


Figure 3.12: Block diagram of the cubic Lagrange interpolator

the possible values of μ_n. If the input and output oversampling ratios N_s and N_s' are incommensurate, the coefficients have to be recalculated in every step. A control unit is used to ensure that the required samples are loaded into the shift register. Special care has to be taken, if the fractional index exceeds $\mu_n > 1$, then a sample has to be inserted into the shift register and μ_n must be decremented. This process has to be repeated until $\mu_n < 1$.

In the following, simulation results show the performance of the implemented algorithm. QPSK is used as modulation scheme, the roll-off factor is set to $\alpha = 0.2$ and the SNR to $\gamma_s = 50\,\text{dB}$. In Figure 3.13 the time domain results (real part) using cubic interpolation for decimation from $N_{s1} = 8.5$ to $N_{s2} = 4.0$ are shown. The rate change would correspond to resampling the signal at $4/8.5 = 0.470588$ of the original rate and could be accomplished by upsampling of $L = 1000$ with subsequent decimation of $M = 2125$. It is obvious that such high rate changes are not favorable or may even not be feasible from the computational point of view. To gain more insight into the accuracy of the interpolator the resampled spectrum should be investigated on a logarithmic scale, as depicted in Figure 3.14. It can be seen that the interpolated spectrum is affected by the interpolating filter, however the attenuation is sufficient for modem and most monitoring applications. As illustrated, the produced images are canceled out by the following MF. If further attenuation of the spectral images is required, an interpolation approach using the Kaiser window could be used [30]. If the decimation ratio corresponds to an integer, the interpolation error will vanish since the fractional index $\mu_n = 0$ and the interpolated values coincide with the supporting points in the shift register. An example is shown in Figure 3.15 for decimation from $N_{s1} = 8.0$ to $N_{s2} = 4.0$.

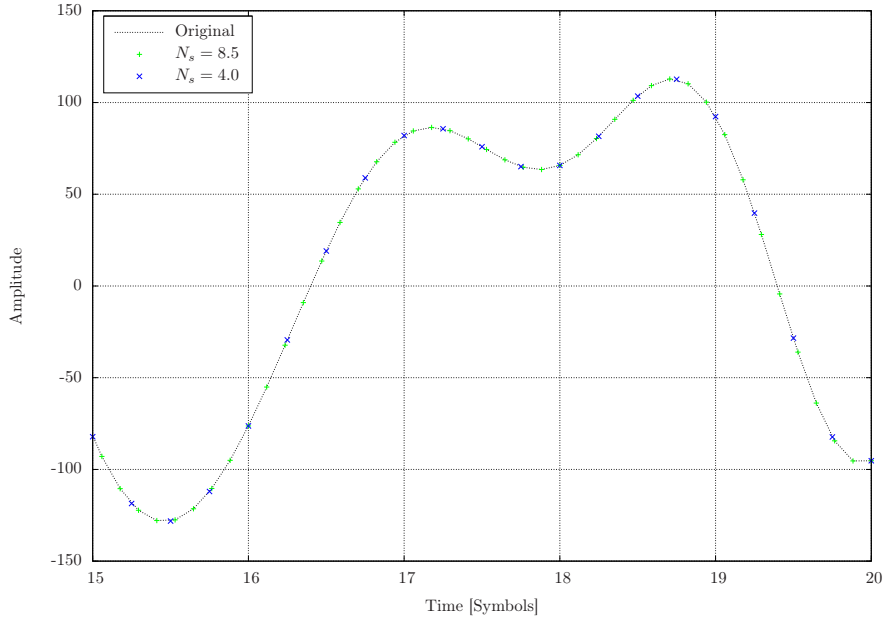


Figure 3.13: Waveform of Lagrange interpolation $N_{s1} = 8.5$ to $N_{s2} = 4.0$

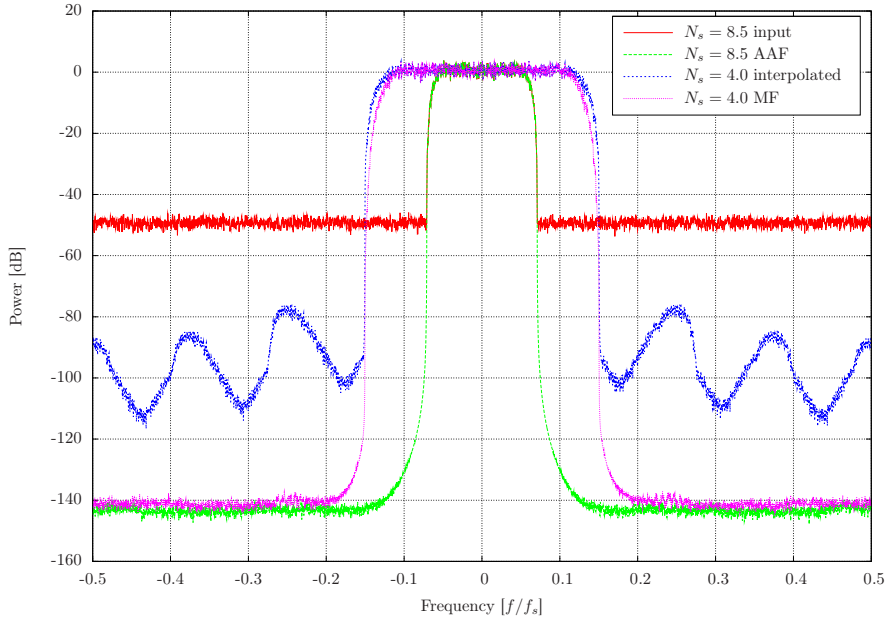


Figure 3.14: Spectra of Lagrange interpolation $N_{s1} = 8.5$ to $N_{s2} = 4.0$

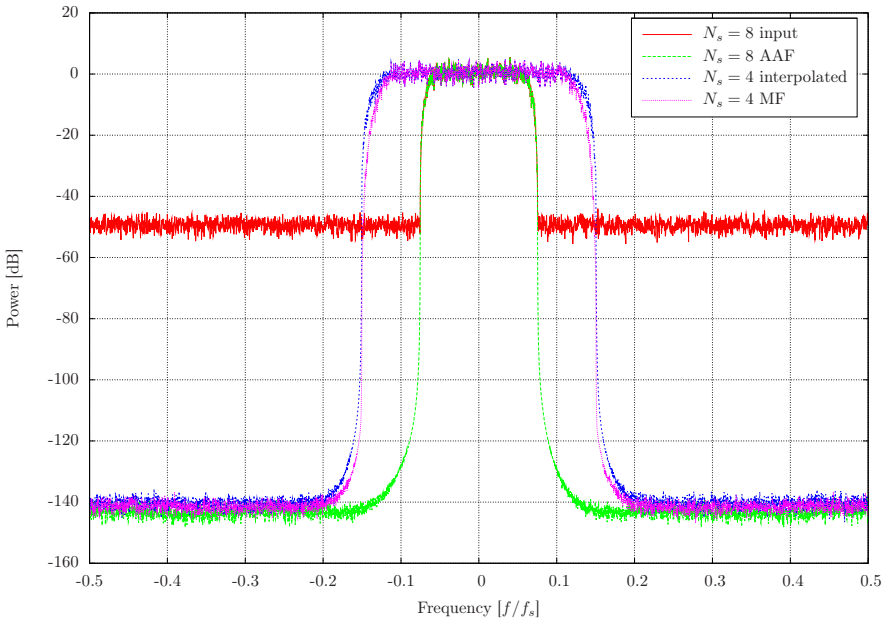


Figure 3.15: Spectra of Lagrange interpolation $N_{s1} = 8.0$ to $N_{s2} = 4.0$

3.5 Roll-off Factor Estimation I

As will be shown in Section 3.6, the only parameter remaining for the design of the receiving filter is the roll-off factor α. Thus the question arises, how to determine α to be able to continue further in the demodulation process? Publications available in the open literature dealing on this topic are very rare. The only paper found is [32], which is dedicated to this task. Basically, the presented method relies on the fact that, when inspecting the magnitude of the RCOS pulse $|g(t)|$, the ratio of the first side lobe to the main lobe exhibits a dependence on the roll-off factor α. The PSD of the received signal $P_r(f)$ is composed of the PSDs of signal $P_s(f)$ and noise $P_n(f)$ as

$$P_r(f) = P_s(f) + P_n(f) \tag{3.14}$$

Since the noise power is assumed to be constant over the whole frequency range, using the histogram method, described in Section 3.1.1, delivers a robust estimate of the former. Thus, the noise power can be subtracted from $P_r(f)$ to get an estimate of $P_s(f)$. Furthermore applying the inverse fast Fourier transform (IFFT) on $P_s(f)$ results in the baseband pulse $g(\alpha,t)$, where α should stress the fact that the shape of the baseband pulse depends on the roll-off. Taking the magnitude of the waveform leads to $|g(\alpha,t)|$, the curve which was used in [8] for coarse estimation of the symbol rate. The maximum of the first sidelobe of $|g(\alpha,t)|$ is located in the range $t_{max} \in [T,2T)$. As mentioned initially, the ratio $R(\alpha)$ defined as

$$R(\alpha) = \frac{|g(\alpha,t_{max})|}{|g(\alpha,0)|} \tag{3.15}$$

is a function uniquely defined by α. The evolution of $R(\alpha)$ is illustrated in Figure 3.16. Previous to the estimation process, $R(\alpha)$ is evaluated over the possible range $\alpha \in [0,1]$ in a resolution of $\Delta\tilde{\alpha} = 0.01$ and stored in a LUT as $R(\Delta\tilde{\alpha}\, n)$ with $n = 0,1,2,\ldots$ As described above, the PSD of the received samples $\hat{P}_r(f)$ and the noise power $\hat{P}_n(f)$ are used for estimation of the waveform $|\hat{g}(\alpha,t)|$. The location of the maximum of the first sidelobe is aided by using zero-padding for the IFFT and parabolic interpolation (see Appendix A.1), such that an estimate $\hat{R}(\alpha)$ can be calculated. The minimum distance between the stored values $R(\Delta\tilde{\alpha}\, n)$ to the estimate $\hat{R}(\alpha)$ is searched and delivers an estimate for the roll-off factor. Since a resolution of $\Delta\alpha = 0.05$ is common practice for designing the filter roll-off, the preliminary result is rounded to the closest value exhibiting the $\Delta\alpha$ resolution. So the estimator can be summarized as

$$\hat{\alpha} = \left\lfloor \frac{\Delta\tilde{\alpha}\, \arg\min_n |\hat{R}(\alpha) - R(\Delta\tilde{\alpha}\, n)|}{\Delta\alpha} + 1/2 \right\rfloor \Delta\alpha \tag{3.16}$$

Where the operator $\lfloor x \rfloor$ corresponds to the floor operation which rounds x to the next smaller integer value.

In the following, simulation runs demonstrate the performance of the above estimator.

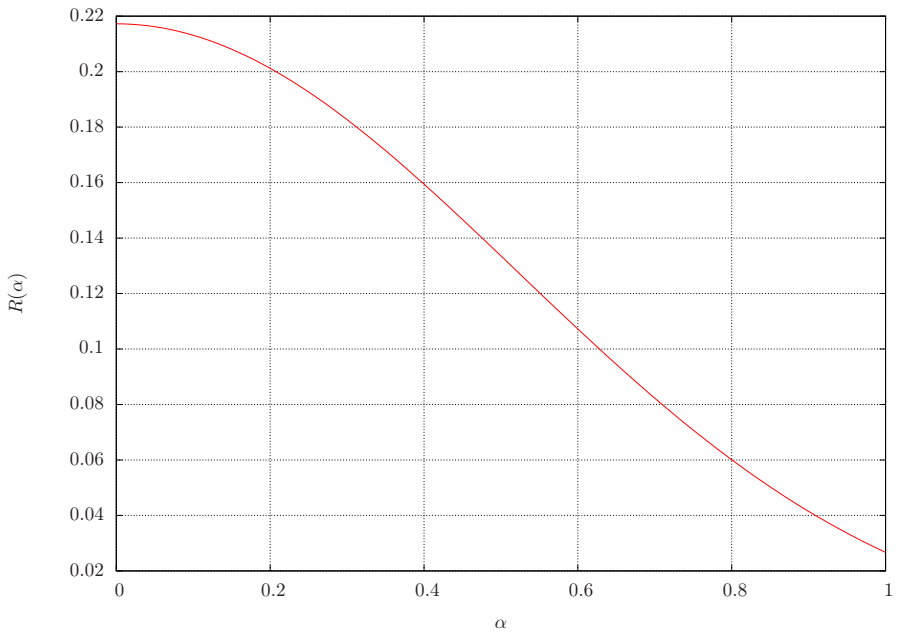


Figure 3.16: Evolution of $R(\alpha)$

The periodogram uses an FFT with $N = 2048$ points, averaged 125 times with an overlap of 75 %. This results in 65536 samples, when oversampling 16384 symbols at a rate of $N_s = 4$. The zero-padding factor for the IFFT is set to $k_{zp} = 2$. For each simulation point 1000 iterations are performed. In Figure 3.17 and Figure 3.18 the mean output and mean square error (MSE) of the estimator are depicted for several signal constellations and a roll-off factor of $\alpha = 0.35$. As expected, the performance is independent of the used modulation scheme, which is in contrast to the SNR based roll-off estimator presented later on in Section 3.11. The estimator mean exhibits a slight bias which grows with decreasing SNR. Nevertheless, the achieved accuracy should be sufficient since a minor deviation from the true roll-off results only in a small SNR degradation, as will be shown in Section 3.6.2. This fact could be used when lacking computational power to apply the pragmatic solution of setting the roll-off factor to a value right in the middle of the expected range, i.e. $\alpha \in [0.2,0.5]$. A more precise estimation could be performed later on in the blind demodulation process. Up to now the performance of the above algorithm was only validated for a single roll-off value $\alpha = 0.35$. Hence, in the following the behavior over the whole roll-off range is investigated. Figure 3.19 and Figure 3.20 show the mean estimator bias and the MSE for different SNR values. As can be seen, there is a trend towards increased bias at the edges of the simulated range. This offset is most pronounced for small roll-off values at an SNR of $\gamma_s = 0\,\mathrm{dB}$. The main reason for this effect is erroneous noise power estimation.

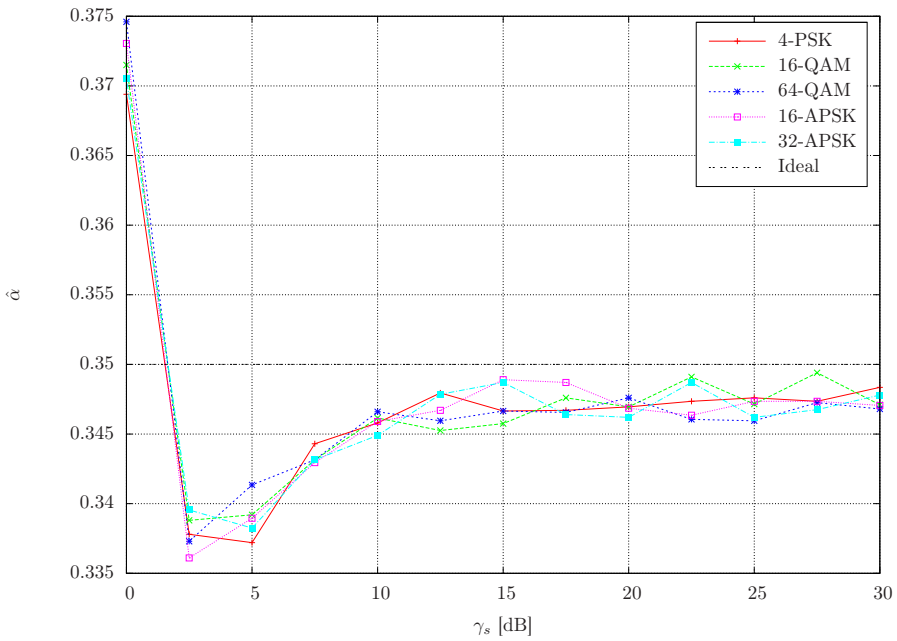


Figure 3.17: Roll-off estimator output for different modulation schemes

The above simulation runs are based on the assumption of a single-carrier scenario with constant noise. Two problems could arise for multi-carrier signals, if narrowband filters are applied for carrier separation. First, the selection of a too large filter bandwidth may result in performance degradation due to interference by adjacent carriers. Second, the histogram method does not work reliably for noise power estimation, since the noise floor is filtered out by the sharp filters. As a way out of this dilemma using the noise level estimate from the pre-estimation stage is suggested. However, the filter characteristics have to be considered when subtracting the noise power from $\hat{P}_r(f)$.

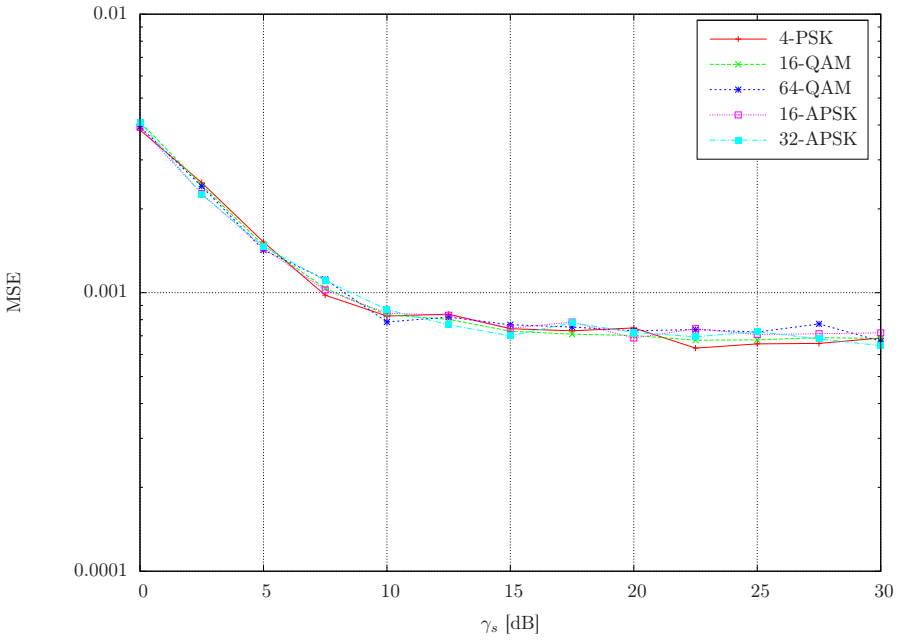


Figure 3.18: MSE of $\hat{\alpha}$ for different modulation schemes

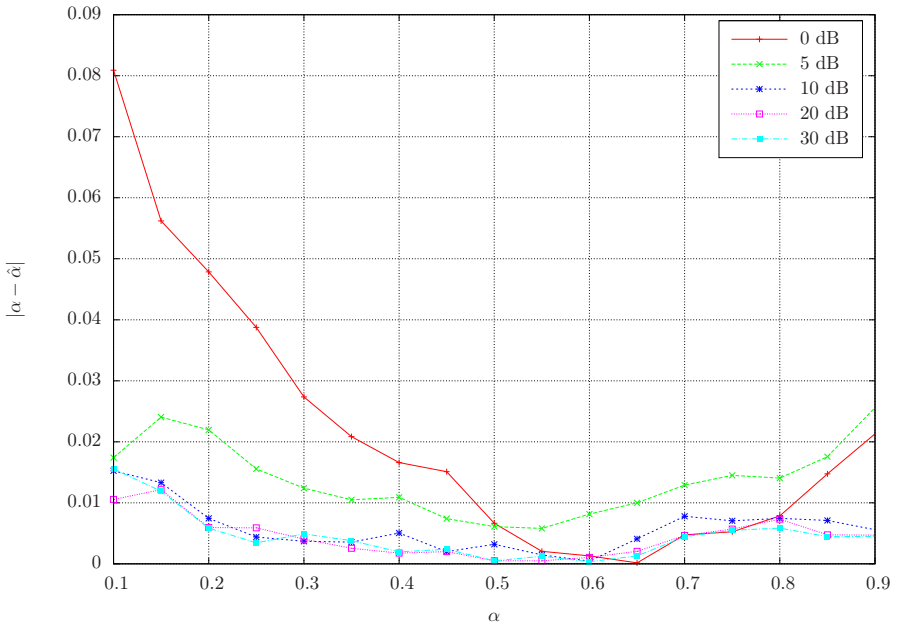


Figure 3.19: Mean roll-off estimator bias for different roll-off factors

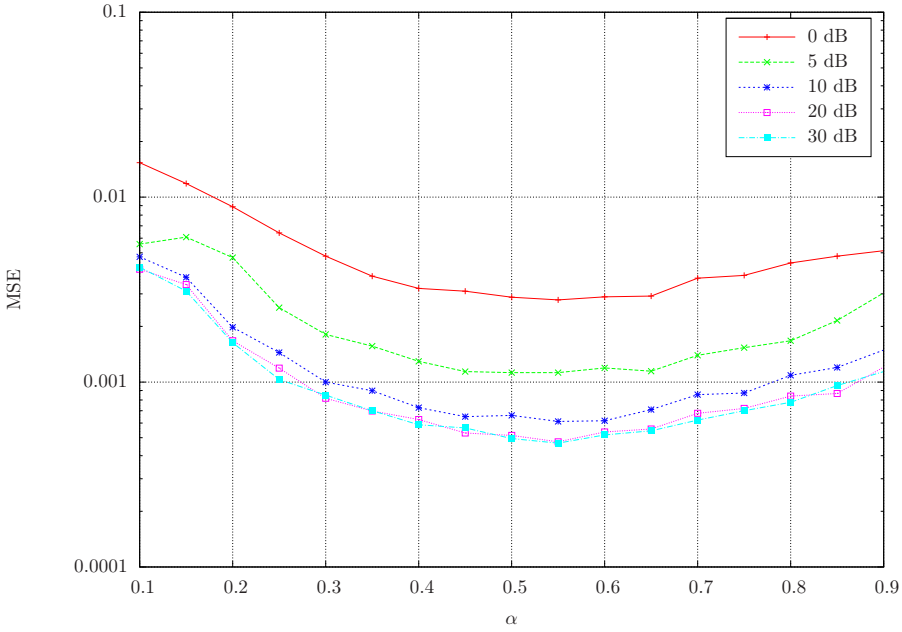


Figure 3.20: MSE of $\hat{\alpha}$ for different roll-off factors

3.6 Matched Filter

The MF in the receiver should provide the maximum SNR at its output for the received waveform. Assuming a flat spectrum of the channel, the impulse response of the MF is a conjugate-mirrored version of the received signal [33]. Since the satellite channel can be viewed as an LTI system with the received signal as a superposition of baseband pulses, the impulse response of the MF can be written as

$$h(t) = h^*(-t) \tag{3.17}$$

3.6.1 Raised Cosine Filter

For baseband filtering, RCOS pulses are established as the de facto standard in satellite communications. The spectrum $G(f)$ is defined as

$$\begin{aligned} G(f) &= |H(f)|^2 = \\ &= \begin{cases} T, & |f\,T| \leq \frac{1}{2}\,(1-\alpha) \\ T \cdot \cos^2\left(\frac{\pi}{2\,\alpha}\,\left(|f\,T| - \frac{1-\alpha}{2}\right)\right), & \frac{1}{2}\,(1-\alpha) < |f\,T| \leq \frac{1}{2}\,(1+\alpha) \\ 0, & |f\,T| > \frac{1}{2}\,(1+\alpha) \end{cases} \end{aligned} \tag{3.18}$$

and depicted in Figure 3.21 for different values of the roll-off factor α. The related impulse response is

$$g(t) = \mathrm{sinc}\left(\frac{t}{T}\right)\,\frac{\cos\left(\frac{\alpha\,\pi\,t}{T}\right)}{1-\left(\frac{2\,\alpha\,t}{T}\right)^2} \tag{3.19}$$

and is illustrated in Figure 3.22. It should be pointed out that $g(mT) = 0$ for $m \neq 0$. Hence, the data symbols are not distorted by adjacent pulses, if the optimum timing instant is recovered successfully. The effect when neighboring pulses are influencing each other is termed ISI. From the above illustrations, showing $G(f)$ and $g(t)$, the effect of the roll-off factor α becomes obvious. It is clear that the latter can be used to tradeoff bandwidth efficiency and sensitivity to ISI. In practical realizations the filter is split up into two RRCOS filters, with frequency response $H(f) = \sqrt{G(f)}$ each, located in the transmitter and receiver, respectively. This guarantees that the overall baseband pulse $g(t) = h(t) \otimes h^*(-t)$ becomes RCOS and thus is free of ISI.

3.6.2 Filter Mismatch

It turns out that a small deviation from the roll-off factor in the receiving filter to that used in the transmitting filter results only in a minor degradation of the SNR at the filter output. Hence, the determination of α is not as critical as one might expect. In the following, the above stated assumptions are confirmed by simulation runs. In Figure 3.23 an example of the filtered signal at the matched output x_k in the absence of noise is shown.

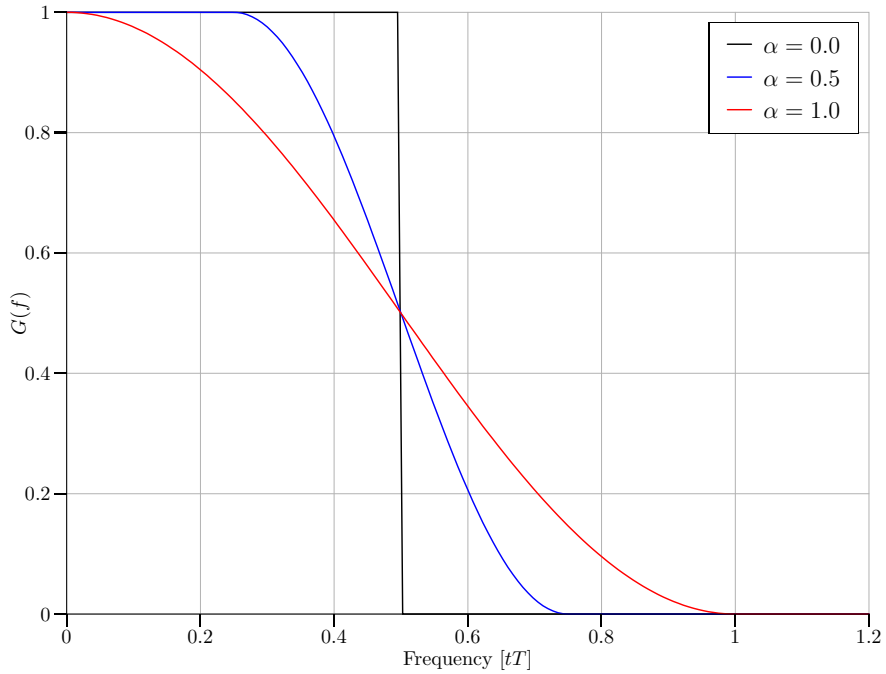


Figure 3.21: Evolution of the RCOS spectrum

Therein the ideal signal, using the same roll-off for transmitting and receiving filter $\alpha = 0.2$, is compared to the signal with wrong roll-off at the receiver $\alpha = 0.35$. To ease comparison the transmitted data symbols c_n are shown as well (blue dots). The mismatched signal exhibits less overshoot between the symbols due to the smaller sidelobes of the impulse response. In the optimum timing instant both signals seem to be quite similar, however, when zooming in the mismatched signal deviates more from the data symbols. To show that the impact of a minor mismatch is negligible for synchronization, an estimation of the SNR is performed at the output of the MF. Figure 3.24 illustrates the evolution of the estimated SNR at the MF output for mismatched receiving filters. The simulation is repeated for different values of the transmitting excess bandwidth $\alpha \in \{0.2, 0.25, 0.35, 0.5\}$. As expected, the estimated SNR has a maximum when transmitting and receiving filters match. The shape of the maximum is rather flat such that a small roll-off mismatch does not affect the SNR significantly. For the estimation of the SNR, the algorithms detailed in Section 3.10 are used, which require recovery of the symbol timing as will be explained in the following.

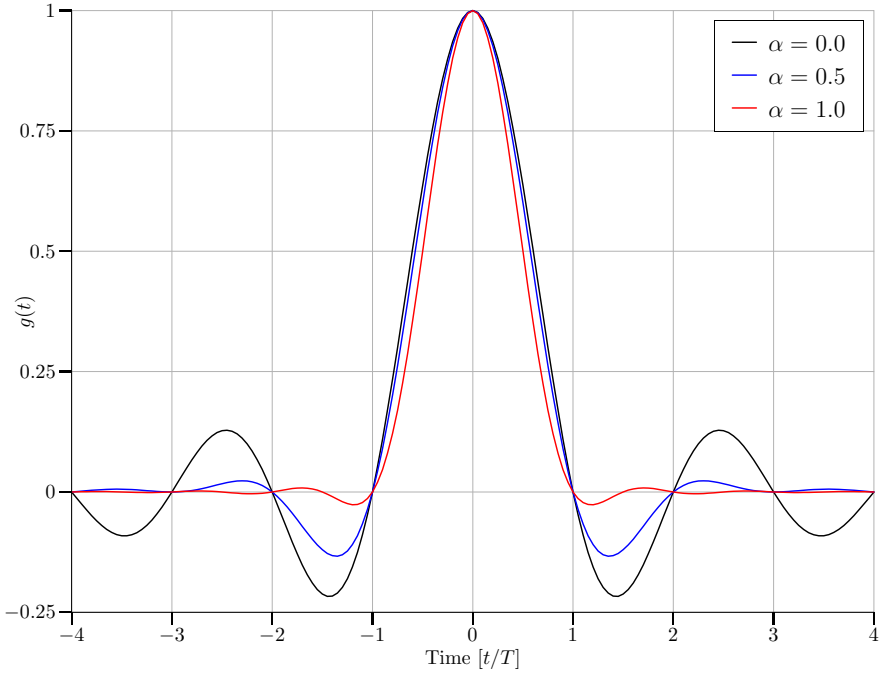


Figure 3.22: Evolution of the RCOS pulse $g(t)$

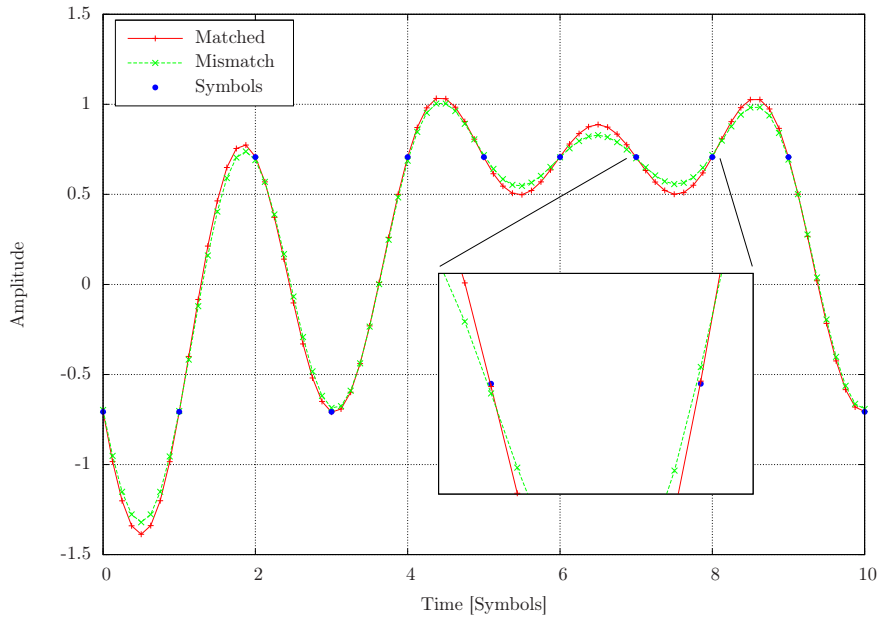


Figure 3.23: Impact of matched ($\alpha = 0.2$) and mismatched ($\alpha = 0.35$) receiving filters

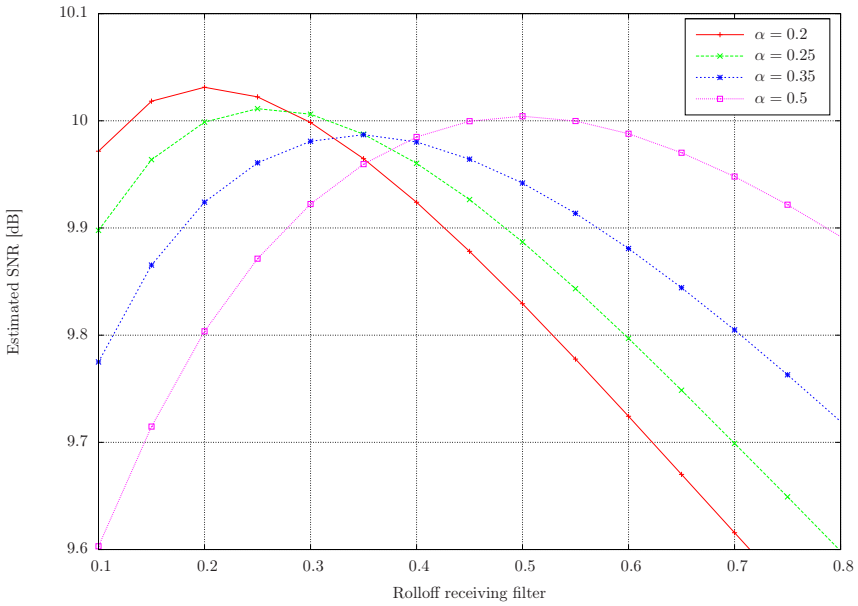


Figure 3.24: SNR at MF output for different roll-off factors

3.7 Rescaling

Rescaling of the received waveform is necessary to guarantee proper operation of most of the subsequent stages, e.g. design of filter coefficients in the timing tracker, partitioning used for SNR estimation, frequency and phase recovery or, finally, for the decision of the transmitted data symbol drawn from a multi-level modulation scheme. Using the pre-estimate $\hat{\gamma_s}^*$ for the SNR and an estimate for the second-order moment M_{20}, an appropriate scaling factor a_s can be derived. As investigated in [34], for oversampled signals the second-order moment at the MF output is given as

$$M_{20} = E[|x_k^2|] = (1 - \alpha/4)E_s + E_n \tag{3.20}$$

This shows that the signal power at the MF output can be expressed as

$$E_s' = (1 - \alpha/4)E_s \tag{3.21}$$

Using the relationship for the SNR, i.e. $\gamma_s = E_s/E_n$, in Equation 3.20, the signal power ahead of the MF can be calculated as

$$E_s = \frac{\gamma_s M_{20}}{\gamma_s(1 - \alpha/4) + 1} \tag{3.22}$$

Now for proper scaling the signal has to be simply divided by $a_s = \sqrt{E_s}$.

3.8 Timing Offset Estimation

From the assigned problem it is clear that only NDA methods should be considered. Normally, timing error estimation and correction consists of two tasks: initial acquisition and tracking. The former is usually carried out by a feedforward (FF) algorithm, e. g. the solution proposed by Oerder and Meyr [25] or a modification thereof. The obtained result serves as a starting value for the following tracker module, frequently implemented as a Gardner detector (GA) [26]. A thorough analysis of the performance of the OM method and the GA can be found in [35]. Since symbol timing recovery is a major part of the demodulation process, a robust algorithm with low complexity and sufficient accuracy should be used. Alternatives to well established algorithms were compared and analyzed. A feedback solution by Moeneclaey and Batsele using only one sample per symbol is investigated in [36], reviewed in [10] and the feedforward counterpart is presented in [11].

3.8.1 Acquisition

Oerder and Meyr Method The Oerder and Meyr (OM) algorithm [25] is independent of modulation scheme and roll-off factor of the baseband pulse. However, it requires at least an oversampling factor of $N_s \geq 3$, for computational reasons setting $N_s = 4$ is frequently

suggested. The design parameter ν is used for calculating $|x_k|^\nu$, which represents the magnitude for $\nu = 1$ or the squared magnitude when $\nu = 2$. For both values of ν the resulting signal contains a spectral component at the symbol rate $f_d = 1/T$. The discrete Fourier transform (DFT) can be used to extract the spectral component X_L by

$$X_L = \sum_{k=0}^{LN_s-1} |x_k|^\nu \, e^{-\jmath\, 2\,\pi\, k/N_s} \tag{3.23}$$

The magnitude of X_L is of minor interest in this context, however, the argument can be used to extract an unbiased estimate of the timing offset using

$$\hat{\tau} = -\frac{T}{2\,\pi} \arg(X_L) \tag{3.24}$$

As already stated above, it becomes evident that the OM method is totally independent of modulation scheme and baseband pulse roll-off and thus can be used at this stage of the blind demodulation process. Figure 3.25 illustrates the usage of the OM method in a feedforward manner to estimate and correct the timing offset. The block diagram includes an additional delay and an interpolator which can be of Lagrange type as introduced in Section 3.4.2.

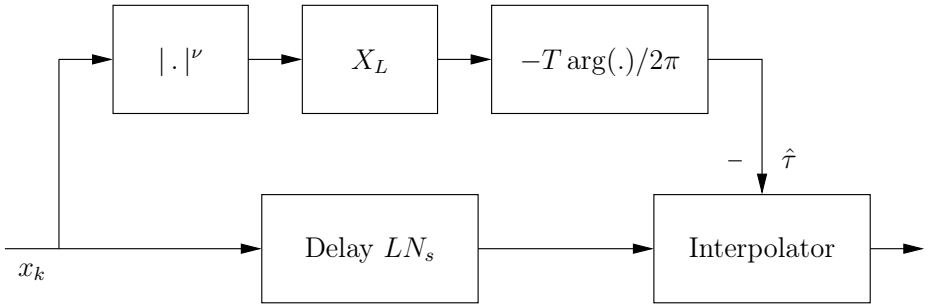


Figure 3.25: Feedforward timing recovery using OM method

Moeneclaey and Batsele In [11] the attempt is made to operate the feedback timing recovery scheme by Moeneclaey and Batsele (MB) [36], working on only one sample per symbol, in a feedforward manner. The main idea is to combine the advantages of feedforward schemes with the low complexity due to baud-rate sampling. Two possible algorithms exhibiting similar jitter performance are introduced.

Estimator 1 Using the error signal u_k of the MB feedback method stated in Equation 3.38, an estimator operating on an observation interval of L symbols can be furnished

as

$$U_0 = \frac{1}{L}\sum_{k=0}^{L-1} u_{kN_s} \tag{3.25}$$

$$U_1 = \frac{1}{L}\sum_{k=0}^{L-1} u_{kN_s - N_s/4} \tag{3.26}$$

$$\hat{\epsilon} = \frac{1}{2\pi}\arg(U_1 + \jmath U_0) \tag{3.27}$$

To achieve a satisfactory performance two iterations of the algorithm must be carried out with the drawback of increased computational complexity. Another deficiency of the introduced method is, when assuming equidistant sampling, the requirement of at least an oversampling ratio of $N_s = 4$. However, it should be pointed out that the algorithm is independent of the modulation scheme and the roll-off factor of the baseband pulse.

Estimator 2 The second estimator operates on U_0 and uses the threshold V_0 to be able to distinguish between $|\epsilon| \leq 1/4$ and $1/4 < |\epsilon| \leq 1/2$. The calculation of V_0 requires two samples per symbol and can be formulated as

$$V_0 = \frac{1}{L}\sum_{k=0}^{L-1} \left(|x_{kN_s}|^2 - |x_{kN_s+N_s/2}|^2\right) \tag{3.28}$$

The expected value of V_0 is proportional to $\cos(2\pi\epsilon)$, thus the second estimator can be written as

$$\hat{\epsilon} = \begin{cases} \arcsin(2\pi U_0/K_d)/(2\pi) & \text{if } V_0 \geq 0 \\ 1/2 - \arcsin(2\pi U_0/K_d)/(2\pi) & \text{if } V_0 < 0,\ U_0 \geq 0 \\ -1/2 - \arcsin(2\pi U_0/K_d)/(2\pi) & \text{if } V_0 < 0,\ U_0 < 0 \end{cases} \tag{3.29}$$

The estimator necessitates an oversampling ratio of $N_s \geq 2$, however three iterations of the algorithm have to be applied to achieve similar performance as the first estimator. For evaluation of the above formula the knowledge of the detector slope K_d at the stable equilibrium point $\epsilon = 0$ is required, thus the used modulation constellation and the roll-off factor α have to be known, as can be seen from Equations 3.39 and 3.40. Simulation results using non-matching parameters for timing recovery are shown further below. The oversampling ratio requirement for both MB estimators is in striking contrast to the initial intention of baud-rate sampling, however it is necessary to generate an appropriate detector characteristic.

3.8.2 Tracking

Gardner Detector Feedback schemes are attractive for timing tracking in many cases because of low complexity from the computational point of view. A prominent and widely

spread algorithm is the GA suggested in [26] which computes an error signal u_k according to

$$u_k = \mathrm{Re}\Big\{(x_k - x_{k-1})\, x^*_{k-1/2}\Big\} \tag{3.30}$$

Figure 3.26 shows the application of the GA for feedback timing recovery using a first-order loop. The same interpolation scheme as used for the feedforward scheme can be applied for

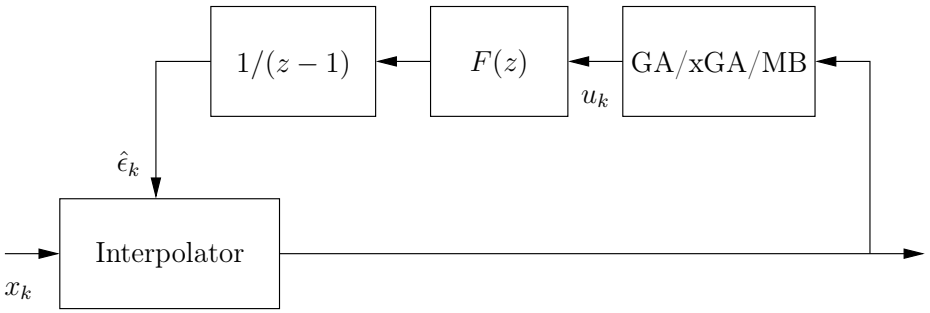


Figure 3.26: Feedback timing recovery using first-order loop

correction of the timing error. During timing tracking it can happen that the estimated timing error requires the interpolation of a value which is located before or after the current sample interval. Thus, again the importance of a control unit should be pointed out which takes care that the correct sample interval is placed appropriately in the shift register used for interpolation. In contrast to resampling, negative values of the fractional index μ_n can occur during tracking. In this case the samples inside the shift register have to be moved backwards to the appropriate interval.

The detector (or open-loop) characteristic, the so-called S-curve, is given by

$$S(\alpha,\epsilon) = E\left[u_k \,|\, \forall k\colon \epsilon = \text{const}, |\epsilon| \le 1/2\right] = \frac{\sin(\pi\,\alpha/2)}{\pi\,(1-\alpha^2/4)}\sin(2\,\pi\epsilon) \tag{3.31}$$

From the above equation it can be seen that the detector characteristic depends on the roll-off α, however, it is not affected by the SNR. The detector slope K_d at the stable equilibrium point $\epsilon = 0$ can be evaluated by

$$K_d = K_\infty = \frac{2\,\sin(\pi\,\alpha/2)}{1-\alpha^2/4} \tag{3.32}$$

In case of a first-order loop the filter transfer function $F(z)$ corresponds to a constant gain factor K_F. Using the equivalent noise bandwidth normalized to the symbol period, the filter coefficient can be calculated according to

$$K_F = \frac{4B_LT}{(1+2B_LT)K_d} \tag{3.33}$$

Hence, the design of the loop filter requires knowledge of the roll-off factor. The influence of setting α to a differing value, e.g. when the roll-off estimation stage in Section 3.5 is omitted, is discussed further below. For successful tracking of signals drifting in time, a second-order loop is necessary with an appropriate filter design in terms of damping ζ and normalized equivalent noise bandwidth $B_L T$, as shown in Section 3.13.3.

Extended Gardner Detector The main drawback of the original GA is the increasing self-noise for decreasing roll-off factors as required for bandwidth efficient communication. Significant self-noise can even deteriorate the accuracy of subsequent SNR estimation schemes in the high SNR range. A way out of this dilemma is to use an extended version of the GA as proposed in [37]. In this context, a generalized polar form of x_k is used to compute the error signal as

$$u_k = \mathrm{Re}\Big\{ (y_k - y_{k-1})\, x^*_{k-1/2} \Big\} \tag{3.34}$$

where

$$y_k = |x_k|^\mu e^{\jmath \arg x_k}, \mu \geq 0 \tag{3.35}$$

The design parameter μ is used to optimize the performance of the algorithm. For $\mu = 1$ the xGA coincides with the original detector. To achieve a reduced jitter variance, as well as for computational reasons, $\mu = 0$ seems to be an adequate choice. Unfortunately, the advantages of the xGA do not apply for modulation schemes exhibiting a non-constant envelope such as QAM or APSK [27]. In [37] the detector slope is developed as

$$K_d = K_\infty = \begin{cases} 2\dot{g}(-\frac{1}{2}T)T + \mu G(\alpha) & M = 2 \\ 2\dot{g}(-\frac{1}{2}T)T + \frac{1}{2}(\mu+1)G(\alpha) & M \geq 4 \end{cases} \tag{3.36}$$

where $\dot{g}(t_0) = (\partial/\partial t) g(t)|_{t=t_0}$ and

$$G(\alpha) = \sum_{M \neq 0} \dot{g}(mT)T\{g[(m-1/2)T] - g[(m+1/2)T]\} \tag{3.37}$$

From the above equations it gets obvious that the computation of K_d requires knowledge of the modulation scheme which is not available at this point of the blind demodulation process. Moreover, it should be remembered that $g(t)$ corresponds to the baseband pulse, which is assumed to be RCOS. The latter underlines that the xGA is dependent on the roll-off factor as well. If no estimate for the excess bandwidth and no knowledge about the modulation order is available, the xGA can only be operated in a suboptimal way.

Moeneclaey and Batsele A feedback scheme operating on baud-rate samples is proposed in [36] and reviewed in [10]. The error signal u_k is generated as

$$u_k = \mathrm{Re}\left[x^*_{k-1} x_k\right] \left(|x_k|^2 - |x_{k-1}|^2 \right) \tag{3.38}$$

and, as shown in [11], the detector characteristic can be approximated by

$$S(\alpha,\epsilon) = E\left[u_k \,|\, \forall k\colon \epsilon = \text{const}, |\epsilon| \leq 1/2\right] \approx \frac{K_d}{2\pi}\sin(2\pi\epsilon) \tag{3.39}$$

where K_d denotes the detector slope at $\epsilon = 0$ and can be expressed by

$$K_d = \frac{\partial S(\epsilon)}{\partial \epsilon}\Big|_{\epsilon=0} = \frac{4(2-K_4)\cos(\pi\alpha)}{\Lambda_d(1-4\alpha^2)} \tag{3.40}$$

Herein K_4 is the symbol kurtosis $K_4 = E_c[|c_k|^4]$ and Λ_d is a symmetry constant with $\Lambda_d = 1$ for BPSK and $\Lambda_d = 2$ for quadrature-symmetric constellations like PSK ($M \geq 4$), QAM or APSK. The main advantages of the MB are reduced self-noise in the high SNR range and no performance degradation with decreasing roll-off factors α as it happens for the GA. However, similar to the xGA these benefits are only valid for modulation constellations with constant envelope. Other modulation schemes, e.g. QAM or APSK, lead to severe performance degradation, as recently verified in [38].

3.8.3 Comparison of Algorithms

Jitter Performance One figure of merit for estimation algorithms is the evolution of the NMSE over the SNR. For comparison reasons it is common practice to use the Cramer-Rao lower bound (CRLB) or its modified version (MCRLB). This fundamental performance limit for symbol timing estimation is dependent on the SNR, estimator length L or normalized equivalent noise bandwidth B_LT, and the form of the baseband pulse. For feedforward algorithms it can be expressed as [19]

$$\text{MCRLB} = \frac{1}{2L[1/3\,\pi^2(1+3\alpha^2) - 8\alpha^2]\gamma_s} \tag{3.41}$$

and for feedback schemes as

$$\text{MCRLB} = \frac{B_LT}{[1/3\,\pi^2(1+3\alpha^2) - 8\alpha^2]\gamma_s} \tag{3.42}$$

For performance comparison of the different algorithms an enormous number of simulation runs was carried out. As a result the selection of the most appropriate timing recovery scheme depending on the received input signal was achieved. A fair jitter performance comparison of feedforward and feedback schemes is difficult due to their different nature. Feedforward methods perform acquisition in one step and tracking as such is not feasible. In contrast to feedback algorithms, which have to carry out acquisition successfully before entering the tracking mode. To continue the list of differences, feedforward methods are not able to track timing drifts, but they have a defined acquisition time and do not suffer from stability problems as it might occur with tracking loops.

To arrive at a treatable simulation setup, the timing error is set to $\epsilon = 0.0$. Hence

the tracking performance of the FB schemes is compared to the acquisition performance of the FF schemes. The parameters of the algorithms are set in such a way that the MCRLB of FF and FB methods coincide. Thus, the estimator length of the FF schemes is set to $L = 100$ and the normalized equivalent noise bandwidth of the FB methods to $B_L T = 0.005$. The NMSE is specified as average value of $(\epsilon_k - \hat{\epsilon}_k)^2$. In the first simulation run, QPSK is used as modulation scheme and the roll-off factor is set to $\alpha = 0.35$. The result is illustrated in Figure 3.27. It can be seen that OM and GA perform conveniently

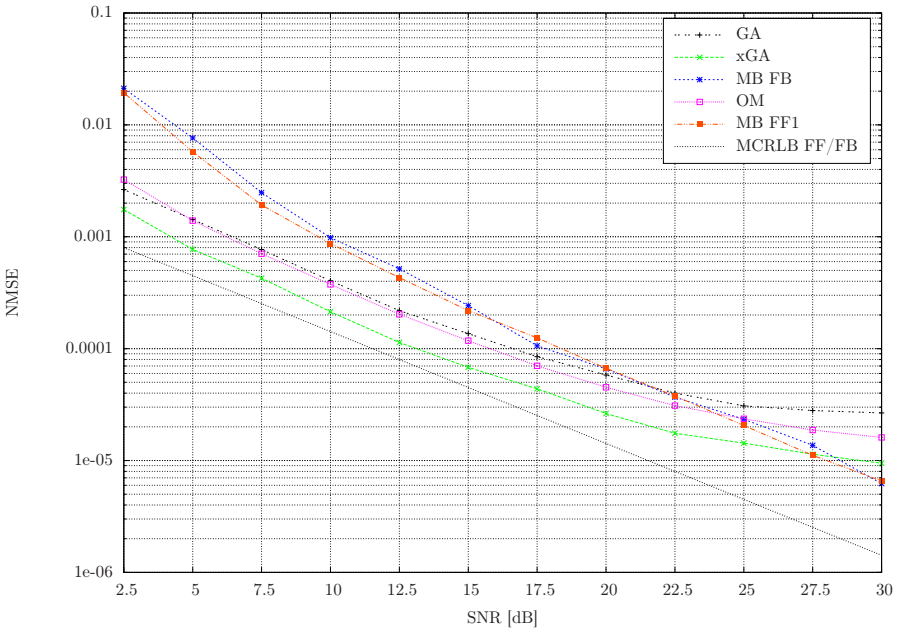


Figure 3.27: FF and FB NMSE of $\hat{\epsilon}$ for QPSK and $\alpha = 0.35$

in the low-to-medium SNR range; however, they suffer from self-noise at large SNR values. The evolution for the xGA has a very similar shape, but with decreased NMSE. Finally, the MB algorithm applied as FF and FB scheme is shown. Only Estimator 1 of the FF MB method is considered, since the characteristics are nearly identically to that of Estimator 2. It is evident that the NMSE is increased in the low-to-medium SNR range, but both algorithms do not exhibit a noise floor for high SNR values. The performance of OM and GA degrades significantly, if the roll-off factor is lowered to $\alpha = 0.2$, as depicted in Figure 3.28. In contrast to this, a minor degradation can be observed for the xGA and even a slightly improvement for the MB applied as FF or FB scheme. The above simulation results suggest use of xGA or MB regardless of the input signal. However, the benefits of xGA and MB vanish for multi-level modulation constellations such as QAM or APSK. Figure 3.29 shows the degradation in case of 16-APSK used as modulation scheme and $\alpha = 0.2$. The performance of GA and OM stays almost the same regardless of

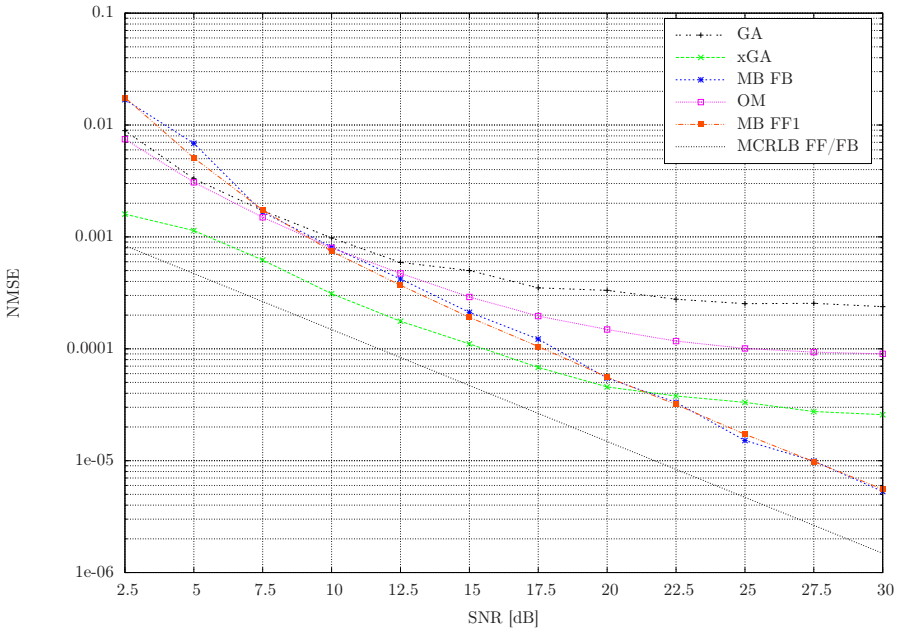


Figure 3.28: FF and FB NMSE of $\hat{\epsilon}$ for QPSK and $\alpha = 0.2$

constant or non-constant envelope of the modulation scheme. Unfortunately this does not hold true for xGA and MB. As can be seen the NMSE increases by nearly two orders of magnitude in the high SNR range.

Since at this stage of the blind demodulation process the modulation constellation is unknown, the optimum algorithm can not be selected. Due to its solid jitter performance for all types of input signals, the OM method is used solely for symbol timing recovery at this point. In later stages when the required parameters are available, a more appropriate algorithm may be selected. Therefore, the FF MB method proposed by the author can be an attractive alternative to established schemes in the medium-to-high SNR range for signals exhibiting constant envelope.

Acquisition Performance As already mentioned above, feedforward schemes exhibit a defined acquisition time in contrast to their feedback counterpart. Thus, if not using a FF scheme to initialize the subsequent FB method, an analysis of the acquisition behavior is inevitable. In [9] it is shown for detectors, exhibiting a sinusoidal shaped S-curve, that the acquisition trajectory for the noiseless case can simply be written as

$$\eta_k = \frac{1}{\pi} \arctan\Big(\tan(\pi\,\eta_0)\, e^{-4\,k\,B_L T}\Big) \tag{3.43}$$

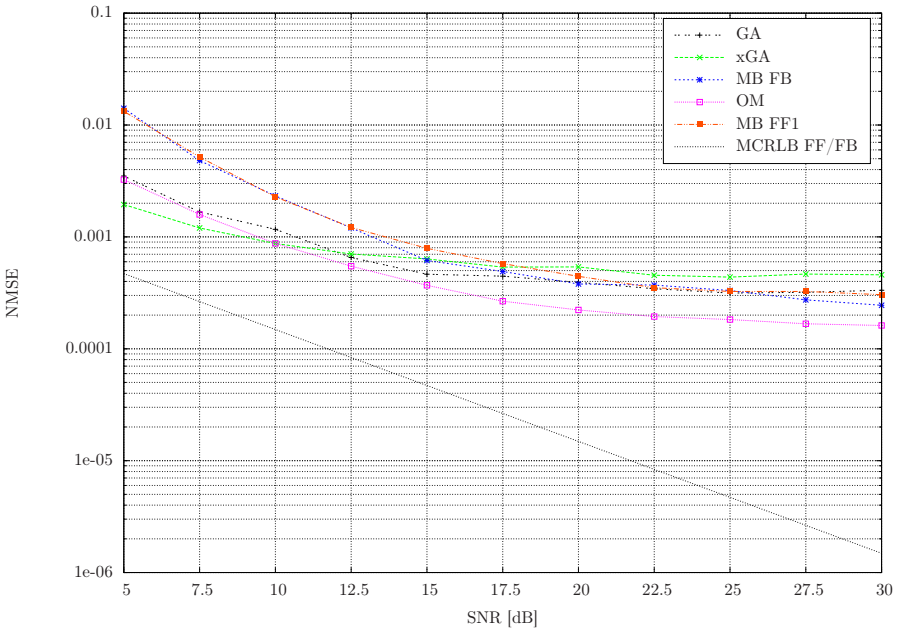


Figure 3.29: FF and FB NMSE of $\hat{\epsilon}$ for 16-APSK and $\alpha = 0.2$

Herein η_0 corresponds to the initial residual timing offset. It was verified that this trajectory does not only apply to GA, but to xGA and MB as well. Slow acquisition behavior, as one of the main drawbacks of feedback schemes, can be mitigated by using adaptive loop filters as investigated in [9] and in more detail in [35] for GA. It is obvious that the decay of the above formula is mainly determined by the normalized equivalent noise bandwidth $B_L T$. The requirements on the selection of $B_L T$ are in striking contrast. On the one hand, large values favor a fast acquisition process, but, on the other hand, small values are better suited to achieve a good steady-state jitter performance. The idea is to vary the bandwidth over time to guarantee a fast acquisition and in the following a good tracking performance. A simple form how to change $B_L T$ over time could be an exponential function $f(t)$ defined as

$$f(t) = 1 + b\, e^{-c\, t} \Leftrightarrow F(t) = t + \frac{b}{c}\,(1 - e^{-c\, t}) \tag{3.44}$$

Herein $F(t)$ corresponds to the integral of $f(t)$, which has to be evaluated due to the time dependency of $f(t)$ in the derivation of the acquisition trajectory. The complete derivation is detailed in [35]. The parameter b defines the starting value of the bandwidth and c the speed of the decay. So finally the adaptive trajectory can be stated as

$$\eta_k = \frac{1}{\pi} \arctan\Big(\tan(\pi\, \eta_0)\, e^{-4\, B_L\, F(k\, T)}\Big) \tag{3.45}$$

In the following, the theoretical acquisition trajectories are verified by simulation runs. 1000 symbols are used to observe the lock-in process. The normalized equivalent noise bandwidth is set to $B_L T = 0.001$ and the adaptive parameters to $b = 9$ and $c = 1$. The initial timing error is set to $\epsilon = \eta_0 = 0.35$ and QPSK is the used modulation scheme. In Figure 3.30 the acquisition process of GA, xGA and MB can be seen for a roll-off $\alpha = 0.35$ and in the absence of noise. From the result the conclusion can be drawn that

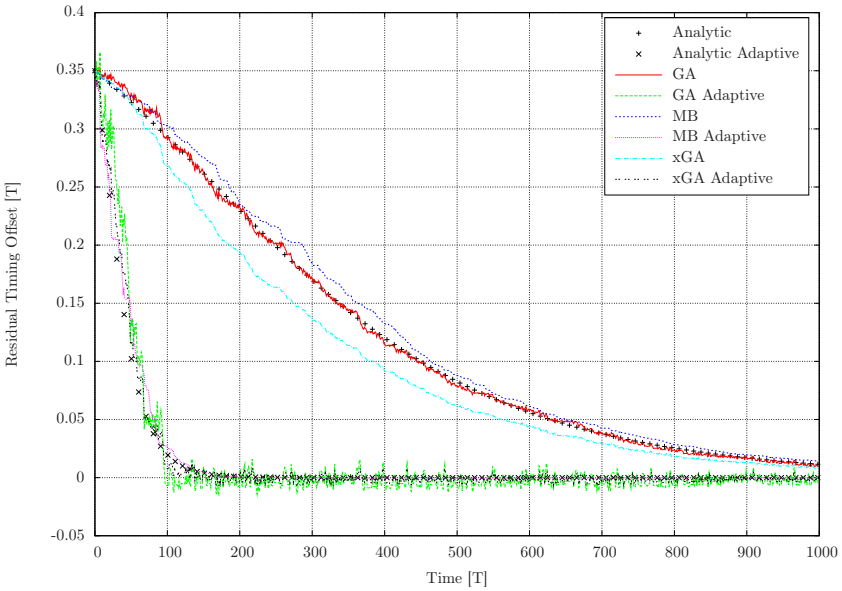


Figure 3.30: Acquisition performance of GA, xGA and MB for $\alpha = 0.2$ without noise

the analytic trajectories coincide very well with the simulated lock-in process. As can be seen, by usage of the mentioned adaptive filter function, the speed up can make the FB schemes an attractive alternative for FF methods when rapid acquisition is required. For the simulation result in Figure 3.31, the roll-off was reduced to $\alpha = 0.2$ and the SNR was set to $\gamma_s = 10\,\mathrm{dB}$. It can be seen that the fluctuations are more pronounced compared to the noiseless case. Especially, due to the reduction of the excess bandwidth, this effect can be observed for the GA most explicitly. Nevertheless, the simulated results follow the theoretical curves very well.

Parameter Mismatch The OM and the first MB FF method are the only timing recovery schemes that do not require any additional knowledge of the input signal. GA, xGA and the remaining MBs may require the roll-off parameter α or even the type of modulation. The question of the impact arises, when designing the recovery scheme including the MF using non-matching parameters.

In Figure 3.32 the evolution of the NMSE is shown for the GA. At first, it is assumed for

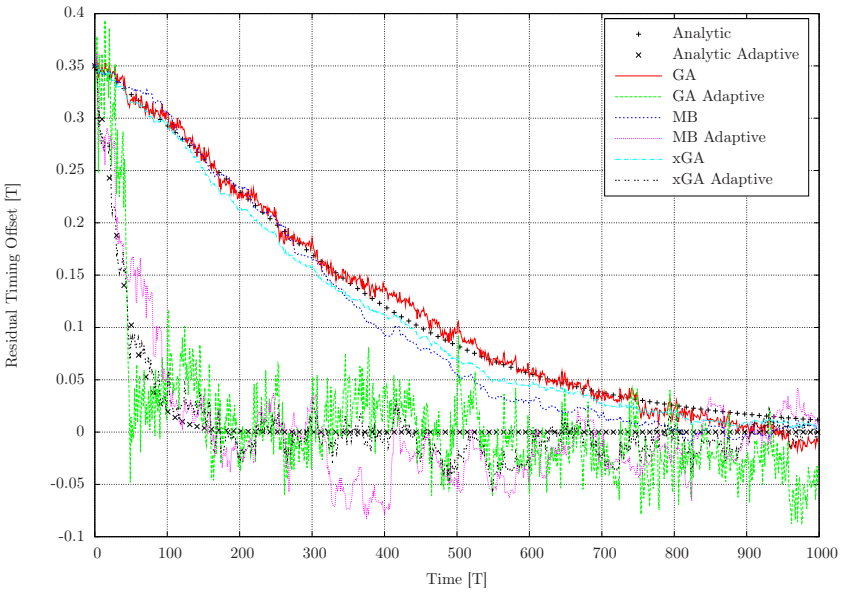


Figure 3.31: Acquisition performance of GA, xGA and MB for $\alpha = 0.2$ and $\gamma_s = 10\,\mathrm{dB}$

$\alpha = 0.2$ and $\alpha = 0.35$ that the value for designing the MF and the tracking loop coincides with the true roll-off. In the following, a wrong value is used for the design. One time the true value is $\alpha = 0.2$ and $\alpha = 0.35$ is used for the design. In the second case it is just the other way round, i.e. $\alpha = 0.2$ for the design and $\alpha = 0.35$ is the true value. In the first case the jitter performance is superior to the true scenario, in contrast to the second case, where the performance degrades. The sensitivity to mismatch of the roll-off for design of MB FB scheme is depicted in Figure 3.33. It gets obvious that the MB FB is very robust in the presence of roll-off mismatch. The sensitivity of the xGA to the mismatch is in between the one of GA and MB FB. The second parameter required for setting up timing recovery properly is the type of modulation. The sensitivity of the algorithms to mismatch of the modulation scheme is conversely to the impact of roll-off deviation. In this case the MB suffers from the highest degradation, especially when mismatching schemes with constant and non-constant envelope, as can be seen in Figure 3.34.

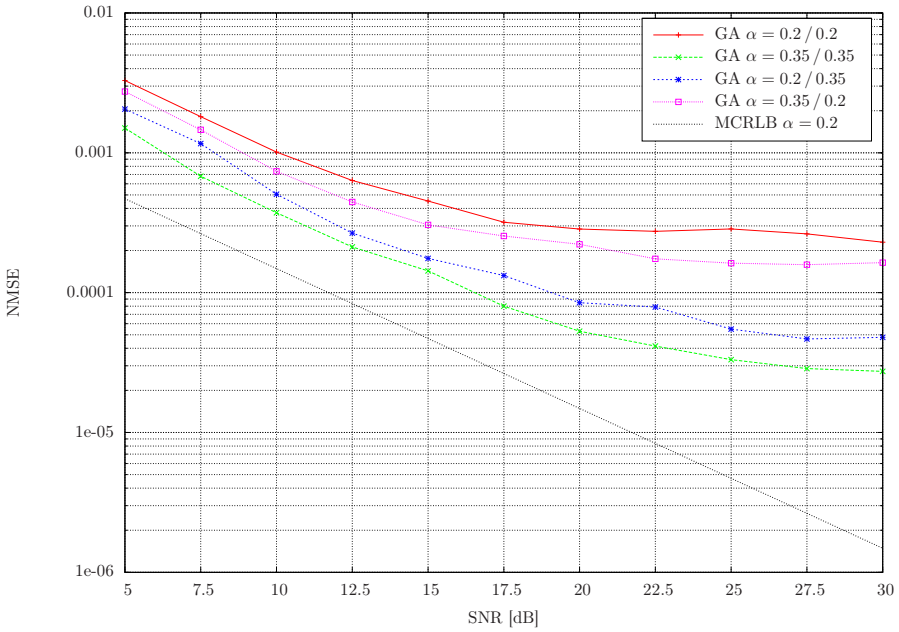


Figure 3.32: GA NMSE of $\hat{\epsilon}$ for mismatch of α

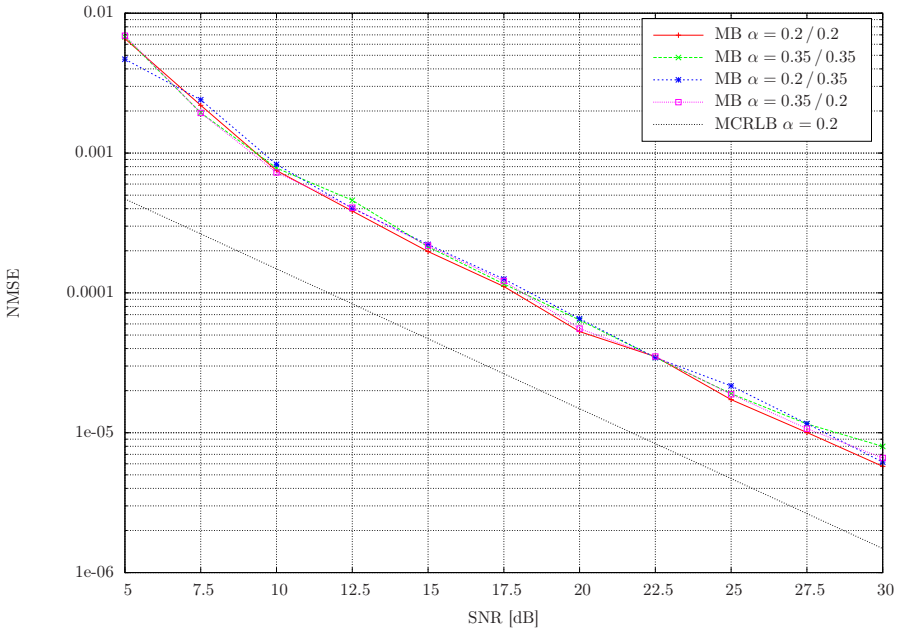


Figure 3.33: MB FB NMSE of $\hat{\epsilon}$ for mismatch of α

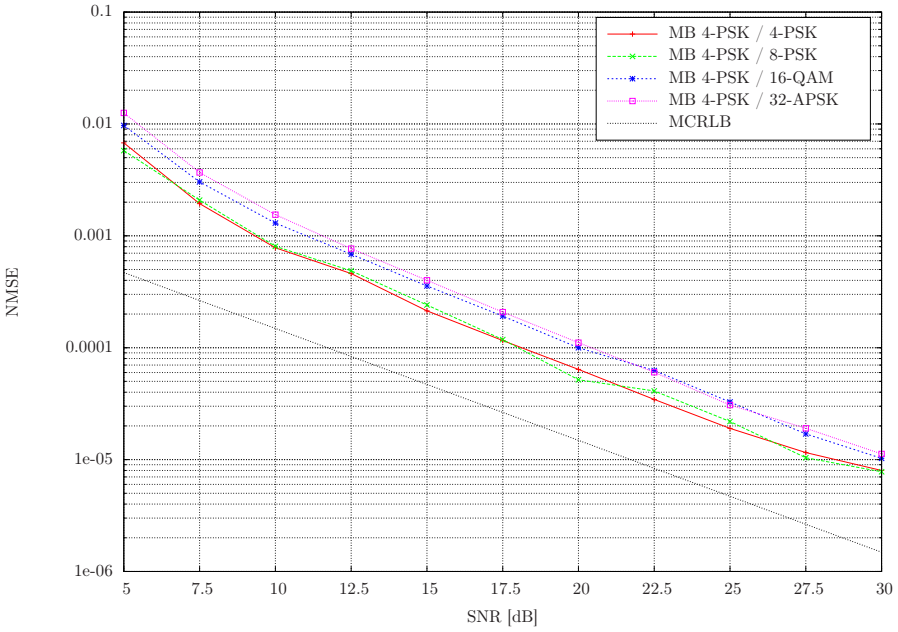


Figure 3.34: MB FB NMSE of $\hat{\epsilon}$ for mismatch of the modulation scheme

3.9 Modulation Classification

The attempt to determine the modulation type of a received signal inside a noisy observation is termed modulation classification. This process can be of cooperative or non-cooperative nature. An example for the first category is the selection of an appropriate demodulator in an SDR environment, whereas monitoring in commercial communication systems or signal interception, as it arises in military communications, are prominent examples for a non-cooperative environment. An overview of existing modulation classification schemes can be found in [39] and, in more detail, in [40]. Two types of classifiers are distinguished: likelihood- and feature-based methods. The main drawback of the likelihood-based approach is both the computational effort and the sensitivity to impairments like phase and frequency offset. In contrast, feature-based methods are suboptimal, but they are more robust and simple.

Many of the existing algorithms rely on the assumption that modulation parameters such as phase, frequency and timing offset are known in advance. Within this work a feature-based modulation classifier was developed that is not limited by this restriction. The only reasonable assumption is the knowledge of the timing offset; as shown in the previous Section 3.8, a modulation-independent method, like the OM algorithm [25], can be used for this task. But the samples at the MF output are still impaired by carrier frequency and phase errors. The established estimation algorithms for determining the latter require knowledge of the used modulation type. In [41] the attempt was made to determine the frequency offset regardless of the used signal constellation. However, so far no modulation-independent algorithms have been found with sufficient accuracy. This leaves us with the requirement to use a modulation classifier that is able to operate properly in presence of frequency and phase offset. The new algorithm, developed by the author [14], is based on the feature-based method in [42] and especially designed to work with modulation schemes typical for satellite communications. It consists of two modulation classifiers in parallel and is extended by an additional SNR restriction feature.

3.9.1 Features

Normalized fourth- and sixth-order cumulants are used as features for the inter- and the intra-modulation class discrimination to derive a hierarchical scheme similar to the one described in [42]. The choice of the features is based on the fact that higher order cumulants (> 2) of Gaussian noise are equal to zero [43]. Thus, these characteristics are very robust in the presence of additive Gaussian noise. Equations for the calculation of cumulants from moments can be found in [43] and [44], where the moments of a signal x_n are defined as

$$M_{pq} = E[x_n^{p-q}(x_n^*)^q] \tag{3.46}$$

Thus, the theoretical values of the features can be pre-computed in the receiver by the calculation of the ensemble averages over the ideal noise-free signal constellation. To arrive at non-ambiguous values, it is assumed that the ideal constellation has unit variance and

all symbols are equiprobable. The determination of the considered cumulants can be performed by

$$C_{20} = M_{20} \tag{3.47}$$

$$C_{21} = M_{21} \tag{3.48}$$

$$C_{40} = M_{40} - 3M_{20}^2 \tag{3.49}$$

$$C_{42} = M_{42} - |M_{20}|^2 - 2M_{21}^2 \tag{3.50}$$

$$C_{63} = M_{63} - 6M_{20}M_{41} - 9M_{21}M_{42} + 18M_{20}^2M_{21} + 12M_{21}^3 \tag{3.51}$$

Since phase rotation has an impact on cumulants which exhibit asymmetry of conjugation, the absolute values of the cumulants are used as features. In Table 3.1 the features of x_n are listed for the considered modulation schemes. In the presence of frequency offset,

Table 3.1: Cumulants of x_n ($w_n = 0$, $E_s = 1$)

Modulation type	$\|C_{20}\|$	$\|C_{40}\|$	$\|C_{42}\|$	$\|C_{63}\|$
2-PSK	1.00	2.00	2.00	16.00
4-PSK	0.00	1.00	1.00	4.00
8-PSK	0.00	0.00	1.00	4.00
16-QAM	0.00	0.68	0.68	2.08
64-QAM	0.00	0.62	0.62	1.80
16-APSK	0.00	0.00	0.77	2.49
32-APSK	0.00	0.00	0.61	1.65

the usage of these features would lead to significant performance degradation; thus, as suggested in [42], the differential signal $z_n = x_{n-1}^* x_n$ could be employed to convert the frequency offset to a fixed phase offset. This transformation was used for the suggested classification schemes in [45] and [46] as well. Table 3.2 summarizes the features of z_n.

3.9.2 Algorithm

From Table 3.1 and 3.2 it becomes obvious that the listed features are distinct for the observed modulation schemes. Thus, a hierarchical classification scheme similar to the one introduced in [42] can be derived. Applying the simplification that the variance of a cumulant is the same for all modulation schemes under consideration reduces the discrimination between two signal constellations to a threshold detector. The required thresholds for discrimination between two different modulation (sub-)classes m and n can be derived from $|C_{uv}^m|$ and $|C_{uv}^n|$. An observation period of L symbols at the output of the

Table 3.2: Cumulants of z_n ($w_n = 0$, $E_s = 1$)

Modulation type	$\lvert C_{20}\rvert$	$\lvert C_{40}\rvert$	$\lvert C_{42}\rvert$	$\lvert C_{63}\rvert$
2-PSK	1.00	2.00	2.00	16.00
4-PSK	0.00	1.00	1.00	4.00
8-PSK	0.00	0.00	1.00	4.00
16-QAM	0.00	0.46	0.26	0.16
64-QAM	0.00	0.38	0.09	0.21
16-APSK	0.00	0.00	0.49	0.82
32-APSK	0.00	0.00	0.08	0.79

previous timing recovery stage is used to calculate a feature estimate $|\hat{C}_{uv}|$. This leads to the formulation of a decision step between two modulation (sub-)classes with $|C_{uv}^m| < |C_{uv}^n|$ as

$$|\hat{C}_{uv}| < \frac{|C_{uv}^m| + |C_{uv}^n|}{2} \quad \begin{cases} \text{yes}, & \text{choose } m \\ \text{no}, & \text{choose } n \end{cases} \tag{3.52}$$

After analyzing the features in Table 3.1, a hierarchical classification scheme based on threshold detection can be developed. For x_n as input signal, at first $|C_{42}|$ is used to discriminate between 2-PSK, PSK ($M > 2$) and all other modulation classes. Then, $|C_{40}|$ chooses among 4-PSK, 8-PSK, 16-QAM, 64-QAM and APSK. Finally, $|C_{42}|$ is taken to decide between 16-APSK and 32-APSK. In a similar manner an algorithm can be developed for the differential signal z_n. The classification steps are formulated as a binary decision tree, illustrated in Figure 3.35 for x_n and in Figure 3.36 for z_n as input signal. To avoid pitfalls, it should be mentioned that appropriate scaling of the input signal, as detailed in Section 3.7, is necessary for reasonable comparison of the estimated features $|\hat{C}_{uv}|$ to the theoretical values $|C_{uv}|$. Otherwise, the estimated features have to be scaled accordingly. Since the SNR estimators available from the open literature, like moment-based methods [47], are not applicable due to the lack of knowledge of the modulation type, the pre-estimate for γ_s from the former stages is taken. In [48] a scheme is presented, which performs joint power estimation and modulation classification to circumvent this dilemma. If no SNR estimate is available, $\hat{C}_{21}$ could be used as a rough solution; however the performance of the algorithm will degrade with increasing noise.

With decreasing SNR, the number of wrong classifications raises more rapidly for the differential scheme as for the standard scheme. This is due to the increased variance of the differential features. It turned out that the amount of erroneous decisions between 8-PSK and APSK based on $|C_{63}|$ of the differential scheme can be decreased significantly

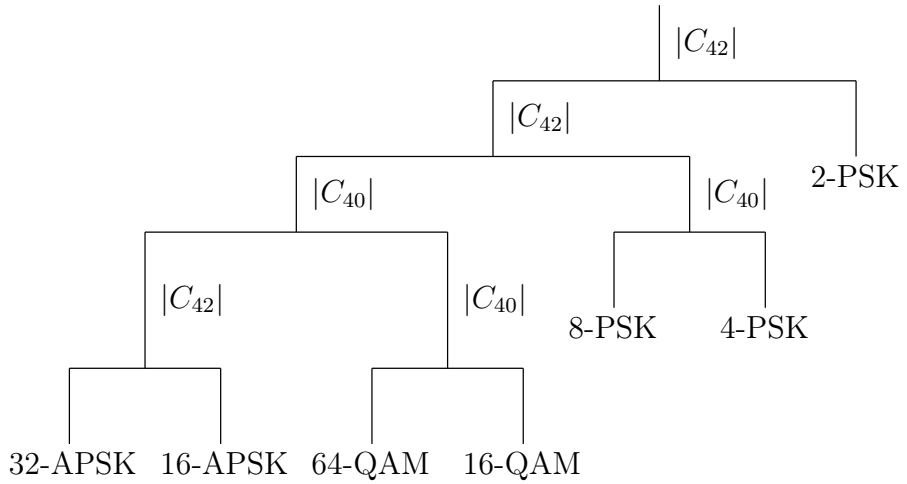


Figure 3.35: Binary decision tree (standard modulation classifier)

by applying the decision of the standard method. Observing the confusion matrices of the differential and normal scheme, it is obvious that the normal scheme is very reliable for distinguishing between PSK and APSK even in presence of a frequency offset. Table 3.3 shows the confusion matrix of the normal scheme at an SNR of $\gamma_s = 5\,\mathrm{dB}$, an estimator length of $L = 10000$ symbols and 1000 performed iterations for the DVB-S2 scenario. Thus, if the differential approach decides for APSK or PSK, the decision is cross-checked

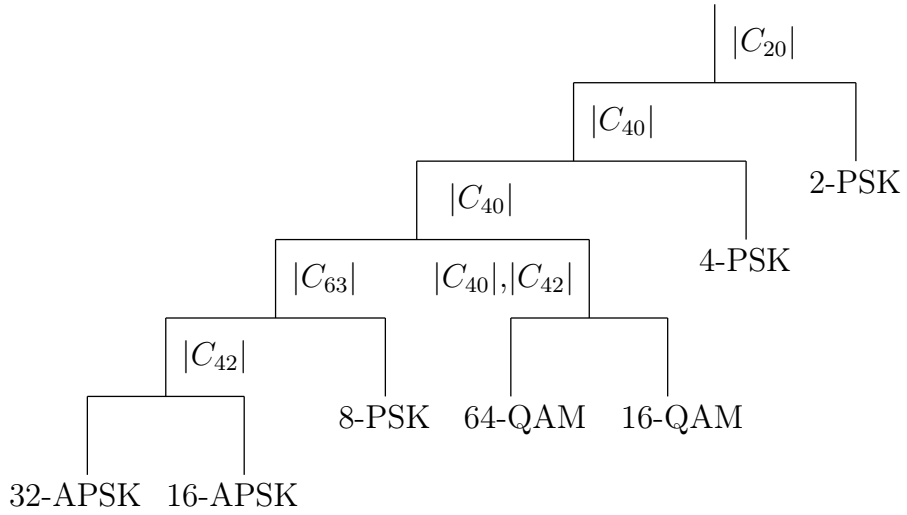


Figure 3.36: Binary decision tree (differential modulation classifier)

Table 3.3: Confusion matrix of normal scheme for DVB-S2 scenario

	Estimated			
True	4-PSK	8-PSK	16-APSK	32-APSK
4-PSK	0	1.0	0	0
8-PSK	0	1.0	0	0
16-APSK	0	0	1.0	0
32-APSK	0	0	0.001	0.999

with a normal scheme running in parallel. As mentioned in [42], the most challenging discrimination is between 16-QAM and 64-QAM, which requires a large amount of data. It was observed that using $|C_{42}|$ for this purpose at high SNR and $|C_{40}|$ at low SNR outperforms a single feature over the whole SNR range. No similar SNR-dependent selection of features was found in the open literature. Simulation results demonstrating the performance gain are omitted due to the limited space.

3.9.3 SNR Restriction

To increase the probability of successful classification, an additional SNR restriction feature is applied. A reliable SNR estimate is available that can be used to narrow the range of possible signal constellations. Depending on the order of the modulation scheme, a certain SNR is required to guarantee reasonable transmission of data. So, by inspection of the Shannon plane [33], a threshold can be derived for each modulation scheme below which error-free transmission is impossible. These SNR limits can be incorporated into the proposed hierarchical classification scheme. The relationship for the capacity boundary, i.e. the bit rate R equals the channel capacity C, can be stated as follows:

$$\frac{E_b}{N_0} = \frac{1}{R}(2^R - 1) \tag{3.53}$$

and is illustrated in Figure 3.37. The solid dots correspond to the E_b/N_0 limits for modulation orders $\nu = \log_2 M = 1,2,\ldots,6$ in the uncoded case. For the coded case $R = r \log_2 M$, r denoting the code rate. Thus, an SNR limit can be derived on

$$\frac{E_s}{N_0} = \frac{E_b}{N_0} r \log_2 M \tag{3.54}$$

for each order of modulation, as summarized in Table 3.4 for the uncoded ($r = 1$) and ideally coded ($0 < r < 1$) case.

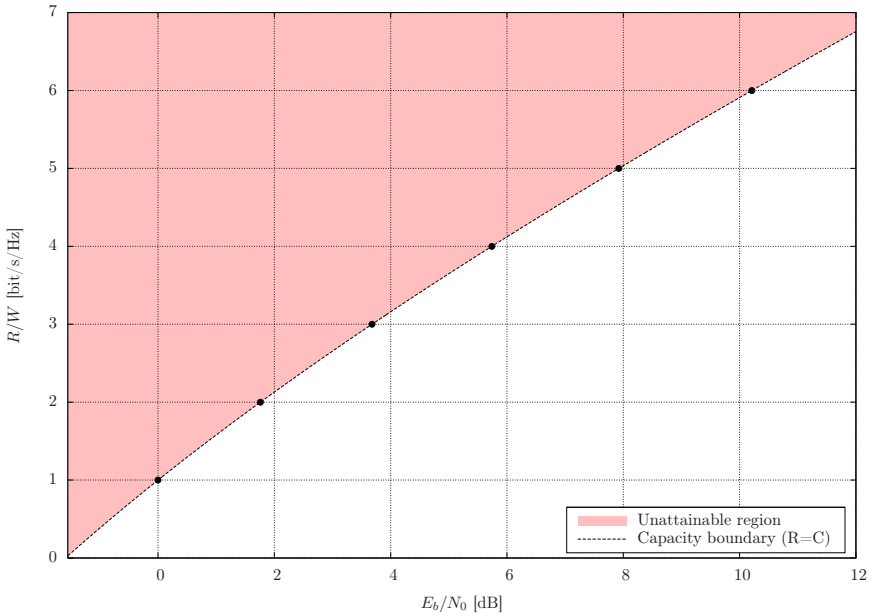


Figure 3.37: Shannon plane showing spectral efficiency

3.9.4 Simulation Results

Numerous simulation runs were performed to evaluate the performance of the proposed algorithm. For each modulation scheme, 1000 simulations were performed. Moreover, if not stated otherwise, it is assumed that the true SNR, required for scaling and SNR restriction, is known. For calculation of the SNR limits the code rate is assumed to be

Table 3.4: Derived SNR limits for the uncoded and ideally coded ($r = 1/2$) case

Modulation order	uncoded [dB]	coded [dB]
1	0.00	-3.83
2	4.77	0.00
3	8.45	2.62
4	11.76	4.77
5	14.91	6.68
6	17.99	8.45

$r = 1/2$. The number of available symbols for classification is set to $L = 1000$, 5000 and 10000. For each iteration, both frequency offset $|\Delta fT| \leq 0.1$ and phase offset $-\pi < \theta \leq \pi$ are selected randomly. To be able to utilize the SNR restriction, modulation schemes are only considered, if the pre-calculated SNR limits allow error-free transmission. Additional simulation results can be found in the paper by the author [14].

Scenario DVB-S2 The first scenario comprises the set of modulation constellations S_1 as defined in the DVB-S2 standard [13], i.e. S_1={4-PSK, 8-PSK, 16-APSK, 32-APSK}. For this purpose, the binary decision tree is modified to avoid choosing modulation types which are not included in the problem set. The success rate P_s, i.e. the number of successful classifications divided by the number of iterations, is depicted in Figure 3.38. The reason for the degradation in the medium-to-low SNR range is mainly due to wrong

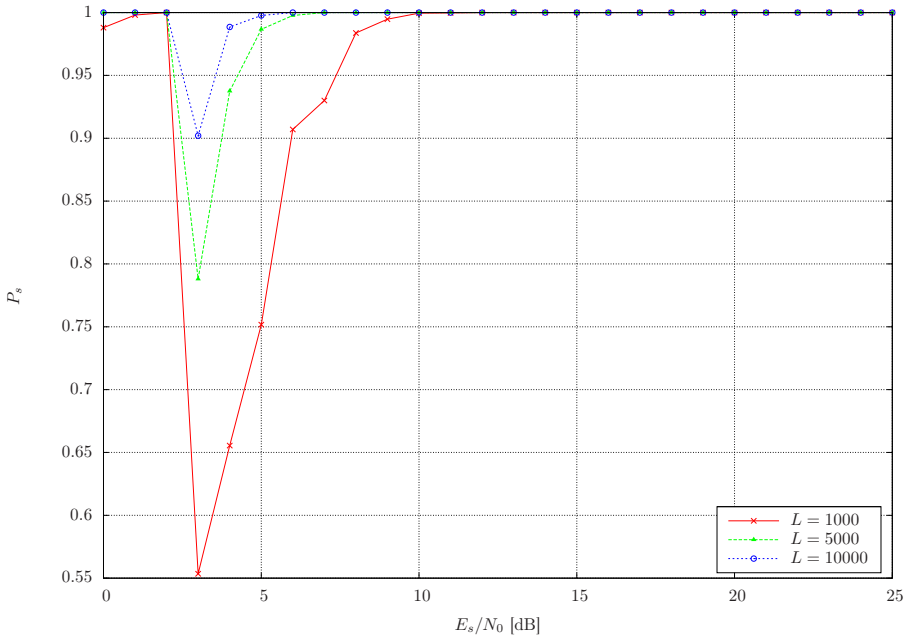


Figure 3.38: Modulation classifier success rate (DVB-S2)

decisions between 4-PSK and 8-PSK. Since below the SNR limit of 8-PSK only 4-PSK remains as valid modulation type, the success rate increases to 1 again.

To address the above mentioned problem of scaling the calculated features appropriately, a separate simulation run was performed. The estimator length is assumed to be $L = 1000$ symbols. The performance degradation in the absence of an SNR estimate can be observed in Figure 3.39. In this case $\hat{C}_{21}$ is used for scaling of the features, instead of the signal power. The impact reduces with decreasing noise and vanishes for SNR values $\gamma_s \geq 15\,\text{dB}$. For SNR values $\gamma_s < 7\,\text{dB}$, it happens that P_s drops down to 0. This is due to the fact

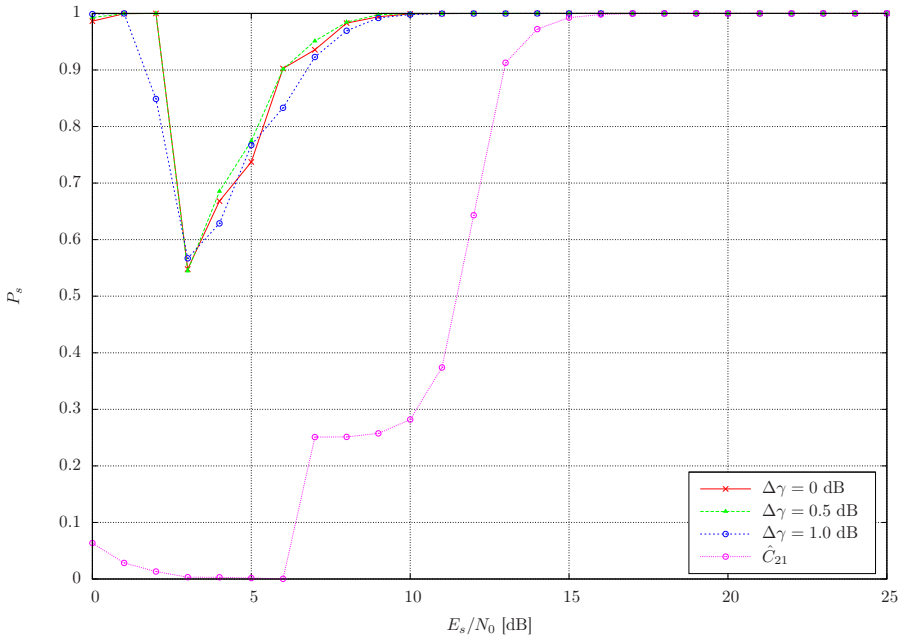


Figure 3.39: Modulation classifier success rate (DVB-S2, SNR mismatch)

that the estimator decides for 32-APSK although this modulation scheme is not used for $\gamma_s < 6.68\,\text{dB}$ (see Table 3.4). Since no SNR estimate is available, the SNR restriction feature is disabled and can not counteract this problem. Moreover, the effect of small errors $\Delta\gamma$ on the SNR estimate is shown, i.e. $\hat{\gamma}_s\,[\text{dB}] = \gamma_s\,[\text{dB}] + \Delta\gamma\,[\text{dB}]$. For $\Delta\gamma = 0\,\text{dB}$ the SNR estimate coincides with the true SNR. It can be seen that the impact of small offsets is nearly negligible. Thus the pre-estimate $\hat{\gamma}_s^*$ should be sufficient for scaling purposes.

Scenario Supported Modulation Schemes The second problem set S_2 consists of all supported modulation schemes, i.e. S_2={2-PSK, 4-PSK, 8-PSK, 16-QAM, 64-QAM, 16-APSK, 32-APSK}. The evolution of the success rate P_s for S_2 is depicted in Figure 3.40. The degradation in the medium SNR range stems mainly from wrong decisions of 16-QAM and 64-QAM, as can be seen from the confusion matrix in Table 3.5. The abbreviation M_1 to M_7 is used for 2-PSK, 4-PSK, 8-PSK, 16-QAM, 64-QAM, 16-APSK and 32-APSK due to the limited space. It can be seen that this degradation can be mitigated by increasing the estimator length, as shown for $L = 50000$.

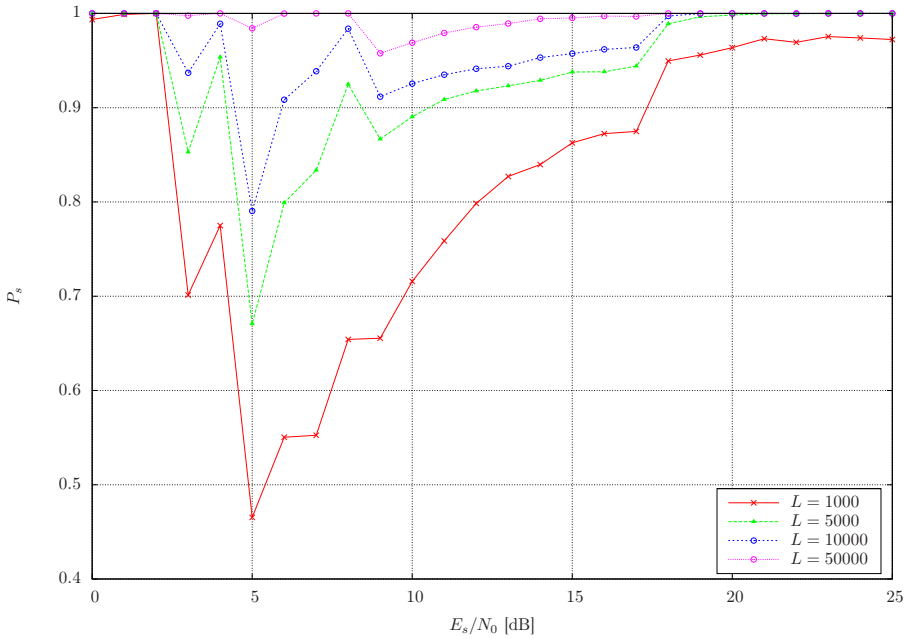


Figure 3.40: Modulation classifier success rate (supported modulation schemes)

Table 3.5: Confusion matrix of differential scheme

	Estimated						
True	M1	M2	M3	M4	M5	M6	M7
M1	1.0	0	0	0	0	0	0
M2	0	1.0	0	0	0	0	0
M3	0	0	1.0	0	0	0	0
M4	0	0	0	0.834	0.166	0	0
M5	0	0	0	0.354	0.646	0	0
M6	0	0	0	0	0	1.0	0
M7	0	0	0	0	0.001	0	0.999

3.10 SNR Estimation

From the pre-estimation stage there is already a measure for the SNR with sufficient accuracy available. So the question arises, why to estimate the SNR again? There are three reasons for this additional effort: (i) the accuracy of the refined SNR estimate may be superior to the pre-estimate; (ii) a failure of the symbol timing recovery stage might be detected by a significant drop in the estimated SNR; (iii) powerful SNR estimation can be used to estimate the roll-off factor in alternative manner, since the SNR at the MF output is a maximum for a properly designed receiving filter.

From the different available approaches for SNR estimation [47], it turns out that moment-based estimators tend to fit best for the task to solve. For example, maximum likelihood (ML) methods seem to be too complex from the computational point of view. The drawback that moment-based estimators are suboptimal is outweighed by the advantages that they are NDA, feature simple implementation and do not require carrier recovery. The standard M_2M_4 estimator [47] exhibits good performance for signal constellations with constant envelope such as PSK; however, it deteriorates completely for multi-level schemes, e.g. QAM and APSK, in the medium-to-high SNR range. An attempt to extend the usable SNR range by applying additionally the sixth-order moment M_6 is made in [49]. Using partitioned subsets of the original signal constellation with equal magnitude is presented in [50]. This approach exhibits good performance in the medium-to-high SNR range, but it degrades for small SNR values due to wrong assignment of the symbols to the subsets. Finally, in [51] an estimator using also the eighth-order moment M_8 is presented to achieve satisfactory performance for multi-level constellations. Moreover, the algorithm allows tuning for specific SNR values. Thus, a combined estimator is furnished, consisting of the standard M_2M_4 estimator for the low SNR range, the partitioned method in the medium-to-high SNR range and the M_8 estimator for the overlapping area when performance is insufficient. The SNR pre-estimate is used to choose the appropriate algorithm.

The above mentioned estimators require correction of the symbol timing offset as well as the removal of oversampling such that the symbols, impaired by frequency and phase offset, are applied as input. A residual timing error will lead to catastrophic behavior of the estimator. An alternative algorithm which does not rely on timing recovery is presented in [34]. The main drawback of this method is the limitation to the medium SNR range for reasons of sufficient accuracy.

3.10.1 Moment-Based Estimator

An observation length of L symbols is used to estimate the second- and fourth-order moment M_2 and M_4:

$$M_2 = E\left[x_i\, x_i^*\right] \approx \hat{M}_2 = \frac{1}{L}\sum_{i=1}^{L} |x_i|^2 \tag{3.55}$$

$$M_4 = E\left[(x_i\, x_i^*)^2\right] \approx \hat{M}_4 = \frac{1}{L}\sum_{i=1}^{L} |x_i|^4 \tag{3.56}$$

Since signal and noise are zero-mean and independent, with independent real and imaginary part of the latter, the above equations can be written as

$$M_2 = E_s + E_n \tag{3.57}$$

$$M_4 = K_4 E_s^2 + 4E_s E_n + 2E_n^2 \tag{3.58}$$

where K_4 corresponds to the symbol kurtosis $K_4 = E_c[|c_k|^4]$. Solving the above equations for E_s and E_n and forming the result to the ratio $\gamma_s = E_s/E_n$ delivers the estimator

$$\hat{\gamma}_s = \frac{\sqrt{2\hat{M}_2^2 - \hat{M}_4}}{\hat{M}_2 - \sqrt{2\hat{M}_2^2 - \hat{M}_4}} \tag{3.59}$$

The normalized CRLB, as a lower limit for the accuracy of this estimator, can be stated as [50]

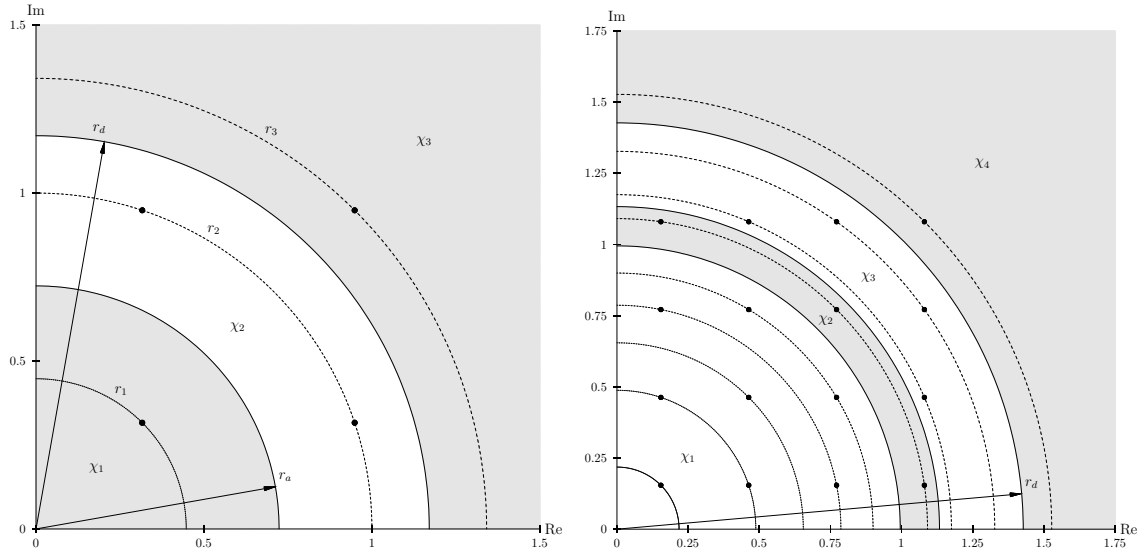
$$\text{NCRLB} = \frac{\text{CRLB}}{\gamma_s^2} = \frac{1}{L}\left(1 + \frac{2}{\gamma_s}\right) \tag{3.60}$$

3.10.2 Partitioned Moment-Based Estimator

With increasing SNR the estimate is dominated by self noise for constellations with varying symbol amplitude levels. So, as presented in [50], the idea is to partition the symbols in areas with constant magnitude. For the estimation process, only symbols of one region are considered for calculation of required moments. When keeping the number of input symbols constant this will lead to a reduction of usable symbols according to the selected subset. Since only symbols located on a circle with radius r_i are processed, the assumption of a constellation with unit variance does not hold anymore. This has to be taken into account in the final estimator formula given by

$$\hat{\gamma}_s = \frac{\sqrt{2\hat{M}_2^2 - \hat{M}_4}}{r_i^2\left(\hat{M}_2 - \sqrt{2\hat{M}_2^2 - \hat{M}_4}\right)} \tag{3.61}$$

16-QAM In Figure 3.41 the first quadrant of the supported QAM constellations is illustrated. In case of 16-QAM, as shown in Figure 3.41a, it can be seen that this modulation type exhibits three levels of different amplitudes, i.e. r_1, r_2 and r_3. In all four quadrants, four symbols are located on both inner and outer circle with radii r_1 and r_3; the remaining eight symbols are on the middle circle with radius r_2. Thus, three regions for partitioning can be defined, i.e. χ_1, χ_2 and χ_3. The decision borders for the regions are assigned by radii r_a and r_d. An analysis on the selection of r_a and r_d depending on the



(a) 16-QAM **(b)** 64-QAM

Figure 3.41: Partitioning of QAM constellations for SNR estimation

SNR is given in [50] as well. A convenient choice and also the solution for the asymptotic case $\gamma_s \to \infty$ is

$$r_a = (r_1 + r_2)/2 \tag{3.62}$$

$$r_d = (r_2 + r_3)/2 \tag{3.63}$$

Again, it should be mentioned that scaling of the received signal is necessary for correct partitioning as described in Section 3.7. Now the question arises, which of the three defined regions should be used for SNR estimation? In [50] the usage of region χ_2 is suggested since it exhibits the lowest NMSE in the high SNR range. It is clear that using χ_2 for large SNR values involves twice the number of symbols compared to the two remaining regions. In Figure 3.42 the mean estimator output and in Figure 3.43 the evolution of the NMSE are illustrated. The estimator length is set to $L = 1024$ symbols. Thus, in the noiseless case and assuming equiprobable symbols, applying partitioning of regions χ_1 or χ_3 results in 256 symbols used for estimation. In case of using χ_2, the number is twice as large on average, i.e. 512 symbols. By inspection of the simulation results it becomes clear that for SNR values $\gamma_s \leq 20\,\text{dB}$ using χ_3 leads to best performance, although increasingly biased with decreasing SNR. However, above this threshold using χ_2 is superior due to unambiguous assignment of the symbols. For comparison reasons, the standard algorithm without partitioning is displayed as well. In the low SNR range, i.e. $\gamma_s \leq 10\,\text{dB}$, this estimator works satisfactorily; however, with decreasing noise, the accuracy degrades more and more.

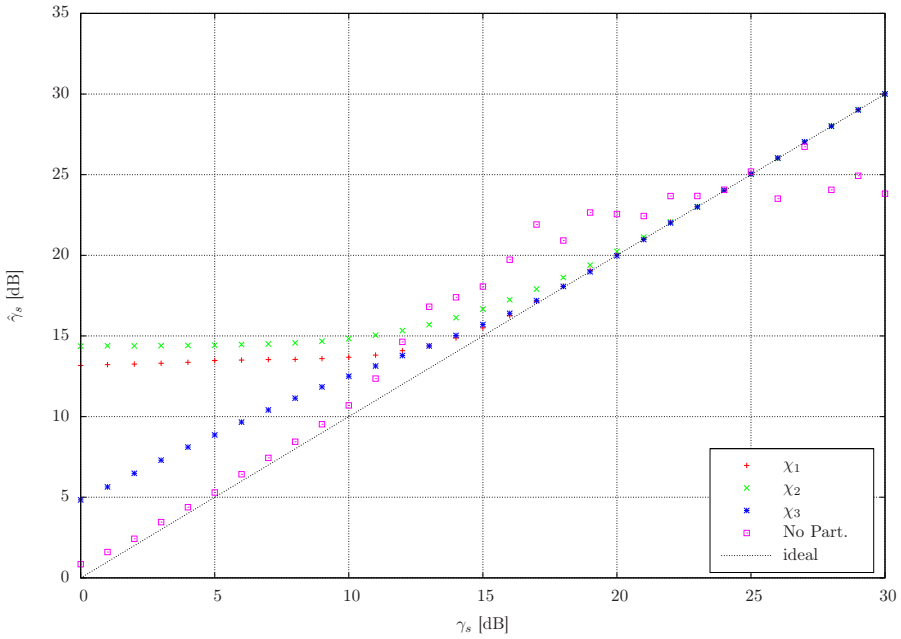


Figure 3.42: Mean SNR estimator output for different partitions (16-QAM)

64-QAM From Figure 3.41b it can be seen that for 64-QAM the magnitudes of the symbols are aligned on nine different levels. The evolution of the estimated mean and NMSE exhibits similar effects as for 16-QAM. To achieve best performance in the high SNR range the symbols from region χ_2 should be taken for estimation. Below a certain SNR value, depending on the estimator length L, the usage of partition χ_4 is recommended. Since only four symbols are located in region χ_4, the estimator length has to be increased appropriately to achieve satisfying results. In Figure 3.44 and Figure 3.45 the performance plots for 64-QAM are depicted. From both illustrations the poor performance due to the reduced estimator length can be seen. Also, the algorithm without partitioning exhibits insufficient accuracy in the low-to-medium SNR range. But, when increasing the estimator length from $L = 1024$ to $L = 8192$ the performance gap can be nearly closed.

16-APSK As illustrated in Figure 3.46a, 16-APSK exhibits two different amplitude levels. From the above results for QAM, it is clear that the recommended partition region χ_2 should contain the symbols on the outer circle. Thus, in average only one fourth of the received symbols has to be omitted from the calculation.

32-APSK Figure 3.46b shows the signal constellation of 32-APSK. The symbols are distributed over three circles. On the outermost circle, 16 symbols are located. So, it should

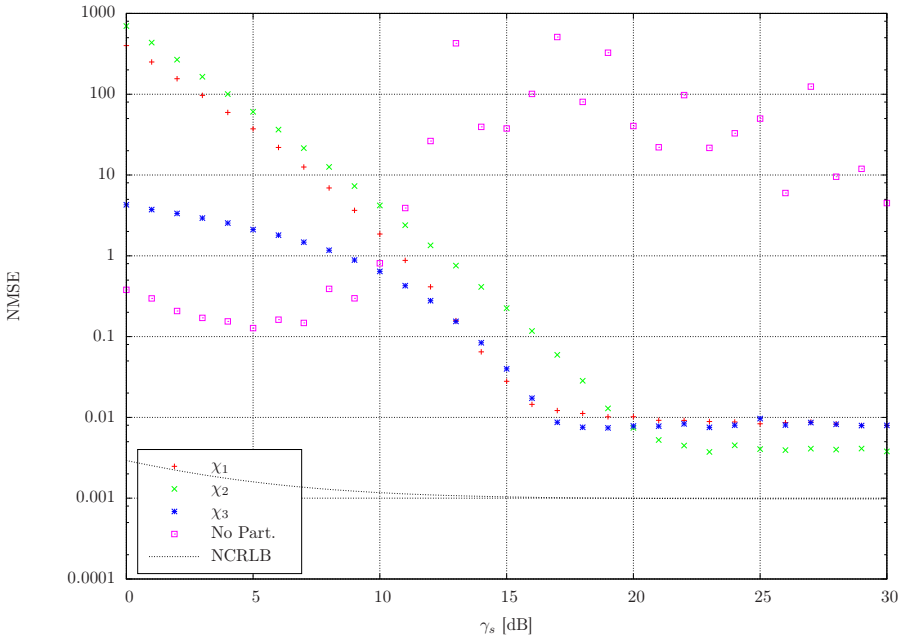


Figure 3.43: Evolution of NMSE for $\hat{\gamma}_s$ at different partitions (16-QAM)

be clear to take symbols with magnitudes larger than r_d for estimation. When choosing χ_2 for partitioning, one half of the symbols will be discarded on average from the calculation process. For performance comparison, the estimator length is set to $L = 1024$ symbols. In Figure 3.47 and Figure 3.48 the mean estimator output and the NMSE evolution are illustrated for 16-APSK and 32-APSK. Again, the estimator omitting partitioning is considered as well. Like for QAM, the latter performs satisfactorily below a certain SNR value; however, above this value the accuracy gets increasingly worse. The partitioned counterpart behaves the other way round.

3.10.3 Multi-level Moment-Based Estimator

From the performance analysis of the above presented algorithms, the idea arises to use the standard estimator in the low SNR range and the partitioned method in the high SNR range. However, in the transition section, where both algorithms are suboptimal, the performance is not satisfactory. A scheme which has adequate performance in this transition zone would be favorable, especially for QAM constellations. The M_8 algorithm presented in [51] seems to be appropriate, since it allows tuning of parameters for specific SNR values. In short, the estimator relies on the fact that a weighted linear combination

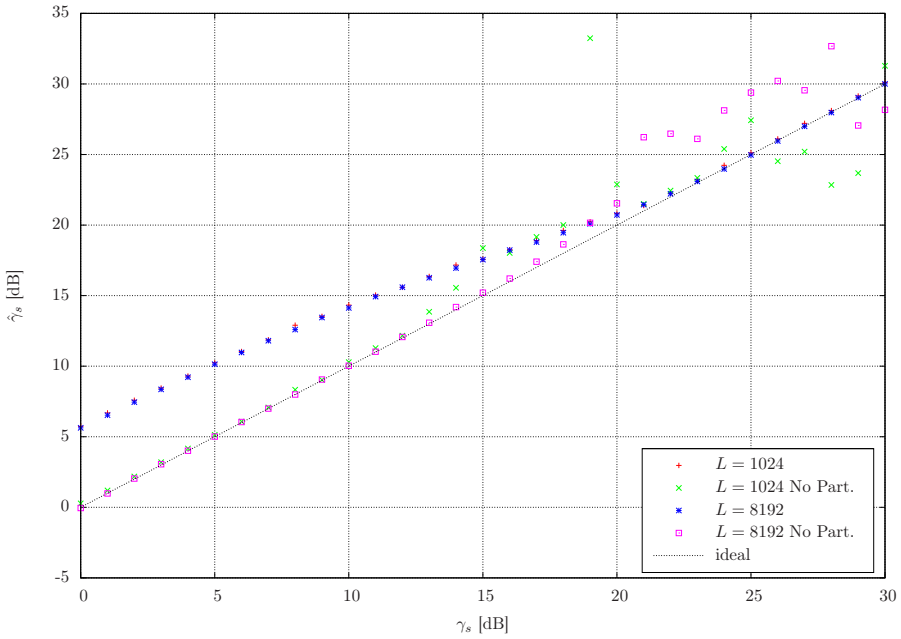


Figure 3.44: Mean SNR estimator output for different partitions (64-QAM)

of quotients of the form

$$f(\gamma_s) = h(\mathbf{m}) = \beta\frac{M_4}{M_2^2} + \gamma\frac{M_6}{M_2^3} + \delta\frac{M_4^2}{M_2^4} + \epsilon\frac{M_8}{M_2^4} \tag{3.64}$$

is a function of the SNR. Herein the vector $\mathbf{m} = [M_2\ M_4\ M_6\ M_8]^T$ and β, γ, δ and ϵ are the weights. It is suggested to use the normalized SNR $z = \gamma_s/(1+\gamma_s)$. Thus, the above equation can be written as

$$F(z) = F_1 z^4 + F_3 z^3 + F_2 z^2 + F_0 \tag{3.65}$$

with the coefficients

$$F_4 = \delta(c_4 - 2)^2 + \epsilon[72(c_4 - 1) - 16c_6 + c_8] \tag{3.66}$$
$$F_3 = (\gamma + 16\epsilon)(12 - 9c_4 + c_6) \tag{3.67}$$
$$F_2 = (\beta + 9\gamma + 4\delta + 72\epsilon)(c_4 - 2) \tag{3.68}$$
$$F_0 = 2(\beta + 3\gamma + 2\delta + 12\epsilon) \tag{3.69}$$

where $c_i = E[|x_k|^i]$ are the constellation moments. So, finally the estimator can be formulated as finding the root of the polynomial $F(\hat{z}) - h(\hat{\mathbf{m}})$. An iterative procedure for

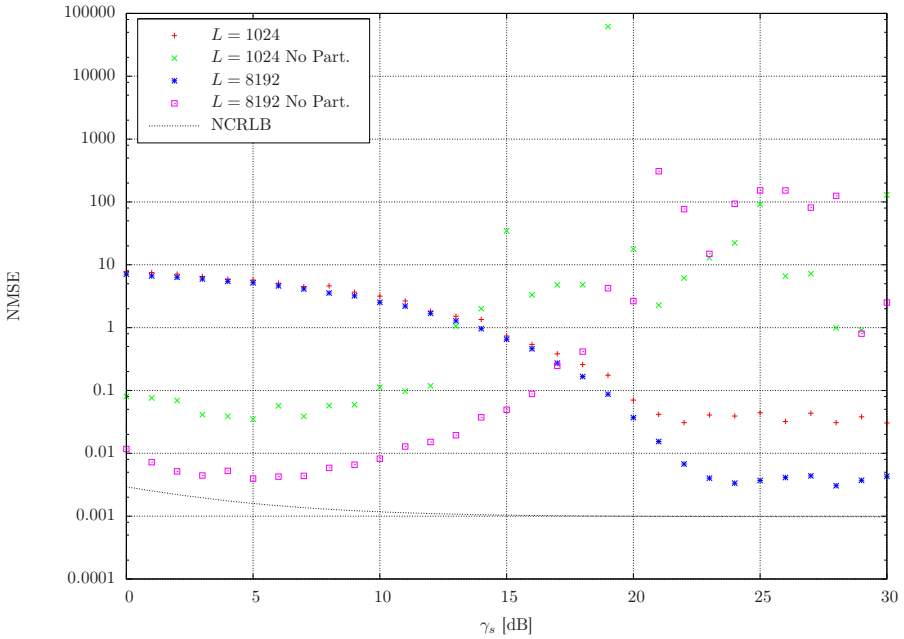
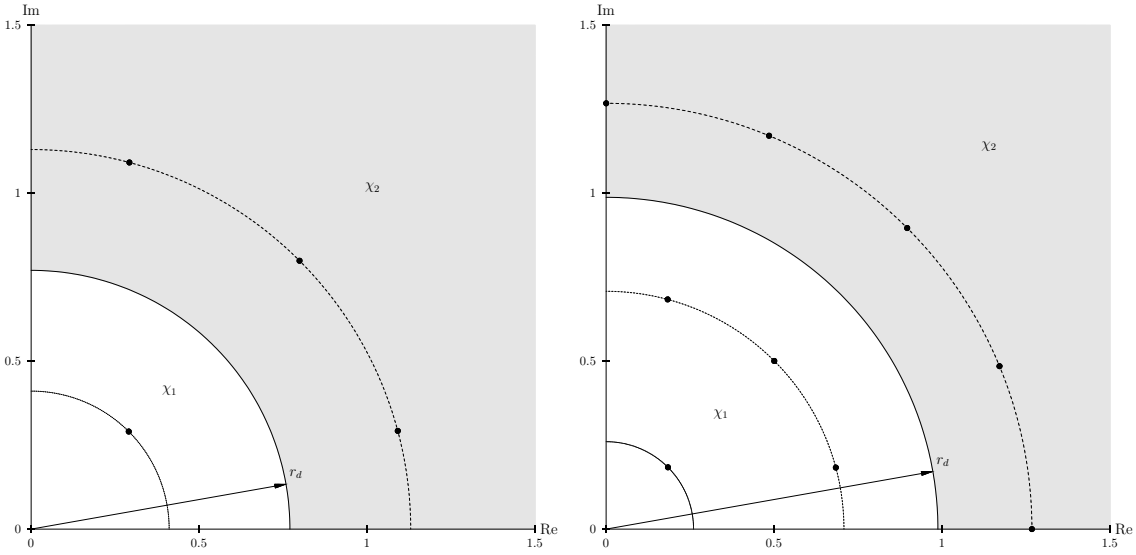


Figure 3.45: Evolution of NMSE for $\hat{\gamma}_s$ at different partitions (64-QAM)

finding the root can be written as

$$\hat{z}_{n+1} = \sqrt{\frac{h(\hat{\mathbf{m}}) - F_0}{F_4 \hat{z}_n^2 + F_3 \hat{z}_n + F_2}} \tag{3.70}$$

with initial value $\hat{z}_0 = 1$. The selection of the weights β, γ, δ and ϵ is depending on the signal constellation and on the SNR operation range. Two criteria exist: one for the high SNR range and one for a nominal SNR value. The authors of [51] provide Matlab and Mathematica source code for computation of the weights for both criteria [52]. The weights for 16-APSK and 32-APSK are calculated according to Criterion 1 and for 16-QAM and 64-QAM are tuned to nominal values in the middle SNR range by Criterion 2 (see Table 3.6). In Figure 3.49 and Figure 3.50 the mean estimator output and the NMSE evolution of 16-QAM is depicted. The performance characteristics of all three introduced schemes are shown, i.e. standard, partitioned and multi-level estimator. Moreover, the results for the combined estimator, choosing the most appropriate algorithm, can be seen. This combined version either takes an SNR hint from the pre-estimation stage as input or first executes the multi-level algorithm and afterwards, depending on the result, the standard or partitioned estimator. The SNR limits in which the multi-level algorithm is employed were derived by numerous simulation runs and are summarized in Table 3.7. Especially, from the NMSE evolution plot it can be seen that the multi-level



(a) 16-APSK **(b)** 32-APSK

Figure 3.46: Partitioning of APSK constellations

Table 3.6: Weights for multi-level SNR estimator

Modulation type	Criterion	β	γ	δ	ϵ
16-APSK	C1	6.1331	-2.8918	-1.4676	1.0
32-APSK	C1	12.0919	-3.8541	-2.7024	1.0
16-QAM	C2 (15.0 dB)	7.3773	-2.5011	-2.8080	1.0
64-QAM	C2 (17.0 dB)	9.0404	-3.9957	-1.6295	1.0

estimator closes the performance gap in the range from 10 to 15 dB. The most challenging constellation in terms of SNR estimation is 64-QAM. From the above results it should be clear that an increase of the estimator length compared to other modulation schemes is required to achieve satisfactory performance. This observation holds true for the multi-level method as well. In Figure 3.51 and Figure 3.52 the performance results for all three introduced algorithms are shown for 64-QAM. In this case the estimator length was set to $L = 8192$ symbols. Even with the increased observation interval, some deviation is recognized around 20 dB. Compared to an estimator length of $L = 1024$, the usable range for the multi-level approach is extended appropriately from 16 dB to 19 dB. When further increasing the number of processed symbols to $L = 32768$, the limit should be set to 21 dB and the performance is satisfactory over the whole range.

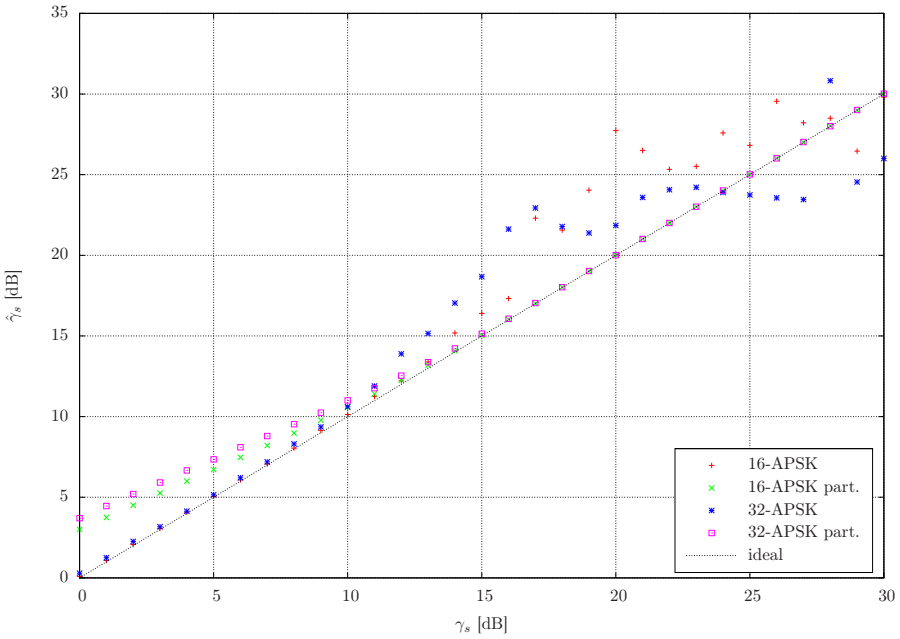


Figure 3.47: Mean SNR estimator output (APSK)

3.10.4 Combined Estimator

Finally, the performance of the combined algorithm that selects the most appropriate SNR estimator is depicted in Figure 3.53 and Figure 3.54 for all supported modulation schemes. The estimator length is set to $L = 1024$ for all signal constellations, except for 64-QAM which is operated with $L = 8192$. For this case, the asterisk should indicate that for a

Table 3.7: Limits for multi-level SNR estimator

Modulation type	Lower Limit [dB]	Upper Limit [dB]
16-APSK	8.0	14.0
32-APSK	7.0	14.0
16-QAM	10.0	16.0
64-QAM ($L = 1024$)	12.0	16.0
64-QAM ($L = 8192$)	12.0	19.0
64-QAM ($L = 32768$)	12.0	21.0

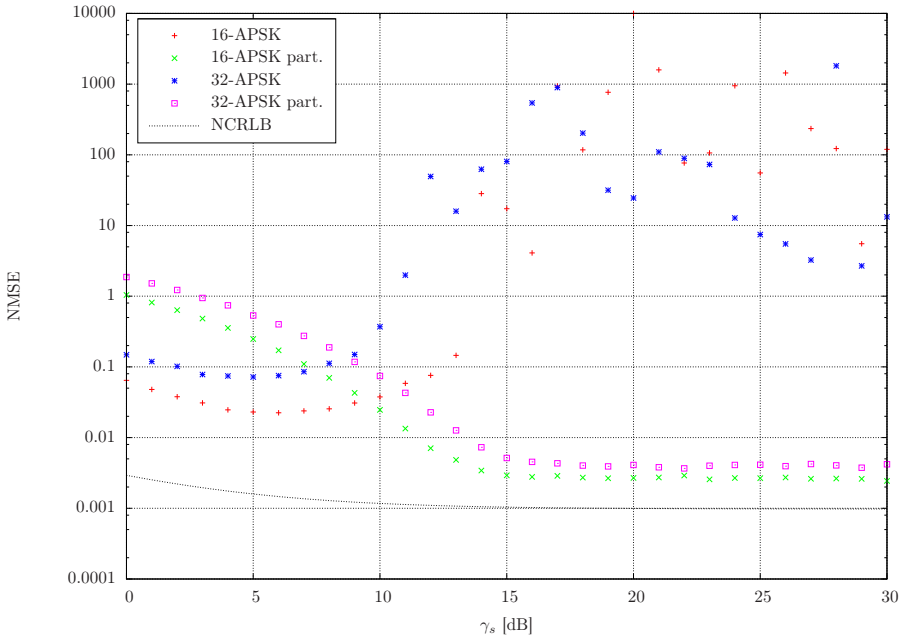


Figure 3.48: Evolution of NMSE for $\hat{\gamma_s}$ (APSK)

fair comparison among the modulation schemes the estimator length has to be adjusted to the same values. It can be seen that the performance of the combined algorithm is quite satisfactory over the SNR range of interest. The only case where the mean estimate deviates from the true value is for 64-QAM around 20 dB.

To avoid pitfalls it should be pointed out that the presence of ISI can completely deteriorate the performance of the presented SNR estimation algorithms. For the previous simulation runs, perfect recovery of the symbol timing was assumed. But, what would happen to the results in presence of a residual timing offset? This could occur when timing recovery omits appropriate interpolation and instead uses selection of the 'best sample'. Figure 3.55 shows the mean estimator output when the symbols are impaired by residual timing offsets. The errors used for the simulation are $\epsilon = 0$, $\epsilon = 1/8$, $\epsilon = 1/16$, $\epsilon = 1/32$ and $\epsilon = 1/64$, where QPSK is chosen as modulation scheme. It is obvious that, with increasing timing offset, the impairment gets more pronounced. This should be kept in mind when deciding for a 'best sample' approach. Of course, this fact has to be considered as well for the selection of appropriate timing recovery and interpolation schemes.

For implementation it is suggested to check the result of the square root which is present in all estimation algorithms (see Equations 3.59, 3.61 and 3.70). It can happen that the radicand gets negative which will lead to a complex number. The effect is especially observed for constellations with non-constant envelope when using the standard

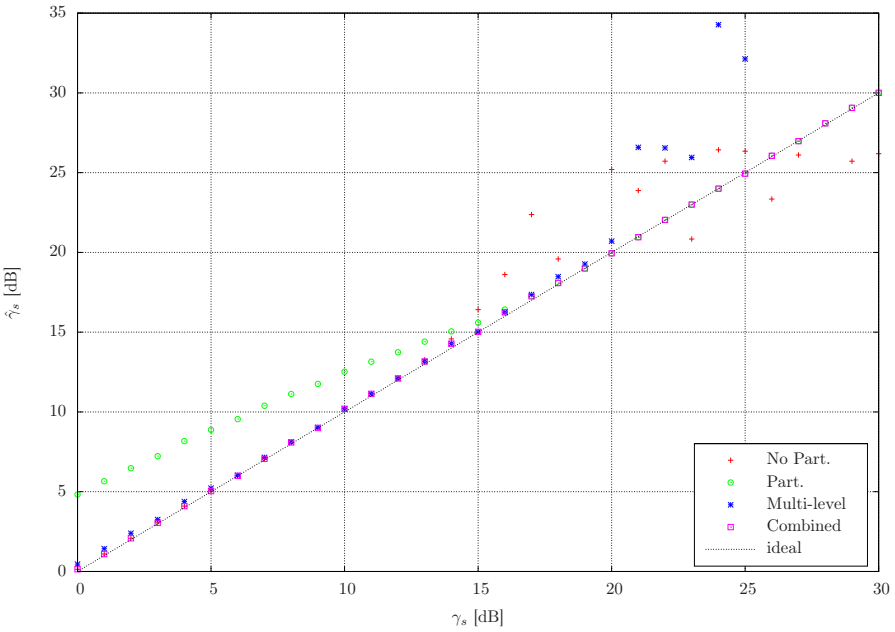


Figure 3.49: Mean multi-level SNR estimator output (16-QAM)

or multi-level estimator; the number of such estimation failures accumulates with increasing SNR.

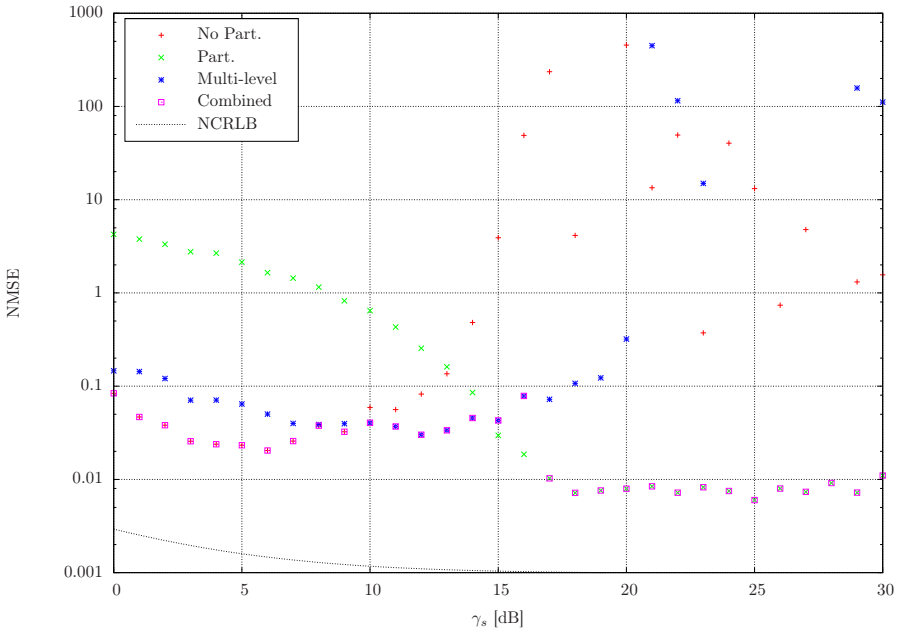


Figure 3.50: Evolution of NMSE for multi-level SNR estimator (16-QAM)

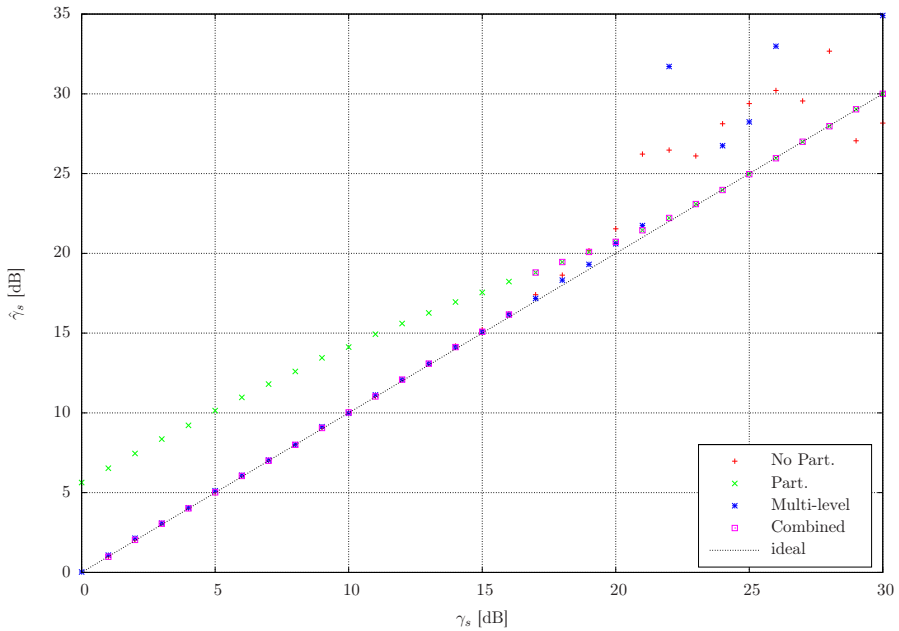


Figure 3.51: Mean multi-level SNR estimator output (64-QAM)

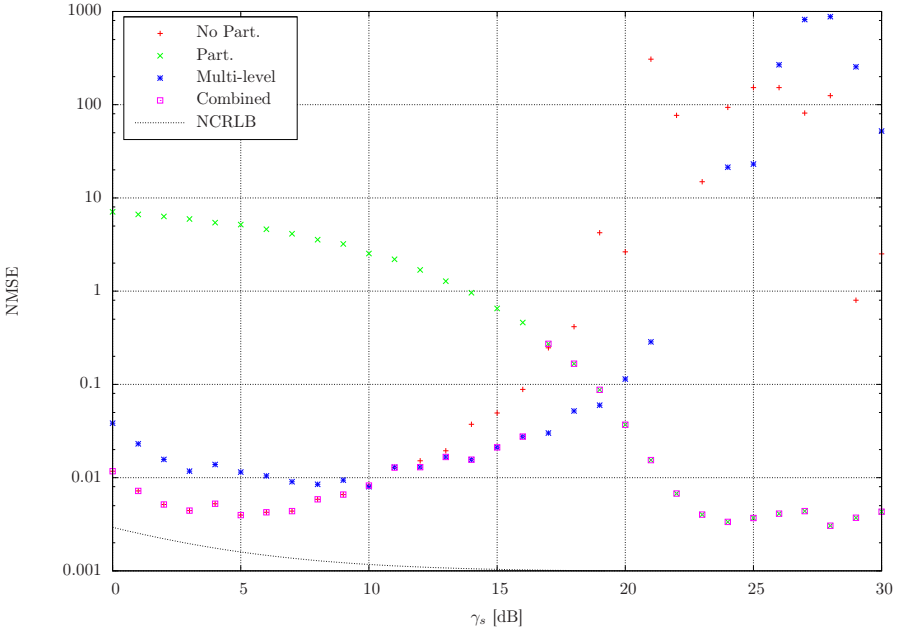


Figure 3.52: Evolution of NMSE for multi-level SNR estimator (64-QAM)

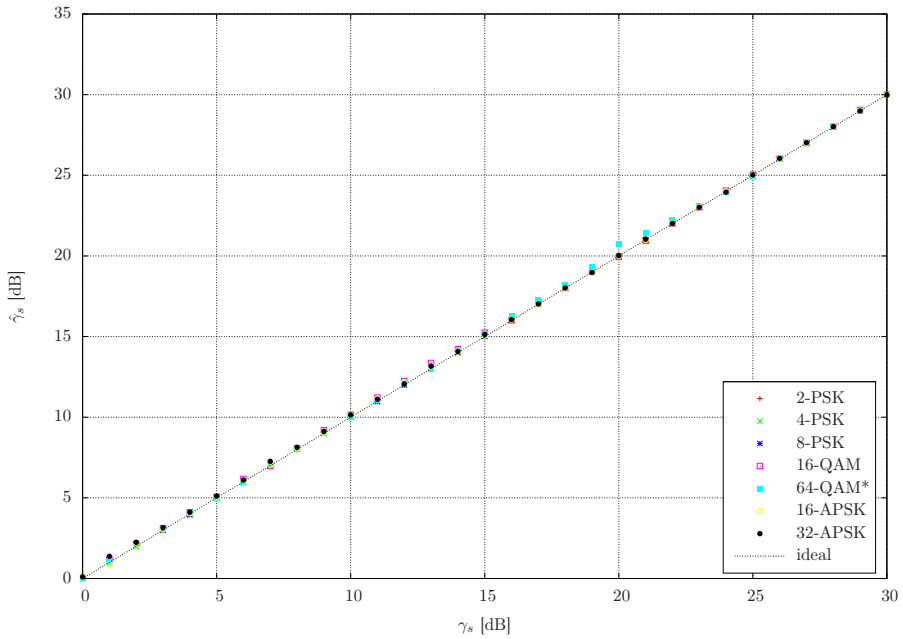


Figure 3.53: Mean SNR estimator output (combined estimator)

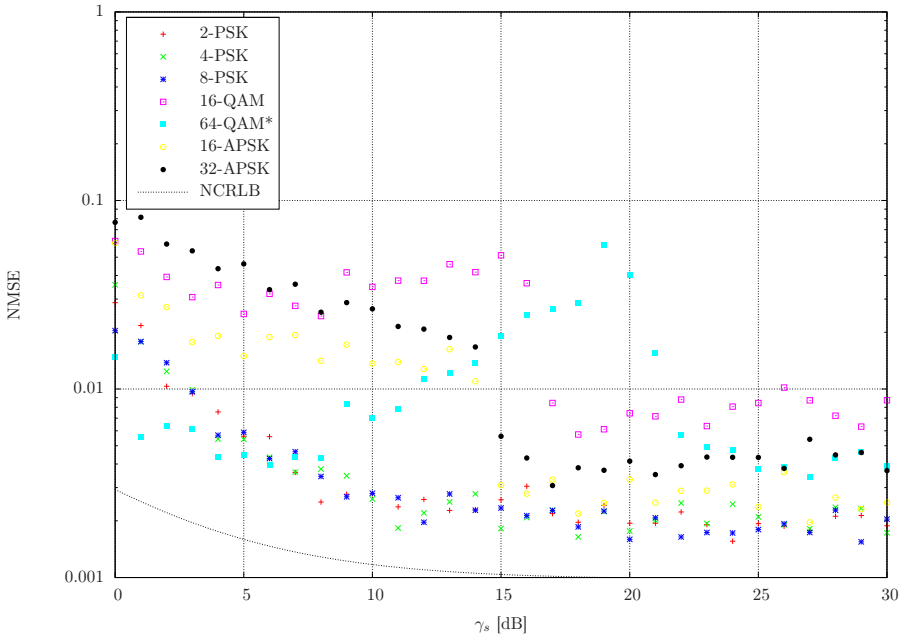


Figure 3.54: Evolution of NMSE for $\hat{\gamma}_s$ (combined estimator)

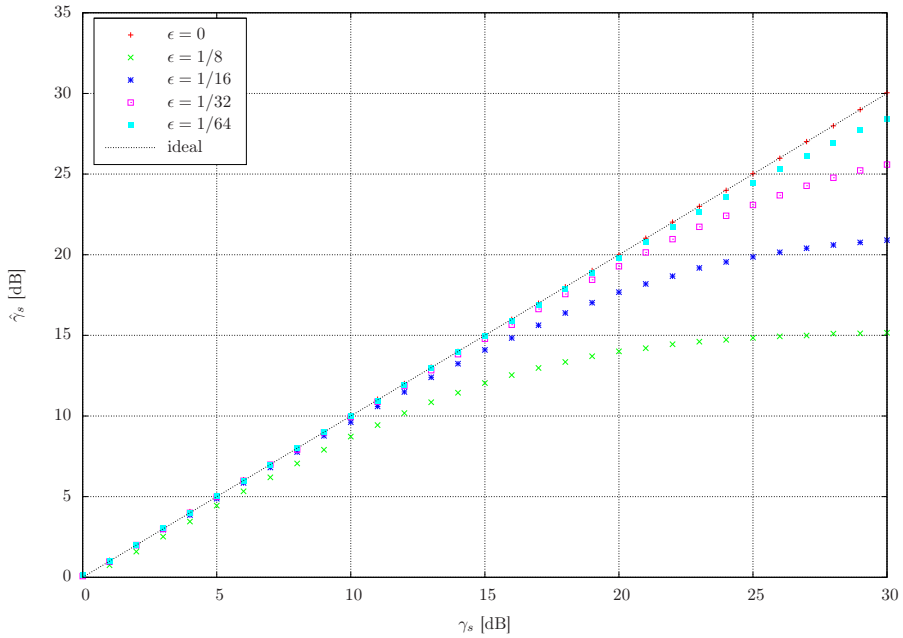


Figure 3.55: Mean SNR estimator output in presence of ISI

3.11 Roll-off Factor Estimation II

At this point of the blind demodulation process, the opportunity arises to determine the filter roll-off using a different approach compared to the one described earlier in Section 3.5. As discussed in Section 3.6, the main purpose of the MF is to maximize the SNR at its output. Considering this fact, the received signal can be filtered using different roll-off factors to determine the excess bandwidth for which the output has maximum SNR. Before applying the filtered signal to the SNR estimator, presented in Section 3.10, the optimum timing instant has to be recovered for proper operation. This necessitates the efficient algorithms introduced in Section 3.8 that exhibit satisfactory performance even for small roll-offs. In Figure 3.56 the block-diagram of the proposed estimator is illustrated.

Figure 3.56: Roll-off determination using SNR estimation

The measurement is repeated for N different roll-off factors such that N estimates for the SNR are available, i.e. $\Gamma_N = \{\hat{\gamma}_{s,1},\hat{\gamma}_{s,2},\ldots,\hat{\gamma}_{s,N}\}$. Thus, an estimate for α can be extracted as

$$\hat{\alpha} = \arg\max_{\alpha_i} \{\Gamma_N\} \tag{3.71}$$

An improvement of the accuracy can be achieved by applying spline interpolation on the values of Γ_N.

In the following, the performance of the SNR-based algorithm is compared to the IFFT method described in Section 3.5. For the required timing recovery the OM algorithm and a cubic Lagrange interpolator are used. The oversampling ratio is assumed to be $N_s = 4$, so that 16384 symbols are applied to the SNR estimator. In Figure 3.57 the mean estimator output of both approaches is depicted for QPSK and 16-APSK. For the IFFT method only QPSK is shown, since the approach is independent of the modulation scheme. It seems that the SNR method exhibits no bias like the IFFT approach. Interesting is the performance degradation for 16-APSK at an SNR of $\gamma_s = 15\,\text{dB}$. This SNR value coincides almost with the switching threshold between multi-level and partitioned SNR estimator, defined in Table 3.7. It seems that wrong assignments of the symbols to the partition sets are responsible for this behavior. This kind of degradation is even more pronounced for QAM constellations and 32-APSK. The MSE of both estimators is shown in Figure 3.58 for all supported constellations. It is obvious that the MSE performance of the SNR method is superior for M-PSK; however, for multi-level constellations the previously discussed SNR estimator problems affect the estimation result. For 16-QAM

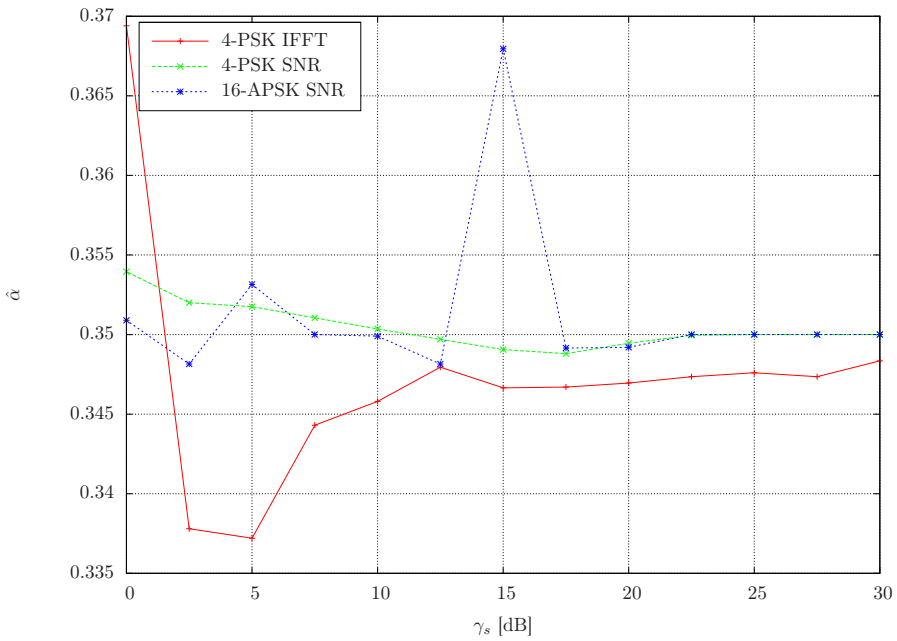


Figure 3.57: Mean estimator output for roll-off estimators

the MSE of the SNR algorithm drops below the one of the IFFT method not before $\gamma_s = 25\,\text{dB}$; 32-APSK improves at $\gamma_s = 17.5\,\text{dB}$ and 64-QAM outside the simulated range. For better understanding of this effect, the SNR estimator output versus the MF roll-off is depicted in Figure 3.59 for 32-APSK as modulation scheme for different SNR values. The estimator length is set to $L = 8192$ and the roll-off to $\alpha = 0.35$. As can be seen, for an SNR of $\gamma_s = 15\,\text{dB}$ the maximum SNR is not located at the true roll-off, but at the end of the simulated range due to wrong partitioning. At a lower SNR of $\gamma_s = 12.5\,\text{dB}$ without partitioning and at a higher SNR of $\gamma_s = 17.5\,\text{dB}$ with negligible partitioning errors, the SNR estimator produces a maximum at the expected roll-off. Thus, in the range of poor performance the use the IFFT method or increasing the estimator length and changing the switching threshold of the multi-level SNR estimator are suggested. However, care has to be taken since, with increasing SNR, stability issues occur more frequently for the multi-level estimator, as discussed in Section 3.10. Another property, identified from the simulation result, is that the SNR evolution for varying roll-off around the maximum is shaped rather flat, which can lead to estimation errors for short estimator lengths or in the low SNR range.

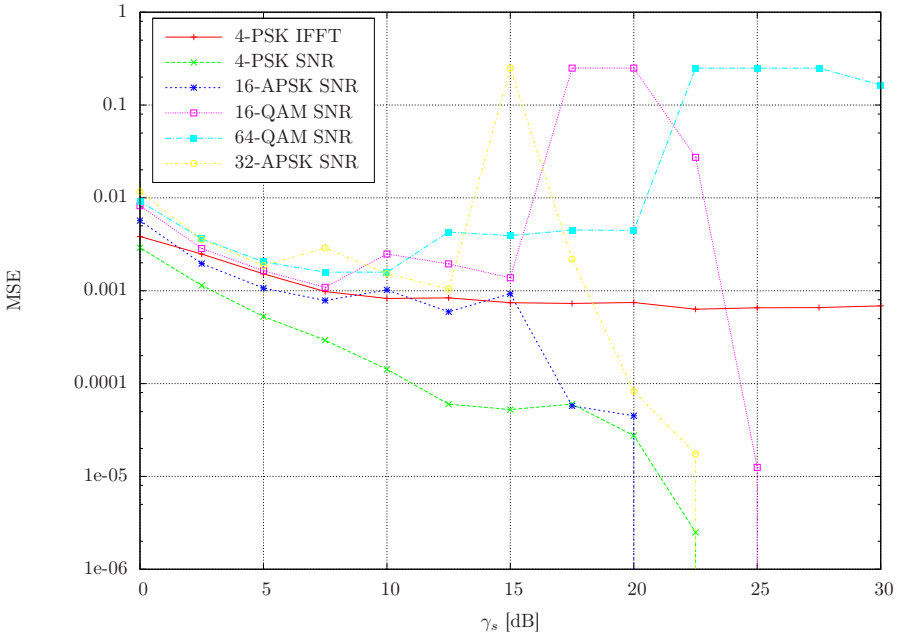


Figure 3.58: MSE comparison for roll-off estimators

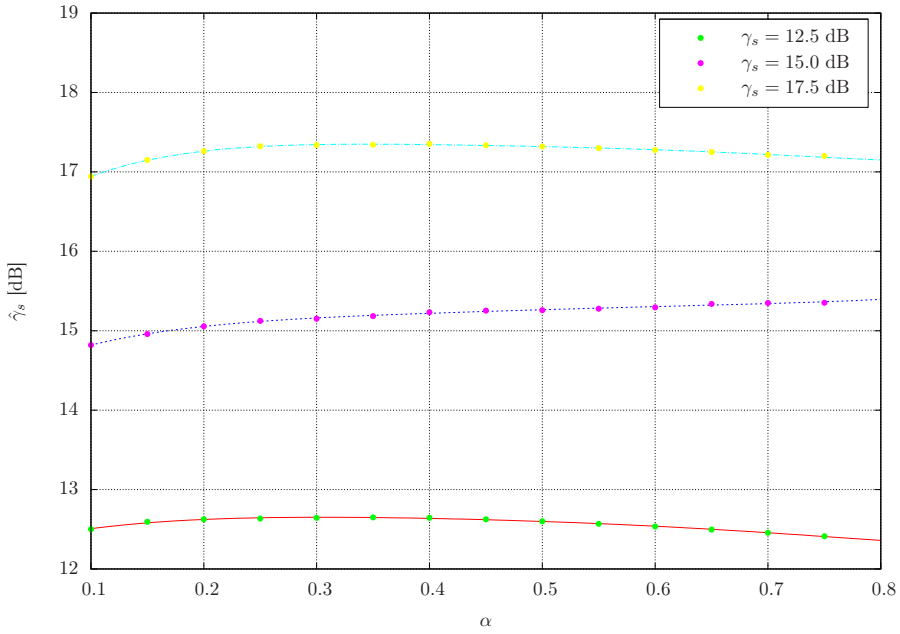


Figure 3.59: Mean SNR-based roll-off estimator output (32-APSK)

3.12 Frequency Estimation

Algorithms for NDA estimation of the normalized carrier frequency offset ΔfT are well established for M-PSK signal constellations [53]. The method by Rife and Boorstyn (RB) [54] or Luise and Reggiannini (LR) [55] are prominent examples in this respect. As described in [56], for determination of the estimate $\Delta\hat{f}T$, a nonlinearity is used to remove the effects of modulation, resulting in a harmonic with constant amplitude affected by additive noise. This sinusoid exhibits both frequency and phase that are multiples of the true frequency and phase offsets.

The nonlinear transformation suggested by Viterbi and Viterbi (VV) [57] is widely established. When writing the symbols, impaired by frequency and phase offset, in polar form as $x_n = \rho_n e^{\jmath\varphi_n}$, the nonlinearity can be formulated as

$$y_n = -\rho_n^{\mu}\, e^{\jmath m\varphi_n} \tag{3.72}$$

The design parameter μ is usually drawn from the set $\mu \in \{0,1,2\}$. For M-PSK signals, m corresponds to the order M of the modulation scheme; however, this is not true for multi-level constellations, as will be shown below.

The least squares solution and the RB algorithm, derived from the ML estimator, are equivalent. The estimator can be written as

$$\Delta\hat{f}T = \frac{1}{k_{zp}Lm} \arg\max_{k} \left| \sum_{k=0}^{k_{zp}L-1} y_k\, e^{-\jmath 2\pi k/k_{zp}L} \right| \tag{3.73}$$

Herein, L is the number of used symbols and k_{zp} the zero-padding factor to increase the resolution of the frequency axis. The DFT, contained in the above equation, can be computed efficiently by using an FFT of length $N = k_{zp}L$. The location of the identified maximum is refined by means of parabolic interpolation (see Appendix A.1). Similar to the symbol rate estimator, introduced in Section 3.3.1, the result will either be accurate or completely wrong, when identifying the highest spectral line as estimate. This type of error is termed as outlier in the open literature. The latter manifests itself as a significant increase of the MSE (threshold effect). One main advantage of the RB estimator is the wide operational range of $|\Delta fT| < 1/(2mT)$. The basic feedforward frequency offset estimation and the correction principle is depicted in Figure 3.60. Therein, the delayed symbols are multiplied by a complex exponential with a frequency $-\Delta\hat{f}T$ for derotation purposes. The dashed pre-processing block performs the required partitioning, presented in the following and necessary for several modulation schemes.

The above described procedure is working properly for M-PSK signals, however, problems arise for multi-level constellations since no promising way for stripping off the modulation from the symbols x_n is available. A way to tackle the problem is to apply partitioning, similar to SNR estimation in Section 3.10.2. This strategy is exemplified in [56] for 16-QAM. A similar method was already suggested in [58], although approaching the

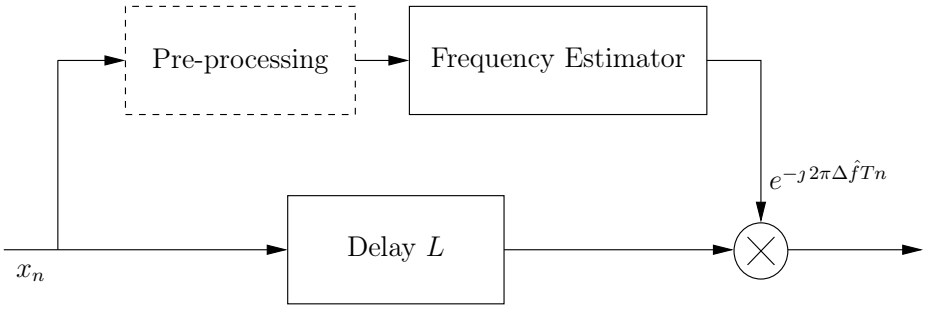


Figure 3.60: Feedforward frequency offset recovery

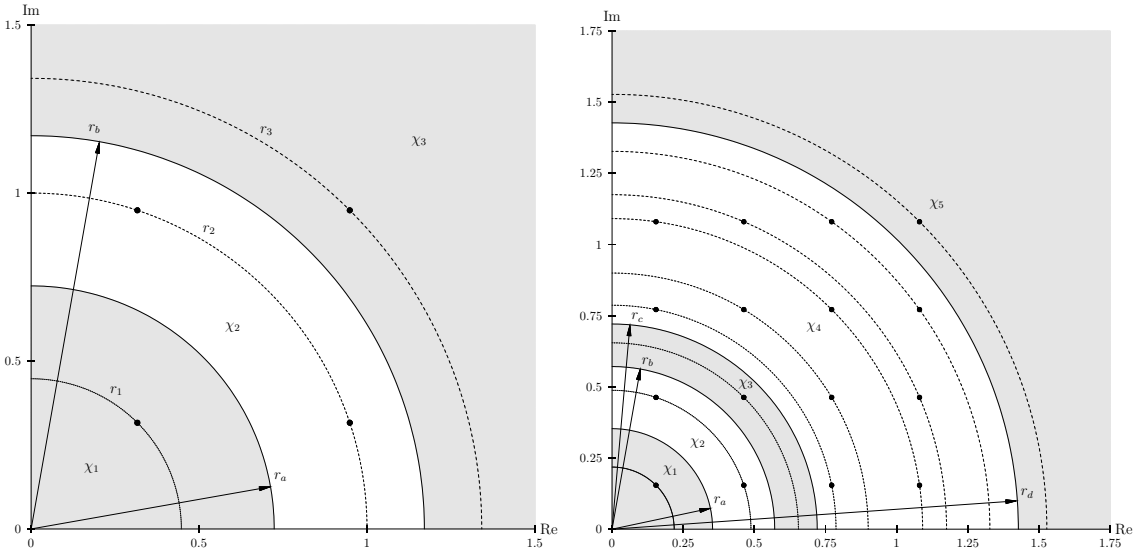
problem from a different perspective. In more recent papers it is tried to find an optimal nonlinear transformation such that the asymptotic variance of the estimator is minimized. This approach is followed in [59] for circular QAM signals and in [60] for general QAM constellations, including 16-QAM and 64-QAM. The optimal transformation depends on the signal constellation as well as on the SNR. To circumvent the complex calculation of the optimum transformation, a linear approximation is applied. The latter can be used for a wider SNR range and turns out to be related closely to the standard partitioning proposed in [56]. In the following, partitioning procedures will be described to support multi-level constellations.

M-QAM The first quadrant of the 16-QAM constellation is illustrated in Figure 3.61a. It can be seen that the symbols located on the middle circle are not spaced equidistantly and thus usage of the VV nonlinearity would lead to adverse effects. In contrast, the data points on the inner and outer circle form a QPSK constellation and can be considered for the estimation procedure. As indicated in the illustration, three subsets of symbols can be formed, i.e. χ_1, χ_2 and χ_3, such that a partitioning procedure can be formulated as

$$\rho_n = \begin{cases} |x_n| & x_n \in \chi_1 \\ 0 & x_n \in \chi_2 \\ |x_n| & x_n \in \chi_3 \end{cases} \tag{3.74}$$

When calculating y_n, the argument of x_n has to be multiplied by $m = 4$. If assuming equiprobable symbols, on average only half of the data is used for estimation and thus will lead to degradation of performance.

Figure 3.61b shows the subsets χ_1 to χ_5 that can be formed for 64-QAM using similar



(a) 16-QAM **(b)** 64-QAM

Figure 3.61: Partitioning of QAM constellations for frequency offset estimation

considerations as for 16-QAM. The partitioning can be summarized as

$$\rho_n = \begin{cases} |x_n| & x_n \in \chi_1 \\ 0 & x_n \in \chi_2 \\ |x_n| & x_n \in \chi_3 \\ 0 & x_n \in \chi_4 \\ |x_n| & x_n \in \chi_5 \end{cases} \tag{3.75}$$

M-APSK For APSK constellations the partition subsets defined previously in Section 3.10.2 can be used as shown in Figure 3.46. Therefore, the following procedure can be used to compute y_n:

$$\rho_n = \begin{cases} 0 & x_n \in \chi_1 \\ |x_n| & x_n \in \chi_2 \end{cases} \tag{3.76}$$

This produces a 12-PSK constellation for 16-APSK and a 16-PSK constellation for 32-APSK residing on the outermost circle. Hence, the multiplier of the argument of x_n has to be chosen as $m = 12$ for 16-APSK and $m = 16$ for 32-APSK.

In the following, simulation results will be presented which demonstrate the performance of the introduced estimator. The MCRLB is used to compare the performance of the

algorithms and is defined as [19]

$$\text{MCRLB} = \frac{3}{2\pi^2}\frac{1}{L(L^2-1)\gamma_s} \tag{3.77}$$

In Figure 3.62 the evolution of the MSE is depicted. The estimator length is set to

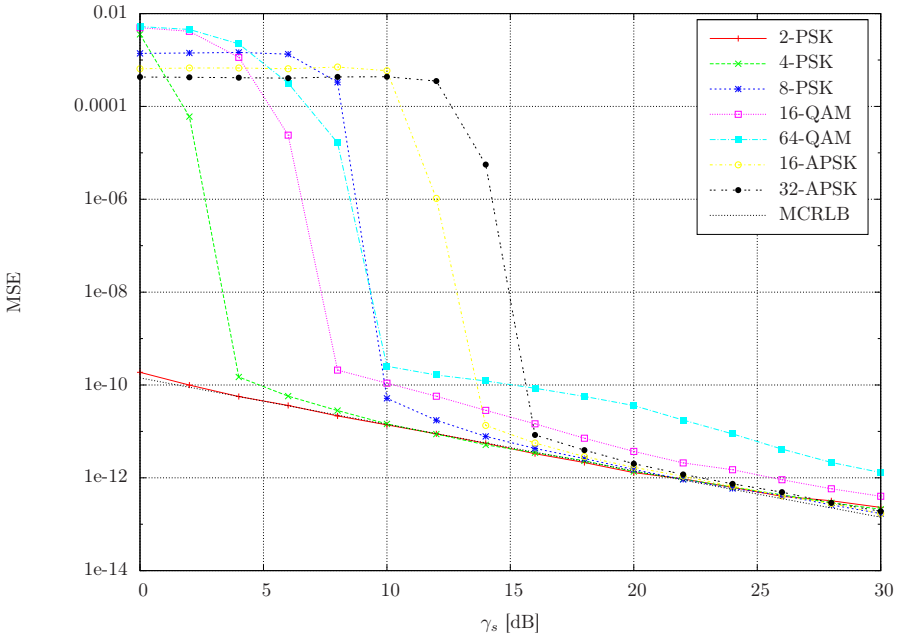


Figure 3.62: Frequency MSE for different modulation schemes

$L = 1024$ symbols, with a zero-padding factor of $k_{zp} = 4$ leading to an FFT length of 4096 points. The roll-off factor is assumed to be $\alpha = 0.35$. The Viterbi parameter is set to $\mu = 1$, since it produces the most promising results. Finally, the frequency offset is assumed to be $\Delta fT = 0.01$. The aforementioned threshold effect can be observed very clearly in the low SNR range. Moreover, the poor performance of both APSK schemes should be emphasized, but no promising alternative algorithm was found up to now. It seems that an SNR of at least $\gamma_s = 16\,\text{dB}$ is required for satisfactory performance of 32-APSK, otherwise the performance is degraded by the threshold effect. Increasing the estimator length L will shift the threshold in direction of smaller SNR values. The latter effect can be observed in Figure 3.63 for 16-QAM and 32-APSK. It is obvious that, although increasing L, the threshold still remains in the medium SNR range. Of course, this behavior is not acceptable and thus future work will focus on this topic.

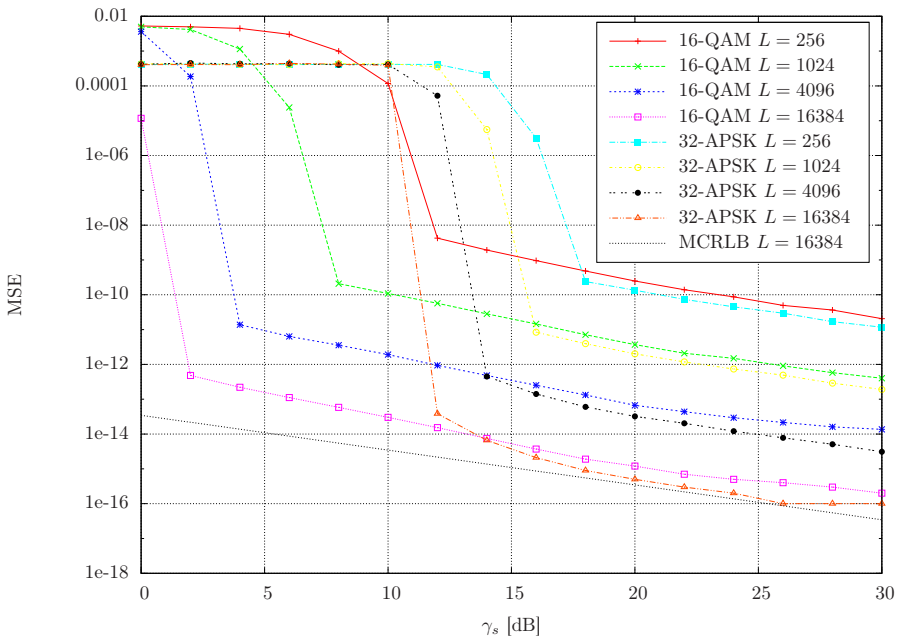


Figure 3.63: Frequency MSE for different estimator lengths

3.13 Phase Estimation

As mentioned in the previous section, by applying a nonlinearity on the symbols a harmonic with frequency and phase being multiples of the true offsets is generated. With $x_n = \rho_n e^{\jmath\varphi_n}$ the same nonlinearity y_n as for frequency estimation, proposed by Viterbi and Viterbi [57], can be used:

$$y_n = -\rho_n^{\mu}\, e^{\jmath m\varphi_n} \tag{3.78}$$

In the above formula the minus sign has to be removed, if BPSK or 32-APSK are applied as modulation schemes. Thus an estimate for the phase offset can be derived as

$$\hat{\theta} = \frac{1}{m} \arg\left(\sum_{k=0}^{L-1} y_n\right) \tag{3.79}$$

The basic arrangement of the above estimator in a feedforward manner is illustrated in Figure 3.64. It is clear that multiplication of the phase φ_n by the factor m will lead to an m-fold ambiguity. Typically, the latter will be resolved by use of a known sequence of symbols or by application of differential encoding. This ambiguity is irrelevant for investigation of the estimator performance and, as will be shown in Section 3.14, for the tasks to solve in the presented demodulation framework as well.

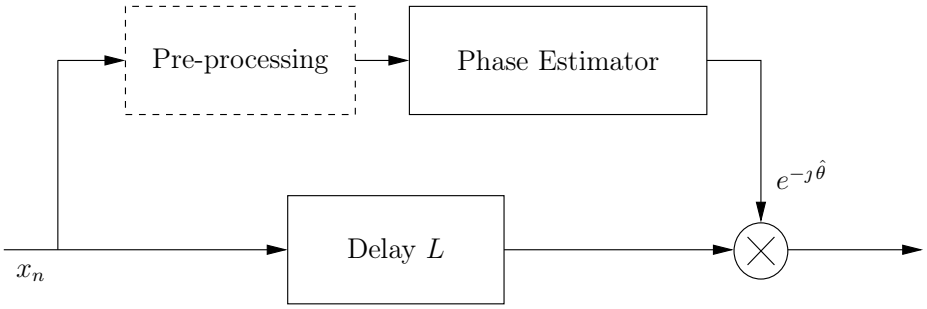


Figure 3.64: Feedforward phase offset recovery

The related MCRLB as the theoretical limit of the jitter variance is given by [19]

$$\text{MCRLB} = \frac{1}{2L\gamma_s} \tag{3.80}$$

For the subsequent simulation runs, $L = 1024$ symbols are processed for phase estimation. For each simulation point, 1000 iterations are performed. It is assumed that timing recovery is ideal and no frequency offset is present such that $\Delta fT = 0.0$. Finally, the roll-off is set to $\alpha = 0.35$ and the Viterbi design parameter to $\mu = 2$.

Figure 3.65 shows the evolution of the MSE of the estimate $\hat{\theta}$ versus the SNR for the supported PSK modulation schemes. It can be seen that the MSE of $\hat{\theta}$ reaches the MCRLB in the medium SNR range. For low SNR values the MSE is highly dependent on the order of modulation. Basically, with increasing order M the angular and radial jitter due to noise is more pronounced, which results in an augmented MSE.

3.13.1 M-QAM

Since the same nonlinear transformation is used as previously for frequency estimation, the same considerations on partitioning apply. This would suggest the same procedure given by Equation 3.74 for 16-QAM and Equation 3.75 for 64-QAM. Another estimation scheme for 16-QAM that applies similar partitioning with an additional iteration using the symbols located on the middle circle is presented in [61]. A promising hybrid NDA/DD scheme for 16-QAM, denoted as two-pass (TP) estimator, is introduced in [56]. At first an estimate $\hat{\theta}$ for the phase offset is derived using the previously described partitioning approach. The result is used to correct the symbols for their phase offset by

$$z_n = x_n e^{-\jmath\hat{\theta}} \tag{3.81}$$

The corrected symbols z_n are still impaired by a residual phase offset $\Delta\theta = \theta - \hat{\theta}$. A second DD iteration tries to estimate the latter. The required decisions $\hat{c}_n$ are derived

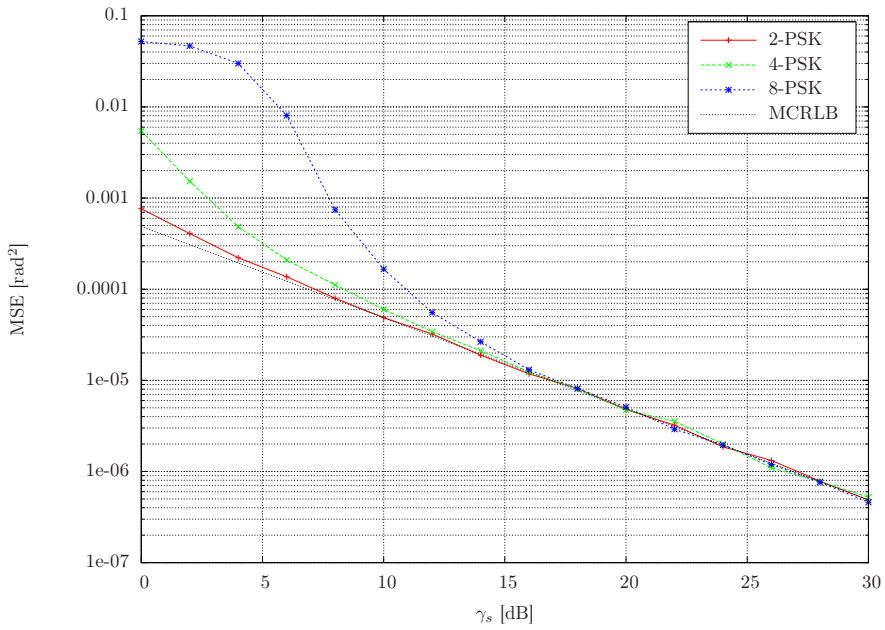


Figure 3.65: FF phase MSE for M-PSK

from z_n such that the estimator can be written as

$$\Delta\theta = \arg\left(\sum_{k=0}^{L-1} z_n \hat{c}_n^*\right) \tag{3.82}$$

The final estimate $\hat{\theta}'$ can immediately be calculated by

$$\hat{\theta}' = \hat{\theta} + \Delta\hat{\theta} \tag{3.83}$$

The blind demodulation framework uses the TP estimator for both 16-QAM and 64-QAM. A performance comparison of the different approaches for QAM constellations is depicted in Figure 3.66. It can be seen that, using the VV estimator without partitioning (denoted by NP) leads to a significant jitter floor in the medium-to-high SNR range. By applying the described partitioning scheme (P) the noise floor can be lowered considerably. However, for both QAM constellations the MSE exhibits an offset to the bound even in the high SNR range. Finally, when using the TP estimator, 16-QAM converges to the bound in the medium and 64-QAM in the high SNR range. Thus, the TP method seems to be an appropriate choice with reasonable complexity.

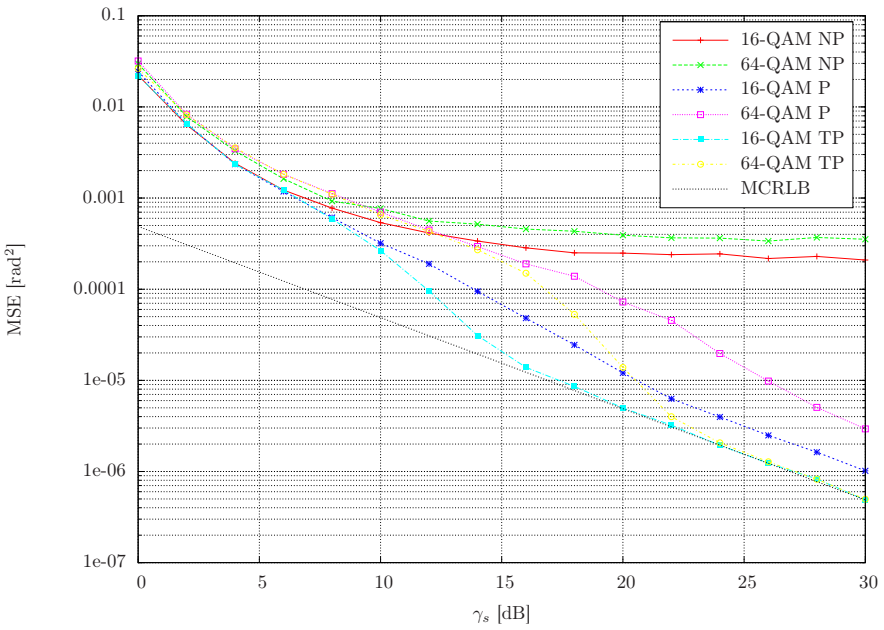


Figure 3.66: FF phase MSE for M-QAM

3.13.2 M-APSK

Algorithms which are dedicated to APSK are surprisingly rare in the open literature. In two recent papers, [62] and [63], feedback hybrid NDA/DD schemes working with 16-APSK and 32-APSK are introduced. A comparison of the algorithms to the VV and power law algorithm with and without partitioning is extensively presented in [64]. Therein it is concluded that the performance of the hybrid algorithm does not differ significantly to that of the VV estimator, when choosing the design parameter $\mu \leq 2$. Thus, it seems convenient to choose the VV algorithm for phase recovery of APSK constellations.

For subsequent symbol decisions on the corrected signal z_n, an issue arises which is unique for APSK. The constellations are generated using Equation 2.10 and Equation 2.14. Thus, the symbols located on the inner circle correspond to a QPSK constellation with the symbols rotated by $\pi/4$ relative to the axis. In contrast, the outermost circle of 16-APSK exhibits an ambiguity of $\beta = \pi/6$ and of 32-APSK of $\beta = \pi/8$. Due to this fact it is likely that the QPSK constellation on the inner ring will remain tilted after phase recovery, which will lead to problems during symbol decision. In Figure 3.67 the symbol constellation for 16-APSK and 32-APSK is re-illustrated for better understanding. The circles correspond to the signal constellation as required for symbol decision. On the other hand, the red crosses denote possible corrected symbols after perfect NDA phase recovery. It is obvious that the symbols on the outer circle coincide with the desired constellation; however, the

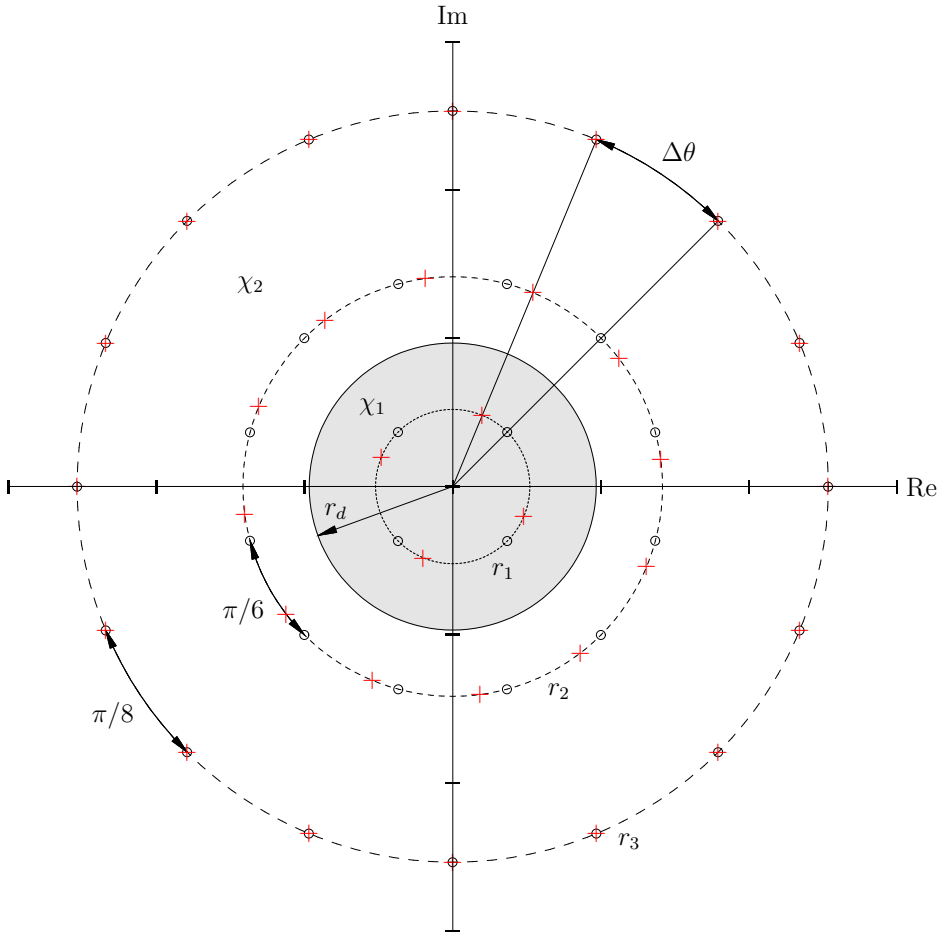


Figure 3.67: Ambiguity scenario for 16-APSK and 32-APSK constellation

QPSK constellation on the inner circle remains tilted by $\Delta\theta$. The only way to tackle this problem is to generate a constellation that is ambiguous by $\beta = \pi/2$, such that the inner QPSK star is properly aligned. For this reason a second iteration is performed on the corrected symbols z_n. After appropriate partitioning, phase estimation is performed on the remaining QPSK constellation to obtain an estimate $\Delta\hat{\theta}$. A subset χ_1 can be formed which includes all symbols with magnitude $0 \leq |z_n| < r_d = (r_1 + r_2)/2$, where r_1 and r_2 denote the radii of the first and second circle. The second subset χ_2 includes all symbols which are not part of χ_1. Thus, the partitioning procedure can be stated as

$$\rho_n = \begin{cases} |z_n| & z_n \in \chi_1 \\ 0 & z_n \in \chi_2 \end{cases} \tag{3.84}$$

Since on average only a fraction of the overall symbols, i.e. 1/4 for 16-APSK and 1/8 for 32-APSK, remains, increasing the estimator length is recommended. The symbols on the inner circle are more sensitive to noise effects due to their small amplitude levels. Hence, directly adding the estimated QPSK offset $\Delta\hat{\theta}$ to the first estimate $\hat{\theta}$ would lead to inferior performance. The phase corrected symbols z_n exhibit negligible residual phase error when assuming an accurate estimate $\hat{\theta}$. Thus, the QPSK estimate $\Delta\hat{\theta}$ has to be a multiple of the ambiguity β of the outermost circle. Keeping this in mind, a quantized estimate can be generated by

$$\Delta\hat{\theta}' = \begin{cases} \left\lfloor \frac{\Delta\hat{\theta}}{\beta} + 1/2 \right\rfloor \beta & \Delta\hat{\theta} \geq 0 \\ \left\lfloor \frac{\Delta\hat{\theta}}{\beta} - 1/2 \right\rfloor \beta & \Delta\hat{\theta} < 0 \end{cases} \tag{3.85}$$

In short, $\Delta\hat{\theta}$ is divided by the ambiguity β and rounded to the nearest integer value, thus when multiplied by β a quantized estimate is generated. In the above equation $\lfloor x \rfloor$ corresponds to the floor operation which rounds x to the next smaller integer value. So the final phase estimate which aligns APSK constellations poperly for subsequent symbol decision can be formed as

$$\hat{\theta}' = \hat{\theta} + \Delta\hat{\theta}' \tag{3.86}$$

Figure 3.68 shows the MSE evolution for the estimator without partitioning (NP), the partitioned counterpart (P) and partitioned estimator plus additional alignment of the QPSK constellation (Amb). It can be seen that the performance of all three algorithms is

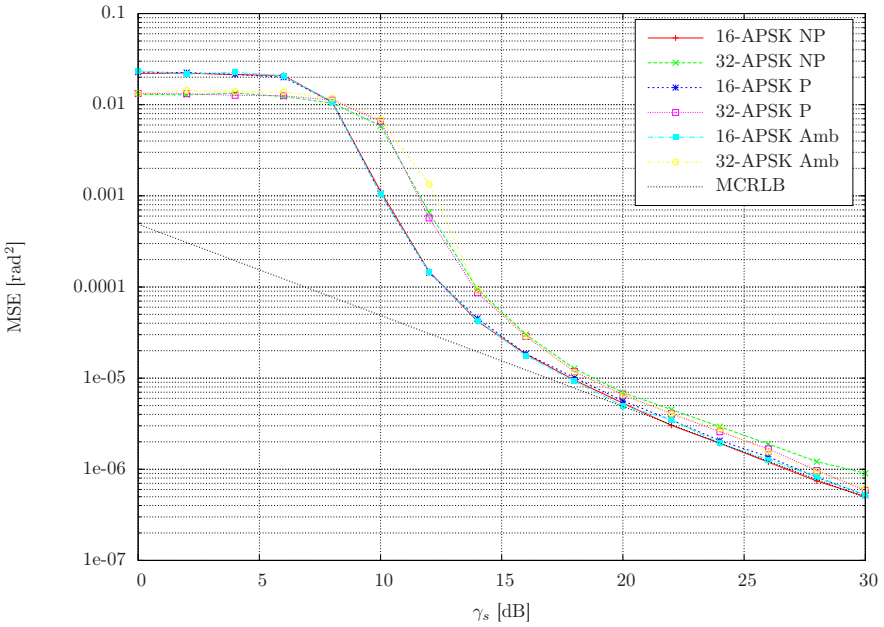


Figure 3.68: FF phase MSE for M-APSK

nearly the same. The only differences that can be identified are a slight improvement in the high SNR range due to partitioning and a slight degradation in the threshold region for 32-APSK when aligning the QPSK constellation.

3.13.3 Tracking

Up to now it was assumed that the phase offset θ is constant in the interval of L symbols. The latter is not true, if a significant residual frequency offset is present due to limited accuracy of the frequency estimator. Problems may arise as well, if the estimator length L is much smaller than the actual size of data that should be processed, such that even a small frequency offset will have an impact on the phase.

A way out of this dilemma is the concatenation of a feedforward scheme and a phase tracker. The feedforward method delivers an initial estimate which will serve the feedback scheme as starting value. A thorough analysis of the VV detector can be found in [65]. The error signal of the detector is calculated according to

$$u_n = \frac{1}{m} \arg\left(y_n\right) \tag{3.87}$$

A requirement for the phase tracker is to handle residual frequency offsets. Using a first-order loop in presence of frequency errors would result in a static phase offset [19]. Thus, an additional integrator is necessary to arrive at a second-order loop which is able to compensate for the rotation. Figure 3.69 shows the application of the phase detector in a feedback manner using a second-order loop. As shown in Section 3.8, the transfer

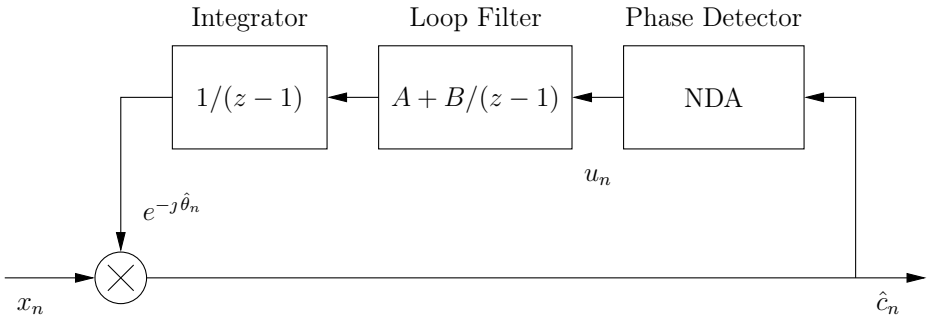


Figure 3.69: Feedback phase recovery using second-order loops

function $F(z)$ corresponds to a constant gain factor K_F in case of a first-order loop. For a second-order loop, $F(z)$ can be stated as

$$F(z) = A + \frac{B}{z-1} \tag{3.88}$$

Specifying the normalized equivalent noise bandwidth $B_L T$ and the damping factor ζ, the

normalized natural frequency can be determined by

$$\omega_n T = \frac{2B_L T}{\zeta + 1/(4\zeta)} \tag{3.89}$$

Reasonable values for ζ are in the range of $\zeta \in [0.6,1.0]$. The determination of the detector slope K_d depending on the design parameter μ and the SNR γ_s requires the evaluation of a complex formula [65], which simplifies for $\mu = 2$:

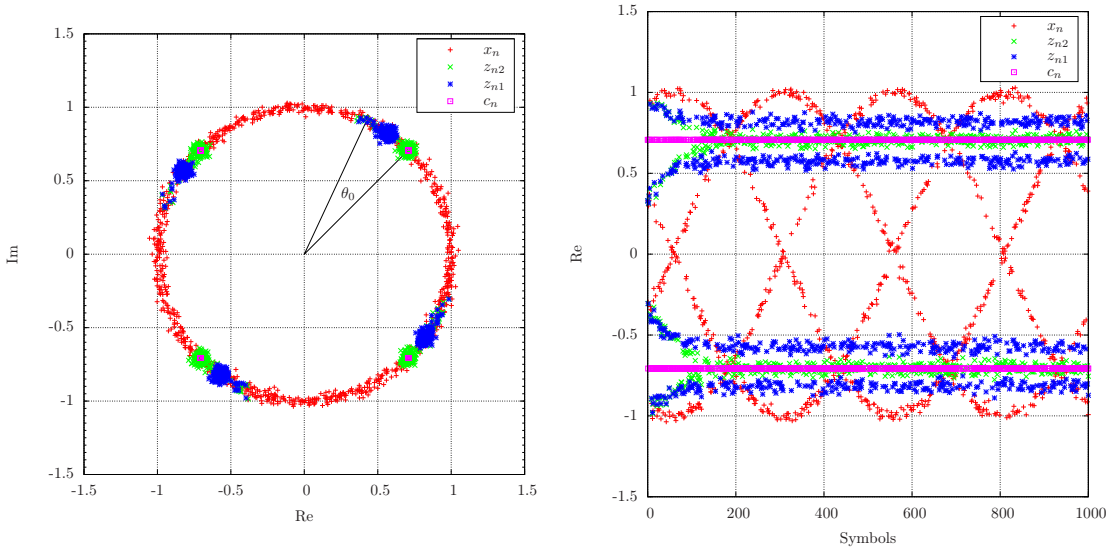
$$K_d(\gamma_s) = \frac{6(e^{-\gamma_s} - 1) + \gamma_s(6 + \gamma_s(\gamma_s - 3))}{\gamma_s^3} \tag{3.90}$$

Using the above results, the coefficients A and B can be computed according to

$$A = 2\zeta\omega_n T/K_d \tag{3.91}$$

$$B = (\omega_n T)^2/K_d \tag{3.92}$$

Figure 3.70 shows the relevant symbols before and after phase recovery using first- and second-order loops in presence of frequency error. For simulation the normalized equivalent



(a) Real/Imag **(b)** Time/Real

Figure 3.70: Phase recovery in the presence of frequency errors

noise bandwidth is set to $B_L T = 0.01$ and the damping factor to $\zeta = 1.0$. QPSK is used as modulation scheme with a roll-off of $\alpha = 0.35$. The rotating red crosses correspond to the received symbols x_n, impaired by an initial phase offset $\theta_0 = \pi/9$ and a residual frequency offset of $\Delta f T = 0.001$. The magenta squares are the original transmitted data

values c_n. When using a loop of order $n = 1$, delivering the symbols z_{n1}, the phase tracker starts at the initial value of θ_0; however, due to the frequency offset a residual phase shift will remain. In case of a second-order loop, with corrected symbols z_{n2}, the algorithm corrects for the phase offset without being affected by the frequency offset. The tracker requires approximately 200 symbols to reach the steady state condition.

Subsequently, simulation results of the MSE performance in the steady state case, i.e. initial phase offset $\theta_0 = 0$, are shown. In Figure 3.71 the MSE evolution when using a

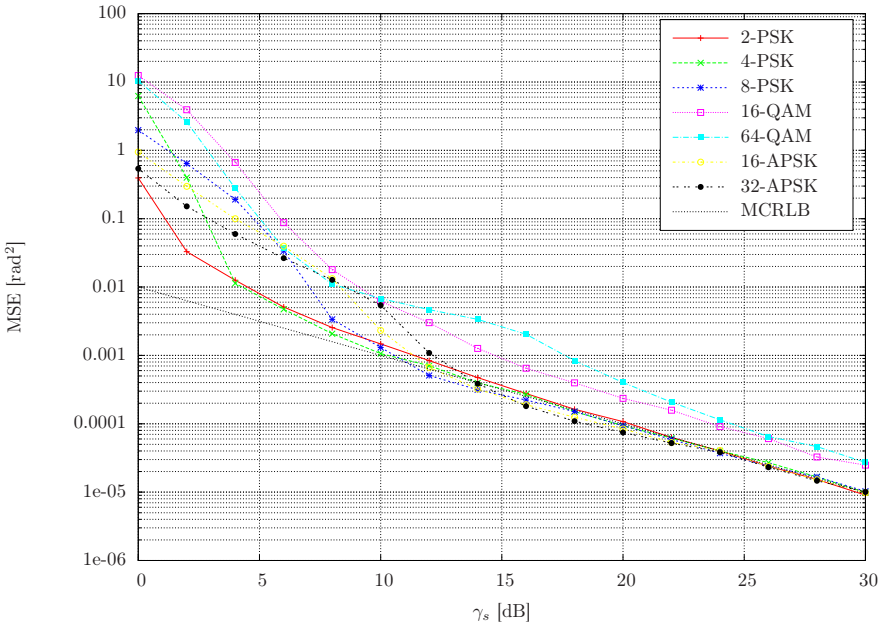


Figure 3.71: FB phase MSE for first-order loops

first-order loop and in Figure 3.72 for a second-order loop are depicted. For comparison reasons the MCRLB is added to the graphs as well. For feedback schemes the latter can be stated as [19]

$$\text{MCRLB} = \frac{B_L T}{\gamma_s} \tag{3.93}$$

Generally, the second-order loop exhibits increased MSE in the low SNR region. Since the TP algorithm is not used, an offset to the MCRLB can be observed for QAM in the medium-to-high SNR range which is more pronounced for the first-order loop. The choice of $B_L T$ is a tradeoff between accuracy of the estimate and the ability to track changes in the signal.

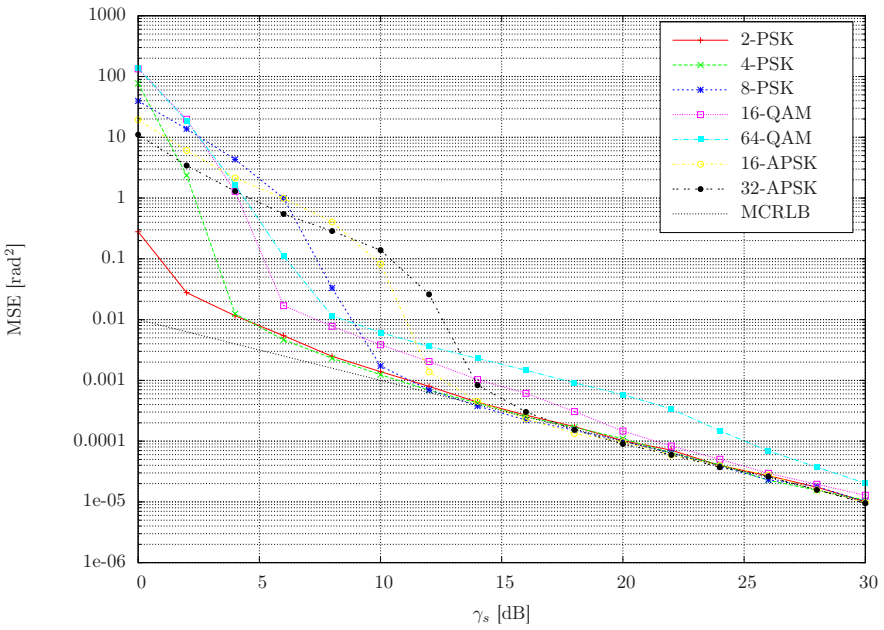


Figure 3.72: FB phase MSE for second-order loops

3.14 Remodulation

After successful execution of the preceding stages, all required parameters for extracting the transmitted symbols c_n are available. Thus, decisions $\hat{c}_n$ can be obtained by comparing the ideal signal constellations to the recovered symbols. For this reason the I/Q plane is divided into M decision regions depending on the specific modulation scheme. If a recovered symbol is located in decision region n, the symbol associated to this partition is assumed to be transmitted. The decision bounds are set to the mean of two adjacent symbol phases $\varphi_{d,i} = (\arg(c_i) + \arg(c_{i+1}))/2$ and additionally, in case of multi-level constellations, to adjacent symbol amplitude levels $r_{d,i} = (|c_i| + |c_{i+1}|)/2$. The obtained decisions can be used to form a remodulated signal $\hat{u}_k$ as

$$\hat{u}_k = \sqrt{\hat{E}_s} e^{\jmath(2\pi k \hat{f}_c T_s + \hat{\theta})} \sum_i \hat{c}_i \, \hat{h}((k/\hat{N}_s - i - \hat{\epsilon})\hat{T}) \tag{3.94}$$

It can be seen that, when assuming ideal synchronization, the above equation corresponds to a discrete version of the received single-carrier signal stated in Equation 2.2. This signal is of great benefit in several applications related to satellite monitoring, as will be shown in subsequent sections. For example, subtraction of $\hat{u}_k$ can be used to identify hidden interfering signals or to improve SNR relationships when calculating the cross-ambiguity

function (CAF).

For generation of the remodulated signal $\hat{u}_k$, at first the decisions $\hat{c}_n$ are zero-padded to achieve a fixed rate of oversampling, e.g. $N_s = 4$, to keep processing complexity low. Afterwards, the signal is multiplied by the scaling factor $\hat{a}_s = \sqrt{\hat{E}_s}$. The following MF adds the required timing offset $\hat{\epsilon}$ by modifying the impulse response appropriately. For timing offsets $\epsilon \geq 1/N_s$ the samples are shifted by multiples of $1/N_s$ and only the remaining offset is integrated into the MF response as $h(t - \hat{\tau})$. The filter output is rotated by $\hat{\theta}$ to take account for the phase offset. It should be mentioned that the m-fold ambiguity does not influence the remodulation process. This is mainly due to the fact that the signal is now rotated back by the estimated offset $\hat{\theta}$ which was corrected before decision making. In the following, the signal is resampled to achieve the desired symbol rate $\hat{f}_d$. Depending on the conversion rate either before or after the resampling an AAF is applied. Finally, the signal is shifted to the center frequency $\hat{f}_c$ by multiplication of a complex tone. If the above steps are successful, an estimate of the received single-carrier signal without noise is available.

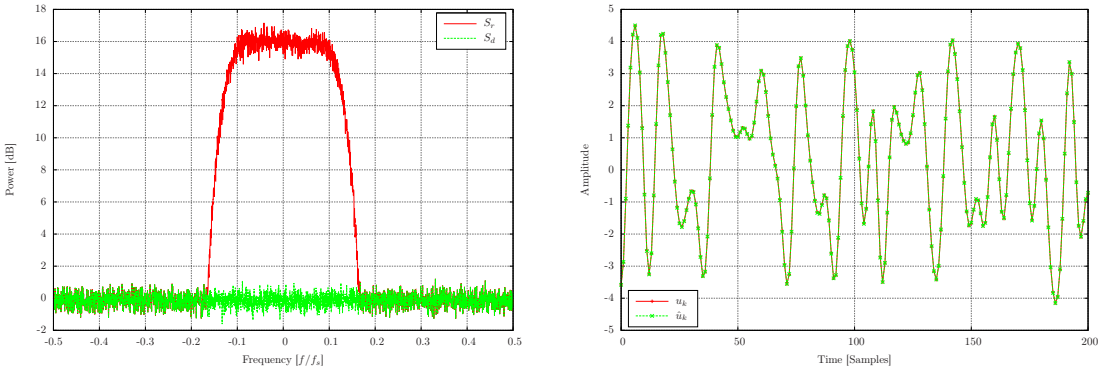
A way has to be found to evaluate the success of the remodulation stage. This can be done by comparing the powers of the corresponding signals. Thus, the power P_r of the received multi-carrier signal r_k subtracted by the power P_u of the remodulated signal $\hat{u}_k$ should be the power P_d of the difference signal $r_k - \hat{u}_k$. In the worst case, the power of the difference signal P_d is not smaller than that of r_k, but rather increased by the power of $\hat{u}_k$. A way to benchmark the remodulation is to define a threshold which will depend on the specific application. For example a criterion for successful remodulation could be stated as

$$P_d < (2P_r - P_u)/2 \tag{3.95}$$

This means that at least half of the expected power has to be subtracted to entitle the remodulation process as successful.

In the following, simulation results will show the subtraction of the remodulated from the received signal. In the first scenario a single 8-PSK carrier exhibiting an SNR of $\gamma_s = 16\,\mathrm{dB}$ is considered as received signal r_k. The associated power spectrum S_r is depicted in Figure 3.73a. The rate of oversampling is set to $N_s = 4$, the roll-off to $\alpha = 0.35$ and the center frequency to $f_c = 0$, to ease the inspection of the signal in the time domain. The detection and synchronization algorithms described within this chapter are applied to arrive at the remodulated signal $\hat{u}_k$. The latter is subtracted from r_k to arrive at $r_d = r_k - \hat{u}_k$ with power spectrum S_d. As can be seen in Figure 3.73a, the subtraction removes the signal component completely such that only the noise floor remains. By inspection of Figure 3.73b it can be seen that the original received signal u_k coincides with the remodulated signal $\hat{u}_k$.

In the next scenario it is tried to generate an example for erroneous remodulation. To generate remodulation errors, the SNR is lowered to $\gamma_s = 8\,\mathrm{dB}$. The according power spectrum S_d is shown in Figure 3.74. The signal component is only partly removed, so that a residual carrier remains. In this case the previously introduced criterion would identify the remodulation process as failed. For better understanding, in Figure 3.75 the



(a) Power spectra **(b)** Waveforms

Figure 3.73: Remodulation for 8-PSK at $\gamma_s = 16\,\mathrm{dB}$

ideal signal u_k and the remodulated version $\hat{u}_k$ are illustrated in the time domain. For demonstration purposes a large observation interval of $N = 524288$ samples was chosen for remodulation. It can be seen from the first 200 samples that, besides from some erroneous decisions, the remodulation procedure works as expected. However, the latter does not hold true for samples at the end of the observation interval, which in contrast show catastrophic behavior. This is mainly due to the fact that the estimate of the center frequency $\hat{f}_c$ exhibits a small error. Thus over a long period of time the remodulated constellation rotates away from the true one. This behavior can be mitigated by applying only chunks of appropriate size to the remodulation stage. An alternative solution could be a refinement of the frequency estimate by applying the accumulated phase error provided by the phase tracker.

The final scenario uses 2-PSK at an SNR of $\gamma_s = 3\,\mathrm{dB}$. The results of the successful remodulation are depicted in Figure 3.76. From the spectral illustration in Figure 3.76a it can be seen that the subtracted spectrum S_d exhibits peaks at the end of the transition band. This suggests that the estimate of the roll-off is too small. From the time domain representation in Figure 3.76b it becomes obvious that overshoots and wrong decisions lead to degradation of the remodulation process. Indeed, the increased fluctuation is caused by an error of the roll-off estimator, which delivers an estimate of $\hat{\alpha} = 0.15$ instead of $\alpha = 0.35$.

As a parameter to determine the quality of the remodulation process the EVM Λ can be used. The latter is usually specified in dB and defined as

$$\Lambda = \frac{\sum_{k=0}^{N-1} |u_k - \hat{u}_k|^2}{\sum_{k=0}^{N-1} |u_k|^2} \tag{3.96}$$

It should be pointed out that this definition does not coincide with the one in [66], relying

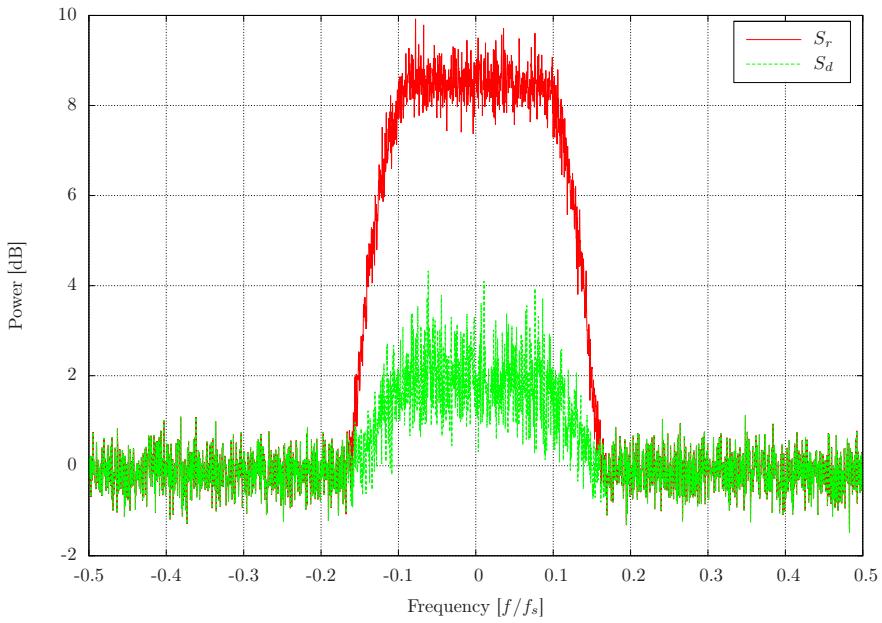


Figure 3.74: Remodulation power spectra for 8-PSK at $\gamma_s = 8\,\mathrm{dB}$

on symbols after the receiving filter, but is applied on the samples of the transmitted signal. This definition enables direct assessment of the remodulation procedure, however lacks an easy connection to the bit error rate (BER), as exemplified in [67]. For the first scenario the EVM results in $\Lambda_1 = -33.03\,\mathrm{dB}$, for the second in $\Lambda_2 = -0.71\,\mathrm{dB}$ and for the final in $\Lambda_3 = -9.28\,\mathrm{dB}$. Thus a significant increase of the EVM can be observed in case of erroneous remodulation. A detailed investigation of the achievable EVM by the blind demodulation framework is presented in the following section.

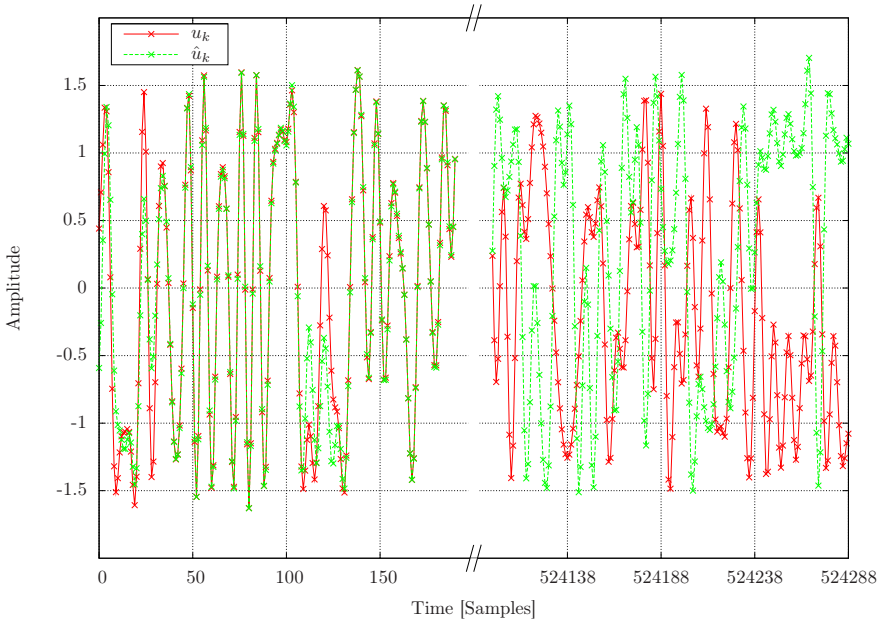
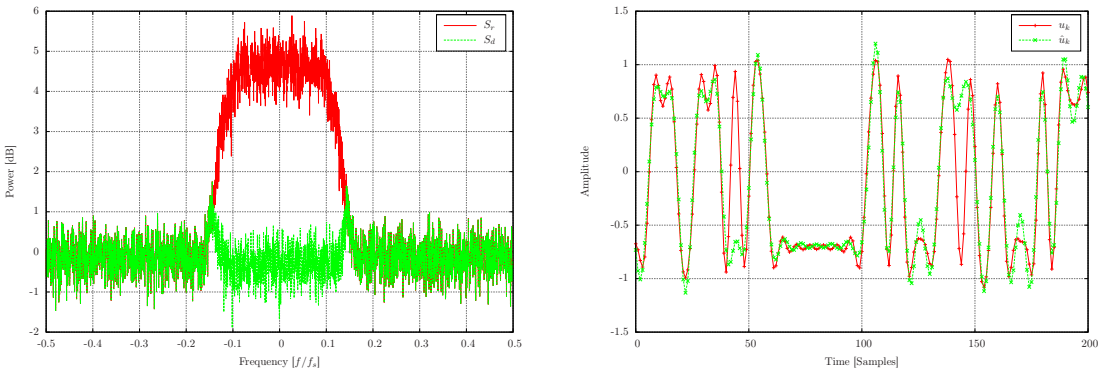


Figure 3.75: Remodulation waveforms for 8-PSK at $\gamma_s = 8\,\mathrm{dB}$



(a) Power spectra **(b)** Waveforms

Figure 3.76: Remodulation for 2-PSK at $\gamma_s = 3\,\mathrm{dB}$

3.15 Overall Performance

In this section simulations demonstrate the capabilities of the developed blind demodulation framework. It is clear that due to the unlimited number of possible input signals, representative scenarios have to be defined to verify the functionality. Moreover, too many parameters have to be determined such that it is unfeasible to present the estimation performance for all in a clear manner. So mainly parameters which were not discussed in detail in the previous sections are considered. The proper functioning of all estimation stages is verified by investigation of the EVM, since it provides an adequate measure.

3.15.1 Single-Carrier Scenario

Error Vector Magnitude As shown previously in Section 3.14, the EVM Λ is directly connected to the quality of the remodulated signal $\hat{u}_k$. So, the EVM can be used for a thorough test of the system, since a small EVM guarantees proper functioning of the previous blind demodulation framework. Of course the question of the EVM performance of the framework compared to a theoretical limit arises. The limit for the minimum achievable EVM in absence of forward error correction schemes can easily be computed. For this reason the received signal is assumed to be synchronized perfectly such that the transmitted symbols c_n remain, only impaired by Gaussian noise n_n. The decisions $\hat{c}_n$ made on this signal are then remodulated and the EVM is calculated. It is clear that, depending on the order of modulation, in the high SNR range the decisions $\hat{c}_n$ correspond almost to the true symbols c_n, resulting in an EVM of almost $\Lambda = -\infty$. However, with decreasing SNR the number of wrong decisions increases and the EVM increases as well. Figure 3.77 shows the evolution of the EVM when assuming ideal synchronization for M-PSK signals.

To assess the remodulation performance of the blind demodulation framework, a simulation is set up. For this scenario a single-carrier signal, oversampled by $N_s = 4$, is assumed. The received signal is impaired by time and phase offset and the center frequency of the carrier is chosen arbitrarily for each iteration. For each run the spectrum is scanned for carriers, the signal parameters of found carriers are pre-estimated and then the carriers are demodulated blindly. The decisions are remodulated and the EVM is calculated. Each simulation point relies on 100 iterations. To guarantee proper carrier detection, $M = 75$ blocks of length $N = 2^{16}$ overlapping by 75 % are processed for generation of a well-averaged periodogram. After successful parameter estimation a block of $N = 2^{14}$ samples is applied for remodulation. The EVM evolution of the proposed framework is also included in Figure 3.77. It can be seen that in the medium SNR range the remodulation performance is dominated by decision errors and thus very close to the ideal EVM. In the high SNR range the EVM deviates from the ideal curve by exhibiting an error floor. The latter is produced by small errors in the described parameter estimation process. Any deviations from the true symbol rate f_d, amplitude scaling factor a_s, center frequency f_c, timing offset ϵ and phase offset θ contribute to this error floor. For small SNR values the

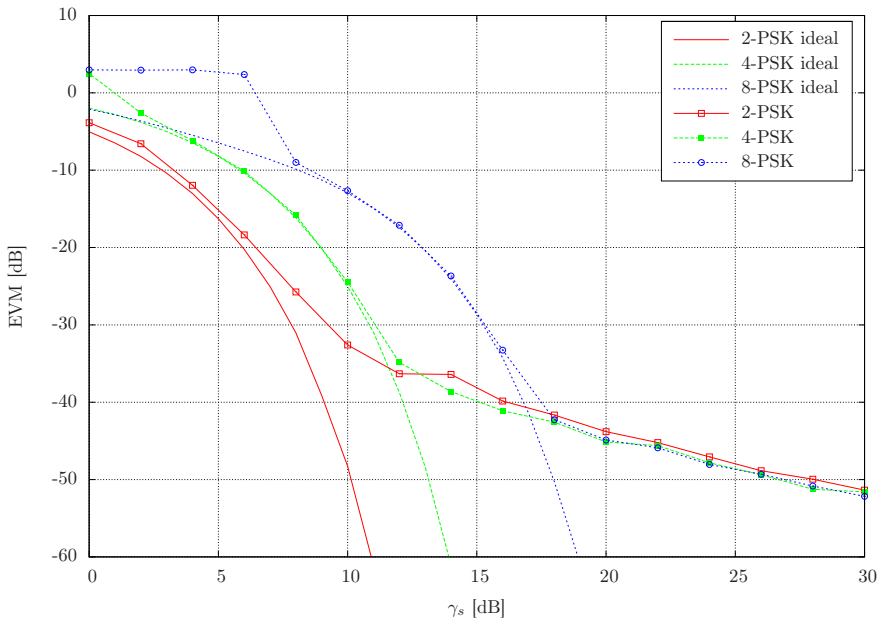


Figure 3.77: Evolution of EVM for M-PSK

EVM increases significantly. This threshold effect is produced by malfunction of the blind demodulation framework. The main impact is due to the frequency estimator degradation for higher order and multi-level modulation schemes, described in Section 3.12. Another cause is the SNR restriction applied for modulation classification introduced in Section 3.9. So, in the low SNR region the true modulation class may be eliminated by the restriction and thus a wrong modulation scheme is used for remodulation.

The same effects can be observed for M-QAM constellations as illustrated in Figure 3.78. The ideal curves are shifted in direction of higher SNR due to the increased probability of false decisions. The noise floor in the high SNR region is approximately increased by one order of magnitude. The threshold effect occurs at smaller SNR values compared to 8-PSK because of the solid frequency estimation performance of M-QAM.

Finally, Figure 3.79 shows the EVM performance for APSK signals. It is obvious that the threshold effect is pronounced most significantly due to the poor frequency estimator performance for APSK constellations. The noise floor remains at a similar level as for M-QAM.

Remodulation Success Rate The introduced criterion for successful remodulation in Equation 3.95 should be verified as well. The success rate according to the criterion is depicted in Figure 3.80 for all modulation schemes. It can be seen, that this loose definition for remodulation success matches the above EVM evolutions very well. Especially the

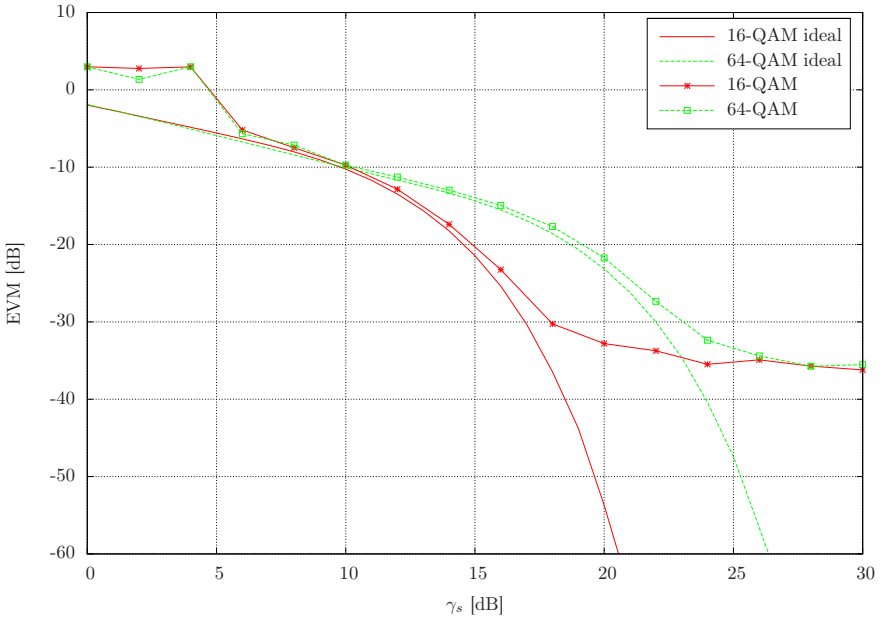


Figure 3.78: Evolution of EVM for M-QAM

EVM threshold effect can be observed together with a significant drop in the success rate. Thus the criterion can be used to judge whether remodulation was successful or not.

Interestingly, 16-QAM and 64-QAM exhibit the same success rate. The alert reader may have noticed this effect already for the EVM evolution in Figure 3.78. The reasons for this behavior are the used modulation classifier and the similarity of both signal constellations in terms of quadrature symmetry. To further investigate this effect, a simulation run is setup using 64-QAM at $\gamma_s = 8\,\text{dB}$. Basically, the modulation classifier excludes 64-QAM from the possible constellations for an SNR $\gamma_s \leq 8.45\,\text{dB}$ (see Table 3.4). Thus, the cumulant-based estimator decides for 16-QAM instead of 64-QAM. The blind demodulation framework tries to demodulate the signal and remodulate the decisions, assuming a wrong modulation scheme. The only parameter that indicates this failure is a significant drop in the second SNR estimate. The moment-based SNR estimator evaluates the SNR as $\hat{\gamma}_s = 6.74\,\text{dB}$, in contrast to $\hat{\gamma}_s^* = 8.00\,\text{dB}$ of the spectral pre-estimation stage. Then, the remodulated 16-QAM signal is subtracted from the received 64-QAM waveform. In Figure 3.81, the spectra of the involved waveforms are depicted. It is obvious that, although considering a wrong modulation scheme, the subtraction shows a significant signal reduction. Additionally, in Figure 3.82 the real part of the transmitted signal u_k is compared to the remodulated version $\hat{u}_k$. The remodulated signal $\hat{u}_k$ follows the original u_k, including a minor impact because of decision errors and offsets due to the coarser resolution of possible symbol amplitudes.

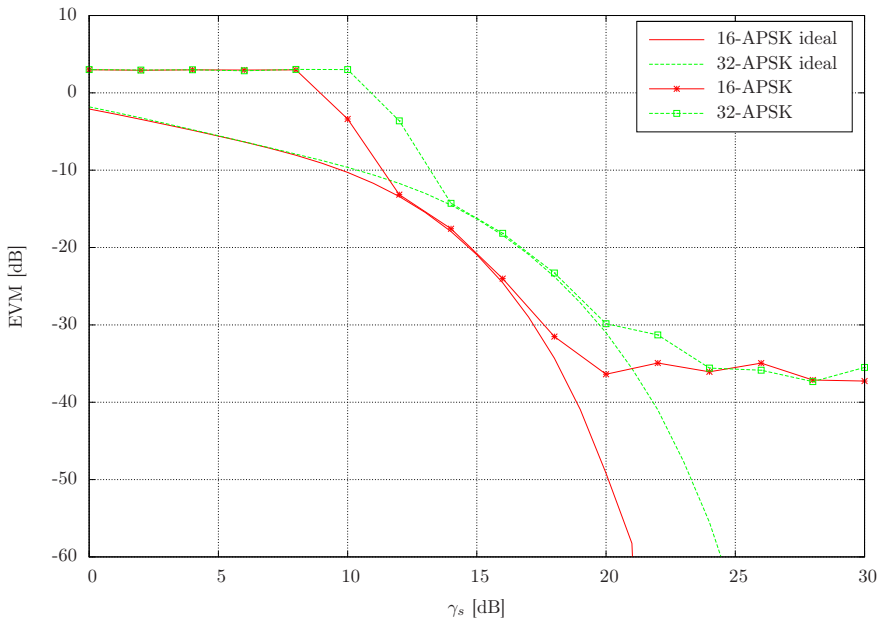


Figure 3.79: Evolution of EVM for M-APSK

SNR pre-estimate Since the SNR pre-estimate $\hat{\gamma}^*$ computed in the carrier detection stage (Section 3.1.1) is of major importance, the accuracy of the latter should be assessed. It finds application in the rescaling stage and is used to choose appropriate algorithms, e.g. SNR estimation, as well. For the simulation runs the previously set up single-carrier scenario is used. In Figure 3.83, the mean estimation error of the pre-estimate (P) is compared to the combined moment-based scheme (M) described in Section 3.10. It can be seen that for multi-level constellations the moment-based estimator exhibits a slight degradation in the medium and severe in the low SNR range. In contrast, the pre-estimate tends to be robust over the entire SNR region, independent of the modulation scheme. From Figure 3.84 it becomes obvious that the NMSE of the rough estimate is very solid and only increases slightly in the low-to-medium SNR range. In contrast, the moment-based method exhibits decreased NMSE in the medium-to-high SNR region. This is especially true for PSK signals. However, in the low SNR range the moment-based method shows inferior performance compared to the pre-estimate. The main reason for this is the increasingly erroneous blind demodulation. So, errors in timing estimation, modulation classification and roll-off estimation will lead to impairments on the symbols which are applied as input to the moment-based SNR estimator. The outlier for 16-QAM at $\gamma_s = 16\,\mathrm{dB}$ is caused by switching between the multi-level and partitioned SNR estimator. Increasing the estimator length L mitigates this effect. It can be summarized, that the pre-estimate $\hat{\gamma}^*$ is a robust and sufficiently accurate estimate of the SNR.

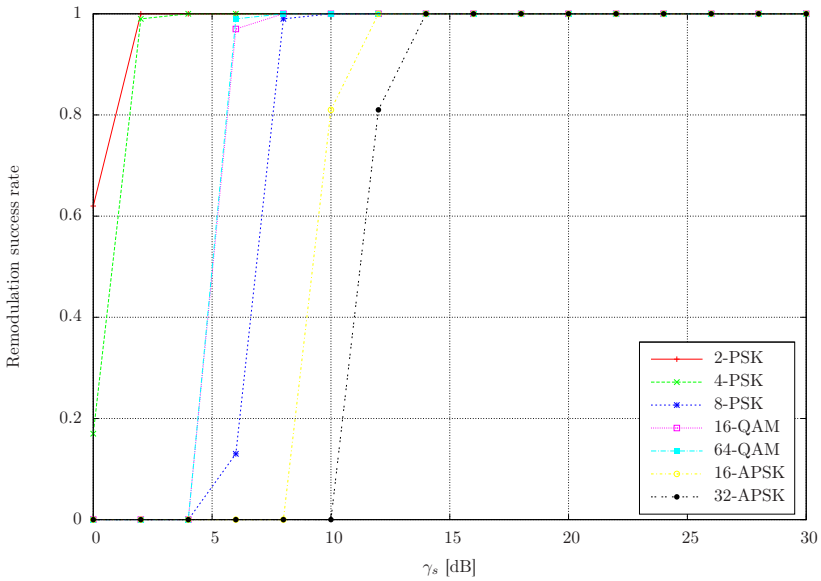


Figure 3.80: Remodulation success rate

3.15.2 Multi-Carrier Scenario

Up to now the performance of the blind demodulation framework was verified by using single-carrier scenarios. In the following it will be demonstrated that multi-carrier scenarios can be handled as well. As for the single-carrier scenario, the carrier detection stage tries to identify all possible carrier candidates within the observed bandwidth, however, with the difference that the number of carriers is greater than one. After successful detection, the blind demodulation process is executed sequentially for each carrier.

Error Vector Magnitude The simulation setup is the same as for the single-carrier setup, but instead of a single carrier, seven carriers are spaced consecutively in the bandwidth of interest. Each carrier features a different modulation scheme, such that all supported constellation types are represented. The oversampling factor is set to $N_s = 12$ for all carriers. The simulation starts at an SNR of $\gamma_{s,q} = 30\,\text{dB}$ and is reduced down to $\gamma_{s,q} = 0\,\text{dB}$. For each simulation point 100 iterations are performed. It should be mentioned that the detection algorithm exhibits solid performance, since no detection failure occurred over the simulated SNR range. The evolution of the EVM is depicted in Figure 3.85. The solid lines correspond to the ideal EVM limits, derived in the previous section. It can be seen, that the EVM performance coincides with the results from the single-carrier scenarios shown in Figures 3.77, 3.78 and 3.79.

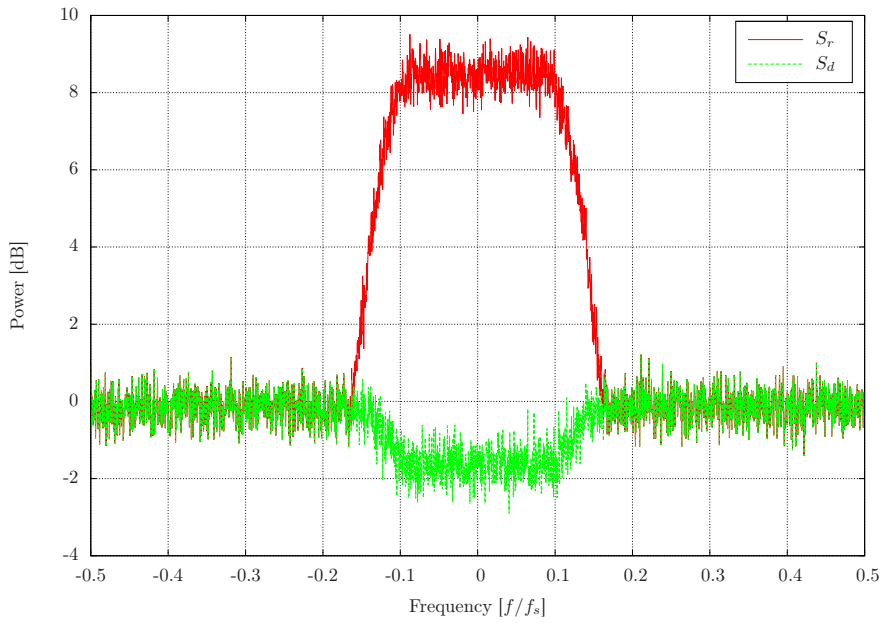


Figure 3.81: Remodulation power spectra for 64-QAM at $\gamma_s = 8\,\mathrm{dB}$

Remodulation Success Rate In Figure 3.86 the evolution of the remodulation success rate, according to the criterion defined in Equation 3.95, is illustrated. When comparing the success rate to the EVM evolution, it becomes obvious that the introduced criterion reliably identifies the threshold effect due to deficient synchronization capabilities.

Finally, the output of the remodulation stage is examined in the frequency domain. Figure 3.87 shows the spectrum S_r of the received signal at an SNR of $\gamma_{s,q} = 16\,\mathrm{dB}$. From left to right, the carriers are modulated using 2-PSK, 4-PSK, 8-PSK, 16-QAM, 64-QAM, 16-APSK and 32-APSK. Moreover, the spectra S_d that remain after subtraction of the remodulated signals $\hat{u}_{k,q}$ are displayed. It can be seen, that at such high SNR the carrier subtraction works successfully. For some carriers a peak at the edge of the transition band can be observed due to erroneous roll-off estimation. The fifth carrier from the left, which corresponds to the 64-QAM constellation, exhibits a decrease of the noise floor after subtraction.

When decreasing the SNR further down to $\gamma_{s,q} = 6\,\mathrm{dB}$, it is clear that the remodulation will produce erroneous results. The according spectral representation is shown in Figure 3.88. As expected from the EVM performance and the remodulation success rate, the remodulation procedure fails for 8-PSK, 16-APSK and 32-APSK. Both QAM carriers are subtracted and feature a decreased noise floor. Again it should be mentioned that the 64-QAM carrier is removed although 16-QAM is applied for remodulation. Still, flawless subtraction can be observed for 2-PSK and 4-PSK.

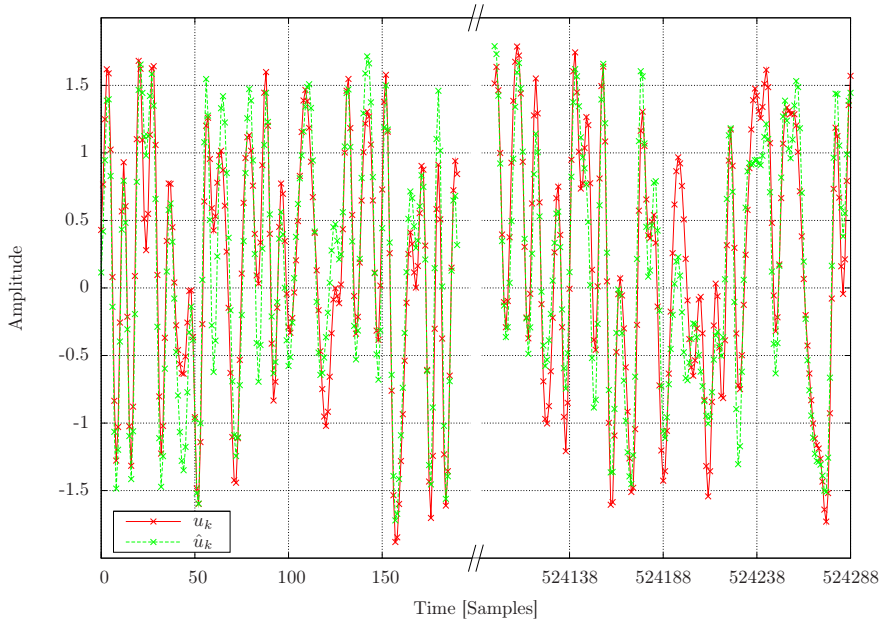


Figure 3.82: Remodulation waveforms for 64-QAM at $\gamma_s = 8\,\mathrm{dB}$

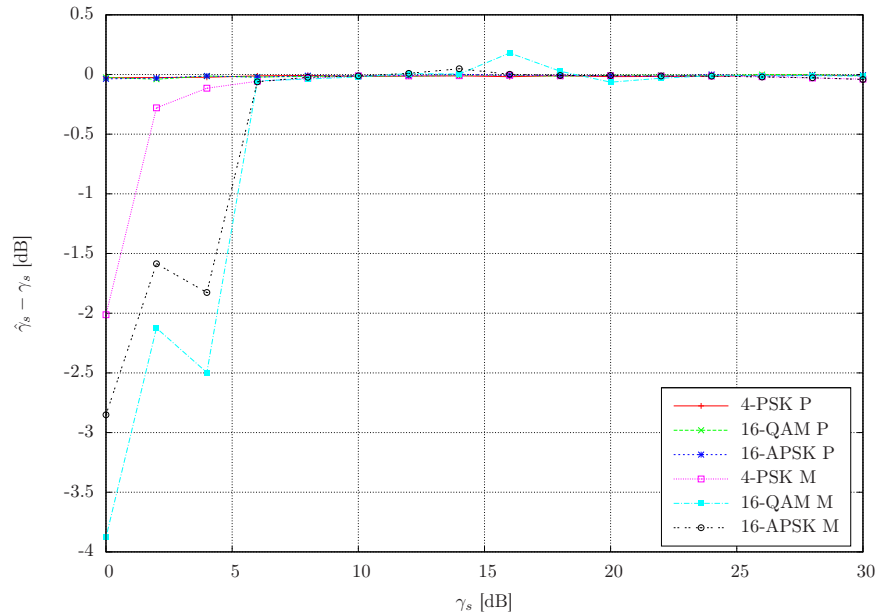


Figure 3.83: Mean SNR estimation error (spectrum- and moment-based)

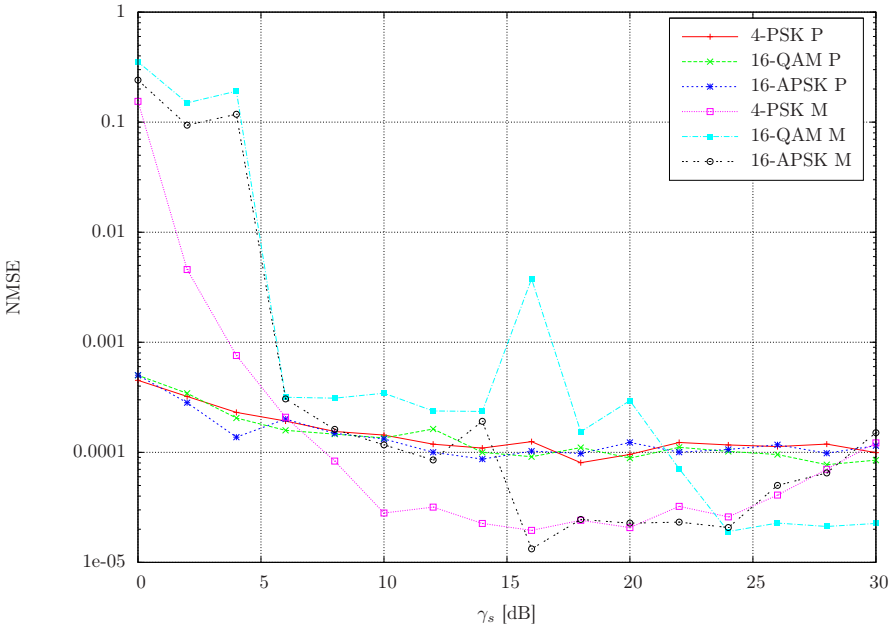


Figure 3.84: SNR NMSE (spectrum- and moment-based)

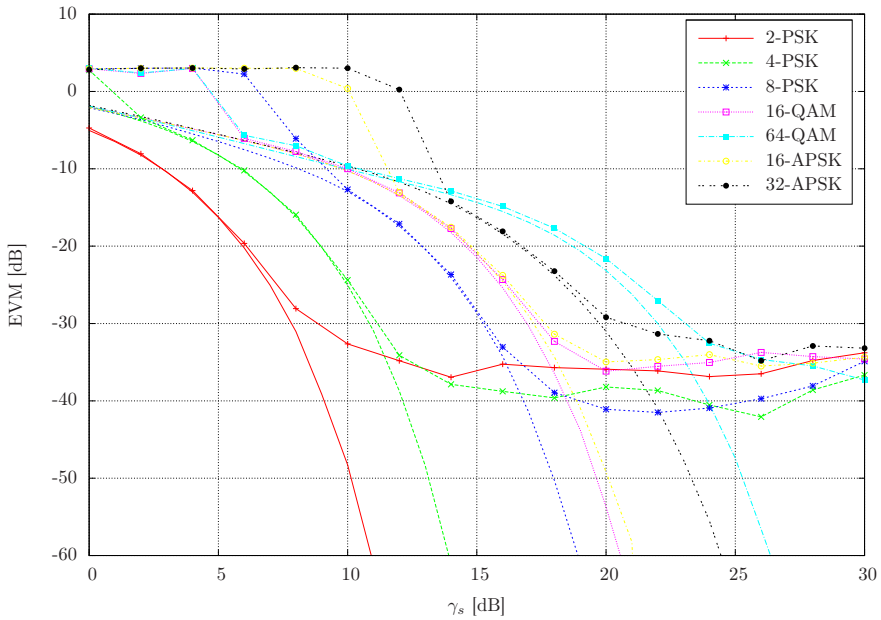


Figure 3.85: EVM evolution for multi-carrier scenario

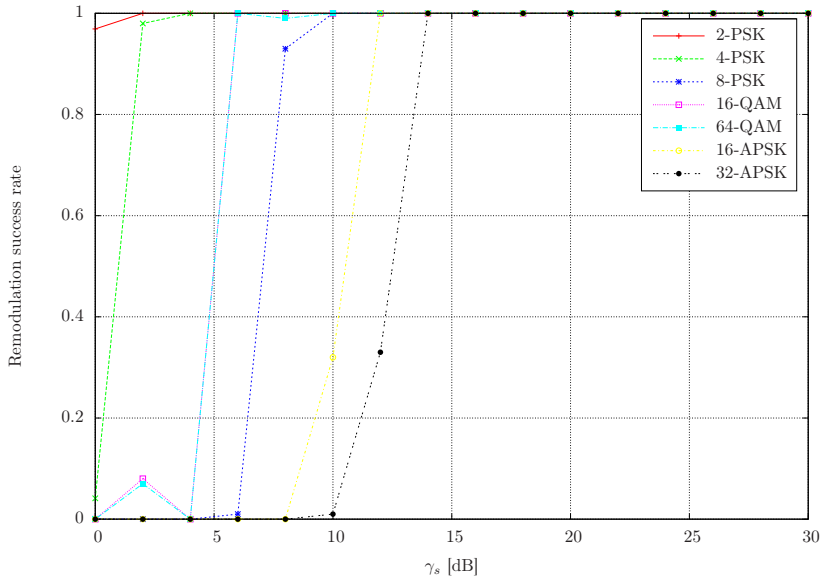


Figure 3.86: Remodulation success rate for multi-carrier scenario

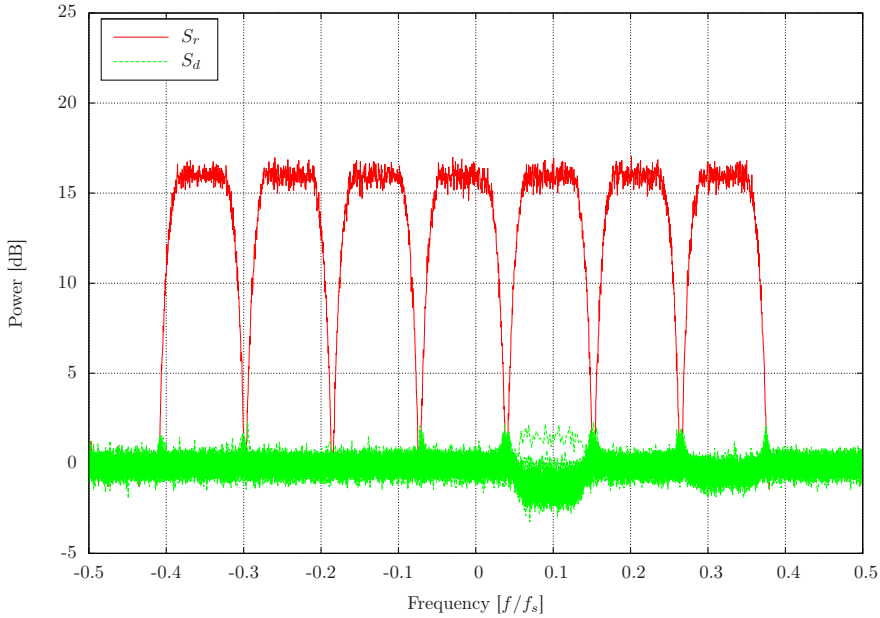


Figure 3.87: Remodulation power spectra at $\gamma_s = 16\,\mathrm{dB}$

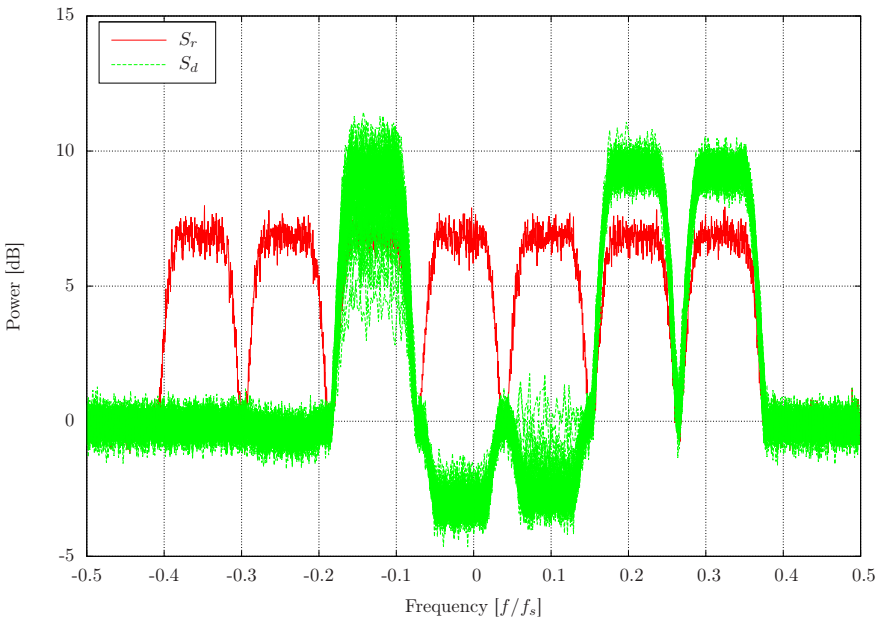


Figure 3.88: Remodulation power spectra at $\gamma_s = 6\,\mathrm{dB}$

CHAPTER 4

Applications

4.1 Cross-Polarization Discrimination Measurement

4.1.1 General Remarks

The polarization of an electro-magnetic wave is defined by the direction of the electric field component [1]. When using orthogonal polarizations, the same frequency can be used twice at the same location. Whereby orthogonal polarizations can either be horizontal/vertical or left-/right-hand circular. In Figure 4.1, the principle of using two orthogonal linear polarizations simultaneously is illustrated. At the transmitter, the electric field components $\mathbf{E}_1$ and $\mathbf{E}_2$ (in blue) are properly aligned in the vertical and horizontal direction. When assuming ideal conditions, the whole power transmitted in one polarization is received in the same polarization plane at the receiver. However, this behavior changes when transmitting and receiving antenna are not properly aligned. In this case, the field components $\mathbf{E}_1$ and $\mathbf{E}_2$ (in red) are received as $\mathbf{E}_{1,c}$ and $\mathbf{E}_{2,c}$ in the co-polar plane and as interference $\mathbf{E}_{1,x}$ and $\mathbf{E}_{2,x}$ in the cross-polar plane. When a signal is transmitted in one polarization plane, cross-polarization discrimination (XPD) is then defined as the ratio between the power received in this polarization to that received in the opposite (orthogonal) polarization. Typically, the XPD is expressed in units of dB and thus can be stated as

$$\mathrm{XPD}\,[\mathrm{dB}] = 20\log_{10}\left(\frac{\mathbf{E}_{1,c}}{\mathbf{E}_{1,x}}\right) \tag{4.1}$$

Unfortunately, misalignment of the transmitting antenna is not the only reason for degradation of the XPD. Antennas which can operate two different polarizations exhibit finite isolation between both planes. So it is clear that there will be interference between both polarization planes. The imperfection of dual polarization antennas is tolerable due to the advantage that a single antenna can handle twice the bandwidth. To mitigate the effects due to finite isolation, the transponders of horizontal and vertical polarization are offset to each other in frequency, as sketched in Figure 4.2.

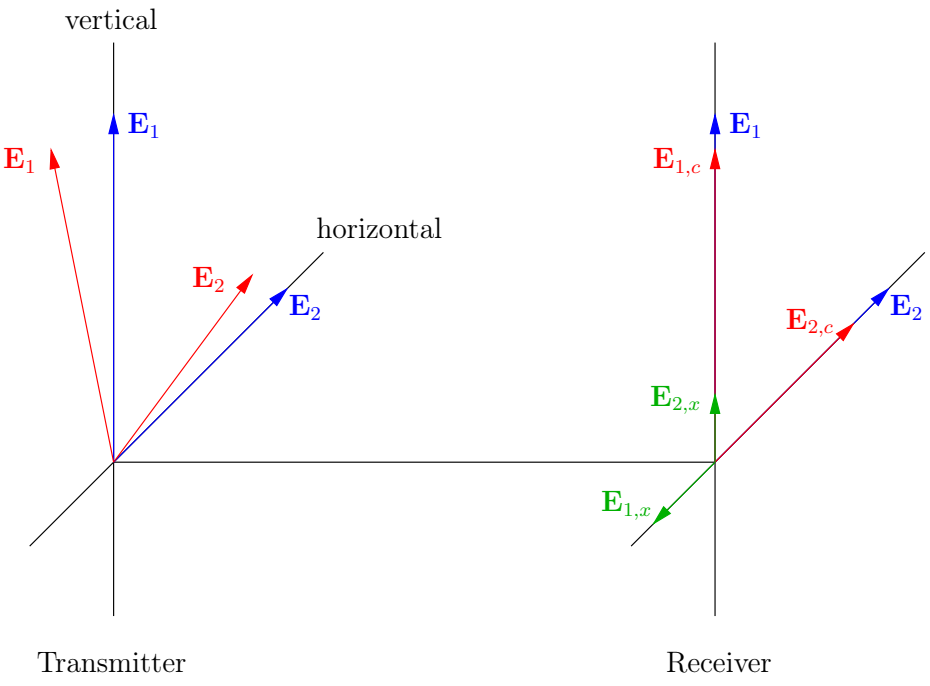


Figure 4.1: Orthogonal linear polarizations

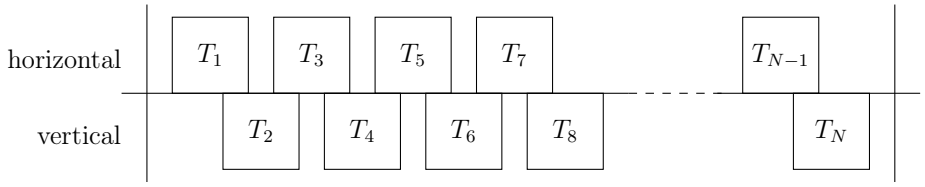


Figure 4.2: Frequency plan for orthogonal polarization

Moreover, propagation effects, such as precipitation or Faraday rotation, can lead to decreased XPD by depolarization of the transmitted wave. However, it is assumed that these effects are small compared to the impairment by antenna alignment.

Before a satellite terminal can be put into service, a line-up procedure has to be performed. By this means, it should be guaranteed that no interference occurs for the present traffic. Typically, a continuous wave signal has to be transmitted which is inspected by the provider control station. To avoid possible interference on the opposite polarization plane, the transmitting antenna is adjusted until maximum XPD is achieved. After successful line-up, the terminal can be put into operation using the assigned frequency slot. From this point on it is not possible to repeat the above XPD measurement in an

easy way with state-of-the-art techniques. The latter would again require the transmission of a continuous wave signal, to re-run the described measurement procedure. It is obvious that the insertion of such a signal would deteriorate the active traffic and thus requires a stop of service. According to Eutelsat, such a measurement procedure has to be performed once a week.

The proposed correlation approach, implemented on a software-defined radio (SDR) platform, can perform the XPD measurement online without interruption of the active services. A ground terminal, which can acquire signals from both polarization planes simultaneously is a prerequisite for applying the proposed methods. For this purpose, a dual polarization receive antenna with subsequent orthogonal mode transducer (OMT) is required. The latter splits up the received signals according to their polarization, such that co- and cross-polar signal are available on separate signal paths.

4.1.2 Proposed Estimator

Coarse Method

In absence of noise, an intuitive approach for estimation of the XPD, or the inverse thereof, can be formulated. The inverse corresponds to the factor a_c, by which the co-polar signal is scaled to arrive at the cross-talked signal. At first, the auto-correlation $R_{l,cc}$ of the co-polar signal $r_{k,c}$ is calculated. The maximum of the auto-correlation $R_{m,cc}$ corresponds to the co-polar signal power P_c. Afterwards, the cross-correlation $R_{l,cx}$ of co- and cross-polar signal, $r_{k,c}$ and $r_{k,x}$, is computed, whereby the maximum $R_{m,cx}$ delivers the cross-talked power $a_c P_c$. Auto- and cross-correlation are calculated according to [68]

$$R_{l,cc} = \sum_{k=0}^{L-1} r_{k,c}^{*} r_{k+l,c} \tag{4.2}$$

$$R_{l,cx} = \sum_{k=0}^{L-1} r_{k,c}^{*} r_{k+l,x} \tag{4.3}$$

Herein L denotes the number of processed samples. The time lag l is set to $l = 0$, for the auto-correlation, whereas for cross-correlation it is set to a value, depending on the expected timing offset between the co- and cross-polar signals. Since the time lag can only be selected in units of samples, interpolation is applied to receive a refined estimate of the cross-correlation maximum (see Appendix A.1). The ratio of the maxima of cross- and auto-correlation corresponds to the estimate of the scaling factor a_c:

$$\hat{a}_c = \frac{R_{m,cx}}{R_{m,cc}} \approx \frac{a_c P_c}{P_c} = \frac{1}{\mathrm{XPD}} \tag{4.4}$$

As initially defined, the inverse of $\hat{a}_c$ delivers the estimate of the XPD. The attenuation factor in dB is denoted by ξ_c and obviously given by

$$\hat{\xi}_c = 20 \log_{10}(\hat{a}_c) \tag{4.5}$$

Unfortunately, this method will only work properly in absence of noise, which is, of course, not the case in practice.

Considering again the above estimation algorithm, the impact of noise becomes obvious. The problem is that the auto-correlation $R_{l,cc}$ calculates the power of the received signal $r_{k,c}$, which actually includes signal power P_c plus noise power P_{nc}, i.e. $R_{m,cc} = P_c + P_{nc}$. In contrast, the cross-correlation $R_{l,cx}$ delivers the power of the signal component common to both polarization planes. As shown, the latter can be formulated as a scaled version of the co-polar signal power P_c, i.e. $R_{m,cx} = a_c P_c$. It should be added that this does only hold true, if the noise contributions of the horizontal and vertical polarization are independent. Finally, the correlation ratio develops to

$$\hat{a}_c = \frac{R_{m,cx}}{R_{m,cc}} \approx a_c \frac{P_c}{P_c + P_{nc}} \tag{4.6}$$

So it should be clear that, with increasing noise, the estimate $\hat{a}_c$ will be increasingly biased as well. For this reason, the introduced method can be applied in the high SNR range reliably, otherwise it only serves as a coarse estimate.

Typically, the mixers for the transponders on opposite polarization planes onboard the satellite and the mixers in the monitoring station, derive their frequency from the same local oscillator. Thus, it can be assumed that the frequency offset between co- and cross-polar signal is negligible. The presence of a frequency offset would degrade the correlation performance. In such a case, standard correlation has to be replaced by calculation of the cross-ambiguity function (CAF) [69], as will be discussed very detailed in Section 4.2.2. The CAF exhibits a maximum at the corresponding frequency and time offset, which will be used instead of the above cross-correlation peak $R_{m,cx}$.

The described correlation algorithm is implemented in such a way that the carrier under consideration is converted down to baseband and filtered according to the carrier bandwidth b_c. If possible, the signal is decimated appropriately to reduce computational complexity of the following correlation stage.

Effective SNR The required number of samples L that has to be processed for generation of a proper correlation peak depends on the actual scenario. According to [70], the effective SNR $\gamma_{s,o}$ of the correlation peak can be evaluated by

$$\gamma_{s,o} = 2BT_L \frac{\gamma_{s,c}\gamma_{s,x}}{\gamma_{s,c} + \gamma_{s,x} + 1} \tag{4.7}$$

where $\gamma_{s,c}$ denotes the SNR of the co-polar carrier and $\gamma_{s,x}$ of the cross-talk signal. The latter depends on the XPD and the noise floor of the cross-polar plane, whereby any cross-polar traffic can be viewed as additional noise contribution. The factor $2BT_L$ is termed as processing gain of the correlation and consists of the noise bandwidth B and the observation interval T_L. As stated in [70], an effective SNR of approximately 20 dB is necessary for unique identification of the correlation peak. Hence, the above equation allows computation of the required number of samples for producing a significant correlation peak according to the specific scenario.

Refined Method

The result of the previously described algorithm is mainly impaired by the contribution of noise in the auto-correlation $R_{l,cc}$. An obvious solution for this problem, is the correction of $R_{m,cc}$ by the noise power P_{nc}. An estimate for the noise power over the entire co-polar bandwidth is computed during the carrier detection procedure, described in Section 3.1.1. The latter can be used to derive P_{nc}, the noise power inside the bandwidth b_c. By this means, a refined estimate for the attenuation factor can be calculated according to

$$\hat{\xi}_c' = 20\log_{10}\left(\frac{R_{m,cx}}{R_{m,cc} - \hat{P}_{nc}}\right) \tag{4.8}$$

Simulation Results In the following, the performance of the coarse and refined method will be compared by simulation runs. For this reason, the knowledge of all relevant parameters required for XPD estimation is available. In the first scenario, a single QPSK carrier is generated. The attenuation from the co- to cross-polar plane is set to $\xi_c = -25$ dB. For both methods, the mean estimator output versus the SNR is depicted in Figure 4.3. It can be seen that the coarse method (C) increasingly deviates from the true value with smaller SNR values. In contrast, the refined algorithm (R), applying the noise power estimate $\hat{P}_{nc}$, delivers an output close to the true value over the whole SNR range. When assuming an oversampling rate of $N_s = 4$ and an estimator length of $L = 262144$ samples, the processing gain becomes 51.18 dB. This results in an effective SNR $\gamma_{s,o} = 23.16$ dB at an SNR $\gamma_{s,c} = 0$ dB and $\gamma_{s,o} = 56.16$ dB for $\gamma_{s,c} = 30$ dB. Hence, the estimator length is sufficient to guarantee proper peak estimation over the whole SNR range. As mentioned above, any traffic on the opposite polarization can be considered as further noise contribution. So, when adding cross-polar traffic with an SNR $\gamma_{s,x} = 15$ dB, the effective SNR evaluates as $\gamma_{s,o} = 8.16$ dB for an SNR $\gamma_{s,c} = 0$ dB and as $\gamma_{s,o} = 41.17$ dB for $\gamma_{s,c} = 30$ dB. The reduced effective SNR will lead to increased jitter of the correlation peak. In the low SNR range the post correlation SNR becomes even significantly smaller than the suggested limit $\gamma_{s,l}$. From the evolution of the mean estimator output in Figure 4.3 for both, the coarse (CT) and refined (RT) method, a tendency to increased error can be observed with decreasing SNR. The augmented jitter, due to the changed SNR relationship, is recognized best from the NMSE evolution, depicted in Figure 4.4. Increasing the estimator

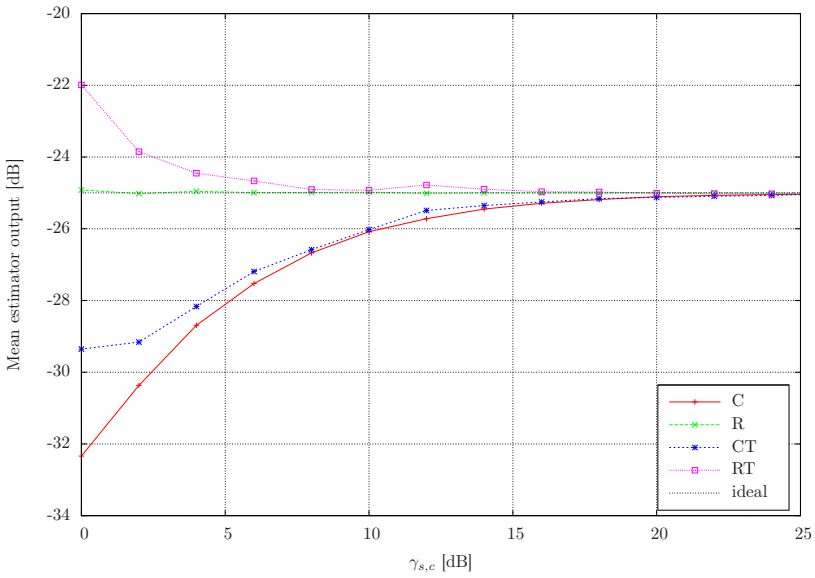


Figure 4.3: Coarse/refined mean XPD estimator output

length will increase the post correlation SNR and thus mitigate the degradation.

Finally, the mean error of the estimator output versus the attenuation factor ξ_c is investigated. The SNRs of the co- and cross-polar traffic, are set to $\gamma_{s,c} = \gamma_{s,x} = 15\,\text{dB}$. This results in an effective SNR $\gamma_{s,o} = 55.64\,\text{dB}$ for an attenuation $\xi_c = -10\,\text{dB}$ and $\gamma_{s,o} = 36.02\,\text{dB}$ for $\xi_c = -30\,\text{dB}$ in absence of cross-polar traffic. In presence of the latter, it amounts to $\gamma_{s,o} = 41.03\,\text{dB}$ and $\gamma_{s,o} = 21.04\,\text{dB}$ respectively. From the simulation result, illustrated in Figure 4.5, it can be seen that the refined method (R, RT) delivers results as expected and the coarse approach (C, CT) exhibits a bias depending on the co-polar noise power. Moreover, in presence of cross-polar traffic, both methods (CT and RT) exhibit an increased jitter over the whole simulated XPD range.

Fine Method

As discussed above, the systematic bias, depending on the noise inherent to the co-polar signal, can be mitigated by using an estimate of the noise power. Moreover, it was shown that this refined method produces satisfying results at low complexity. So the question may arise, why to develop an additional algorithm? The problem is that up to now only cross-talk from the co- to cross-polar side was assumed. Thus any influence from the cross- to co-polar side is neglected. When considering this influence as well, the correlation ratio has to be completed as

$$\frac{R_{m,cx}}{R_{m,cc}} \approx \frac{a_c P_c + a_x P_x}{P_c + P_{nc} + a_x^2 P_x} \tag{4.9}$$

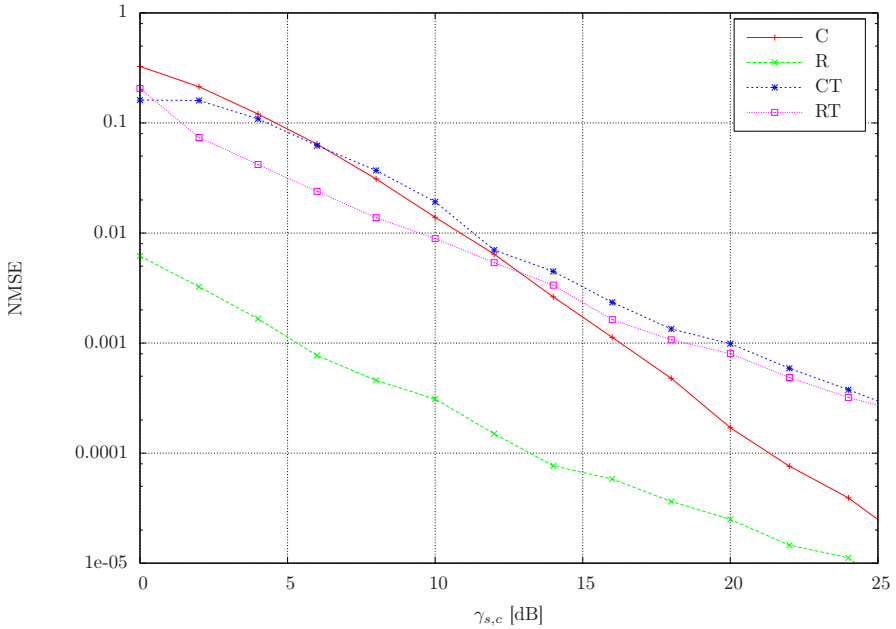


Figure 4.4: Coarse/refined XPD estimator NMSE

Auto- and cross-correlation are extended by scaled versions of the power P_x of the cross-polar signal, whereby a_x considers the amplitude scaling of the cross-polar signal. As can be seen, the extraction of the parameter of interest, i.e. a_c, becomes an intricate problem. The above correlation ratio will only deliver the true value in the absence of a cross-polar signal and co-polar noise.

To find a way out of this dilemma, the benefits of the introduced blind demodulation framework can be brought into play. As demonstrated in Section 3.14, it is possible to generate a remodulated version $\hat{u}_k$ of the received carrier r_k. Assuming perfect remodulation, a version of the received carrier, free of noise and impairments from the opposite polarization plane is available. The obtained signal $\hat{u}_k$ can be used to compute the maximum of the auto-correlation $R_{m,uu} = P_u$ and cross-correlation $R_{m,ux} = a_c P_u$, resulting in a correlation ratio of

$$\frac{R_{m,ux}}{R_{m,uu}} \approx \frac{a_c P_u}{P_u} = a_c \tag{4.10}$$

Herein P_u denotes the power of the remodulated signal. The obtained ratio can be used

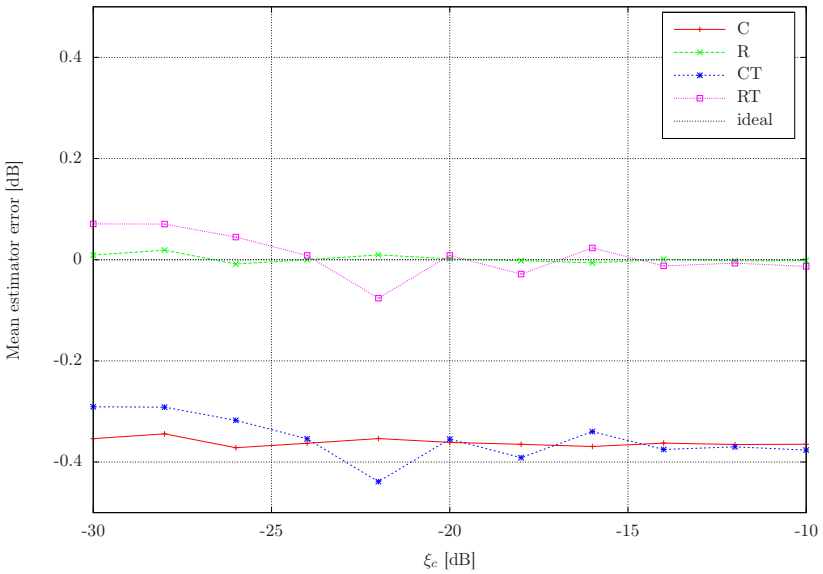


Figure 4.5: Coarse/refined mean XPD error versus attenuation

to generate a fine estimate of the attenuation factor ξ_c by

$$\hat{\xi}_c'' = 20\log_{10}\left(\frac{R_{l,ux}}{R_{l,uu}}\right) \tag{4.11}$$

It is clear that erroneous remodulation will lead to degradation of the proposed algorithm. In the following, the performance of the introduced methods, considering the complete model, will be verified by simulation runs. For this reason, a cross-polar carrier with an SNR of $\gamma_{s,x} = 15\,\text{dB}$ is generated, however, with the difference that the isolation between cross- to co-polar plane is finite, i.e. $\xi_x = -20\,\text{dB}$. The overspill from co- to cross-polar side is still assumed to be $\xi_c = -25\,\text{dB}$. Figure 4.6 shows the evolution of the mean estimator output for the coarse (CX), refined (RX) and fine (FX) algorithm. Now it becomes obvious that neither the coarse nor the refined approach work properly in the presence of significant influence from the opposite polarization. In contrast to this, the fine scheme (FX) shows solid performance over the entire SNR range. Up to now, ideal synchronization and prior knowledge of all parameters were assumed. The performance of the fine algorithm, relying solely on the estimates of the proposed blind demodulation framework (BX), is illustrated as well. The output of the ideally (FX) and blindly synchronized (BX) estimator coincides in the medium-to-high SNR range. As expected, in the low SNR region the blind scheme increasingly deviates from the ideal curve. This is mainly due to erroneous remodulation, which is additionally affected by the overspilled signal. The remodulation performance

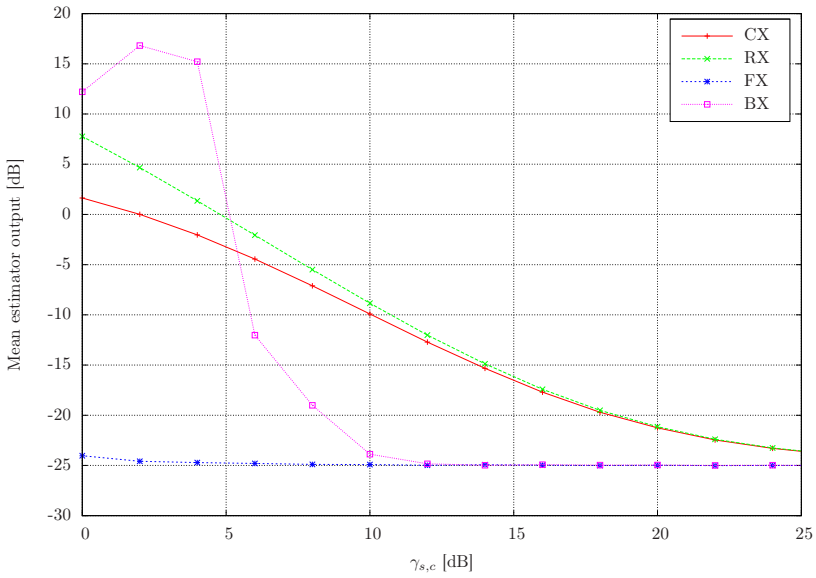


Figure 4.6: Mean XPD estimator output considering the complete model

by means of EVM is investigated thoroughly in Section 3.15. The observed behavior is confirmed by the evolution of the NMSE, as shown in Figure 4.7. Finally, the evolution of the mean estimation error versus the co-polar attenuation ξ_c is depicted in Figure 4.8. The simulation setup is the same as before with a co-polar SNR of $\gamma_{s,c} = 15\,\text{dB}$. For this scenario, the error for the coarse and refined method is in the range from 2 dB up to 12 dB. So the additional effort for remodulation seems to be justified without doubts.

Another capability of the blind demodulation framework could be used to improve the fine method. As previously stated, the post correlation SNR $\gamma_{s,o}$ depends on the specific SNR relationships. In presence of cross-polar traffic, remodulation may be applied to remove the according carriers from the opposite polarization plane. Successful subtraction will result in an increased effective SNR $\gamma_{s,o}$ and thus in less samples required for XPD estimation or improved accuracy. However, from Equation 4.9 it becomes obvious that it is not possible to remove the entire impact of the cross-polar signal. By applying subtraction, solely the term $a_x P_x$ vanishes in the cross-correlation. Unfortunately, the term $a_x^2 P_x$, considering the cross-talked signal in the auto-correlation, will still affect the estimate. Figure 4.9 compares the evolution of the mean estimator output for the refined (RX), fine (FX) and blind fine (BX) method to their counterparts (RS, FS and BS), additionally performing subtraction of the cross-polar carrier. It can be seen that the refined method benefits mostly from the traffic removal, but is biased by the above mentioned factor $a_x^2 P_x$ in the low SNR range. Due to the increased effective SNR, an improvement can be observed for the fine as well as for the blind fine method. Especially for the blind algorithm, the

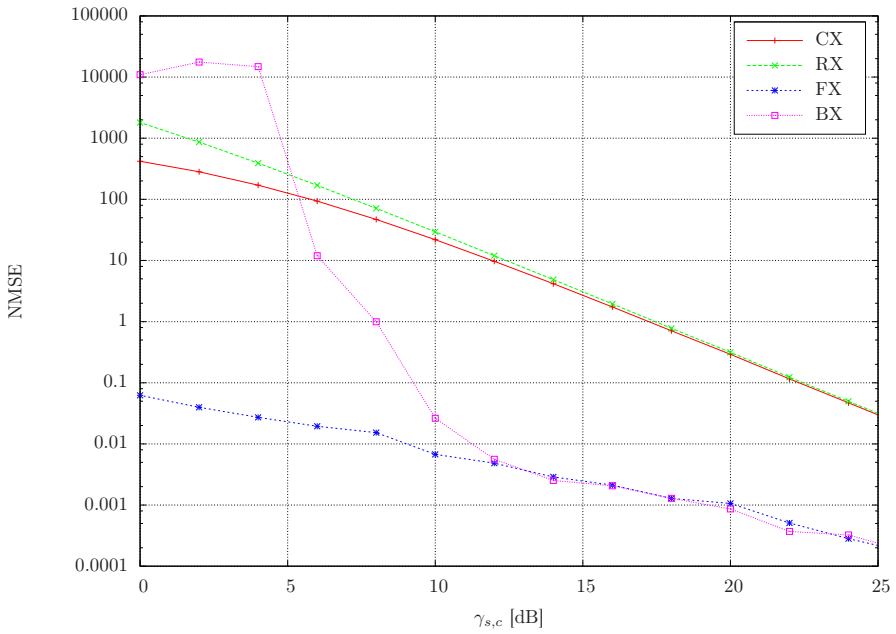


Figure 4.7: XPD NMSE considering the complete model

degradation due to malfunction of the remodulation stage can be shifted towards smaller SNR values. The gain in accuracy is best verified by the evolution of the NMSE, which is illustrated in Figure 4.10. Besides the considerable improvement of the refined method, the NMSE is decreased by more than one order of magnitude for both, the fine and the blind scheme. Finally, the mean error is plotted versus the attenuation ξ_c in Figure 4.11. In this case, the co-polar SNR is assumed to be $\gamma_{s,c} = 15\,\mathrm{dB}$. It can be seen that in this scenario the subtraction leads to acceptable performance for all methods over the entire simulated range. The only exception is an outlier for an attenuation of $\xi_c = -10\,\mathrm{dB}$. The reason for this, is the strong overspill of the co-polar signal to the cross-polar plane. As a result, the affected cross-polar signal is not remodulated properly and thus the benefit by subtraction can not be utilized.

Since the performance of the fine method is strongly affected by the quality of the remodulated signal, alternative approaches functioning properly for small SNR values are desirable. The parameters that form Equation 4.9 could be determined in a different way. As described, the co-polar noise power P_{nc} is estimated by using the carrier detector. Additionally, the signal powers P_c and P_x could be delivered by the rescaling stage, described in Section 3.7. Hence, after calculation of auto- and cross-correlation, two equations are obtained, containing the two unknowns a_c and a_x. The proposed method might work satisfactorily for single carriers on the cross-polar plane. Unfortunately, in presence of multiple cross-polar carriers, exhibiting different XPD values $\xi_{x,i}$, the problem

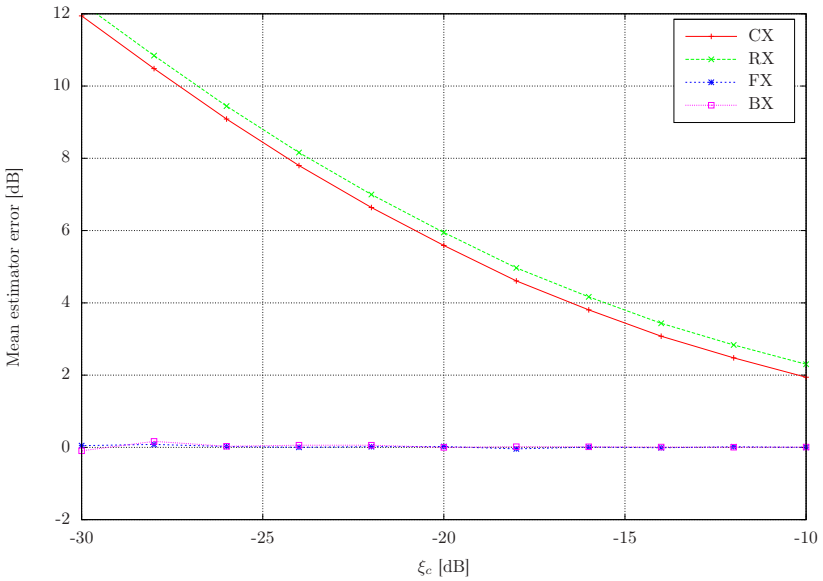


Figure 4.8: Mean XPD error versus attenuation considering the complete model

becomes complicated as well.

4.1.3 Hardware Testing

An in-house developed SDR platform is available for testing the introduced algorithms. The application of this platform enables the verification of the simulation framework and search of errors in the assumed models. Moreover, the devised algorithms can be tested with real signals, additionally impaired by hardware effects, e.g. quantization and phase noise.

Demonstrator Platform

In the following, the used hardware, illustrated in Figure 4.12, will be described briefly. The SDR platform features a sample rate of $f_s = 192\,\text{MHz}$ at a resolution of 12 bit and a bandwidth of $b_o = 36\,\text{MHz}$. All analog components, including power supply, are separated in an external box for electromagnetic compatibility (EMC) reasons. The digitized data are provided to the Gidel PCI card via LVDS. The latter includes a field-programmable gate array (FPGA) (Altera Stratix) and 512 MB of memory for storing the acquired data before transmission over the PCI bus. Inside the FPGA, the samples are decimated by a factor of 4 to arrive at a sample rate of $f_s = 48\,\text{MHz}$. After storing the samples in the RAM of the PC, further processing is performed by the introduced blind demodulation framework implemented in C++ by two Intel Xeon processors. In Figure 4.13, an image

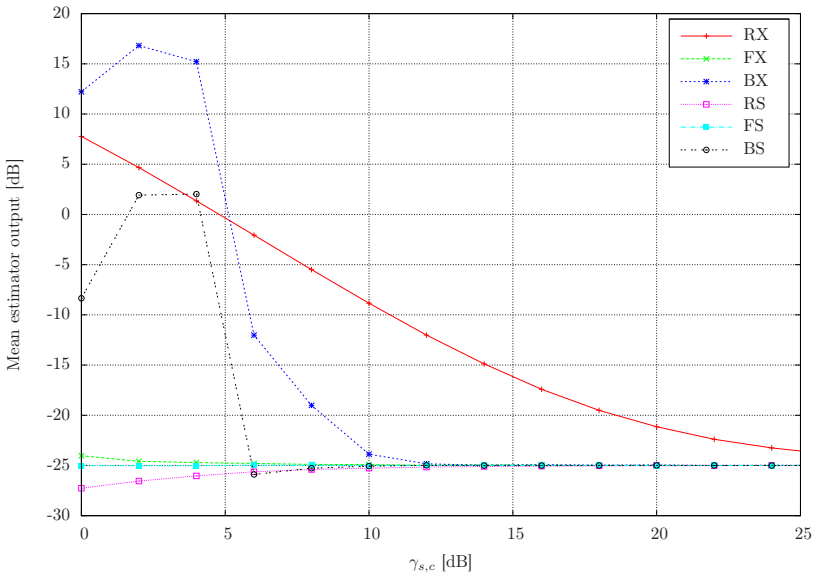


Figure 4.9: Mean XPD estimator output using cross-polar subtraction

of the demonstrator front side is shown. The two sampling units for co- and cross-polar polarization can be seen on top of the industrial PC case. The backside, depicted in Figure 4.14, shows the LVDS connections and the cable for synchronization of the analog-to-digital converter (ADC) clocks.

Testbed

To be able to perform testing of the algorithms, similar to the pure software environment, the testbed sketched in Figure 4.15 was developed. The co- and cross-polar input signals, at an intermediate frequency (IF) $f_{IF} = 70\,\text{MHz}$, are generated by an in-house developed satellite modem (L*IP) and a vector signal generator (Agilent ESG E4438C). The waveforms for the latter are created using the ADS software and downloaded to the generator. By this means arbitrary multi-carrier scenarios can be realized. A switch matrix consisting of attenuators (by Mini-Circuits) permits the adjustment of cross-talk signals. The noise for the co-polar path is produced by a noise generator (Micronetics Nod 5107) and for the cross-polar plane by a self-developed satellite emulator, which is capable of adding time and frequency offset as well. Finally, both signal paths are digitized using the synchronized sampling units. It should be mentioned that the satellite modem, the satellite emulator and the monitoring platform are based on the same flexible SDR platform. Besides the different software, only slight modifications in the hardware are required to allow this reuse.

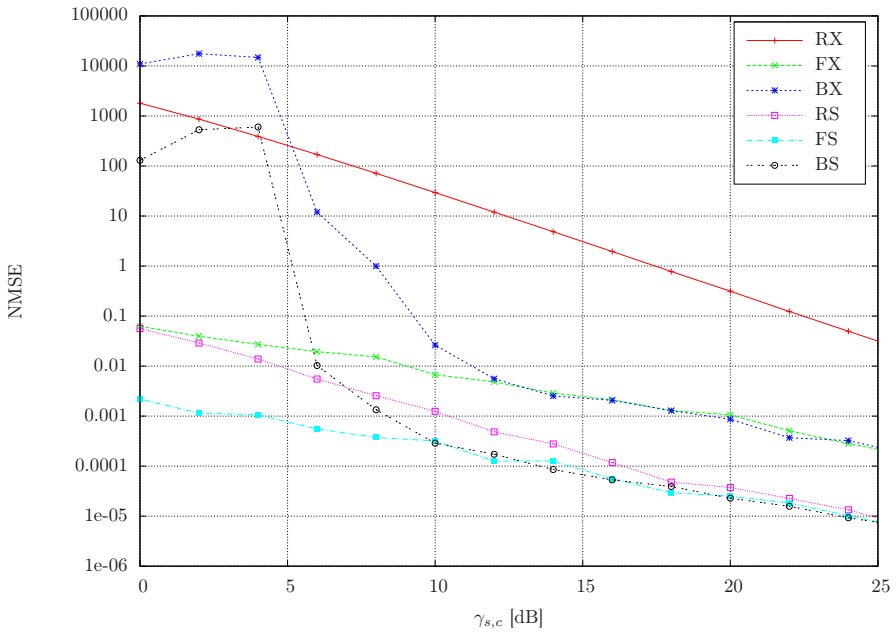


Figure 4.10: XPD NMSE using cross-polar subtraction

Measurement

For the following scenarios, the co-polar traffic consists of a single 1 MSymb/s QPSK carrier, produced by the satellite modem. On the other signal path, four carriers with different symbol rates are generated by the vector signal generator. In the first scenario, there is no traffic on the opposite signal path within the bandwidth of the investigated carrier (see Figure 4.16). To check the proper functioning of the developed XPD estimation algorithms, the attenuation a_c of the cross-talked signal is varied by using the dip switches on the attenuators, with the attenuation of the used dip switches assessed by measurement using a spectrum analyzer, since unsatisfying accuracy of the variable attenuators was observed. The XPD estimation is repeated for different SNR values of the co-polar carrier, i.e. $\gamma_{s,c} = 13.8\,\text{dB}$ and $\gamma_{s,c} = 10.1\,\text{dB}$. An isolation from the cross- to the co-polar plane of roughly $\xi_x = -28\,\text{dB}$ is adjusted. The number of acquired samples is $L = 524288$ for each signal path. Moreover, it should be pointed out that the input of the ADC has to be well scaled to avoid effects of quantization noise or clipping.

In Figure 4.17, the results for the first scenario using the hardware testbed are shown. It can be seen that the performance of the fine estimation method (F HW) features a slight bias of approximately 0.3 dB over the whole XPD range. There may be several reasons for this effect: (i) measurement errors, induced by the determination of all adjusted XPD values using a spectrum analyzer; (ii) the design of the used anti-aliasing filter (AAF) is

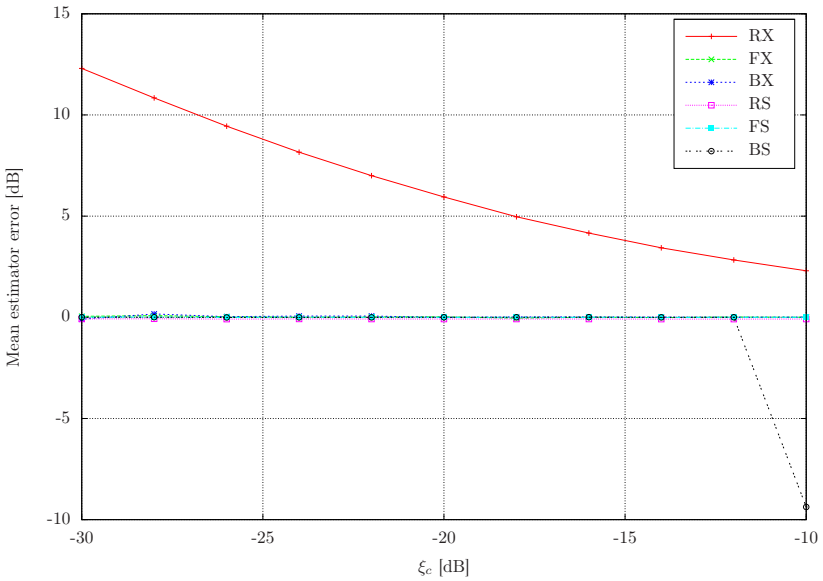


Figure 4.11: Mean XPD error versus attenuation using cross-polar subtraction

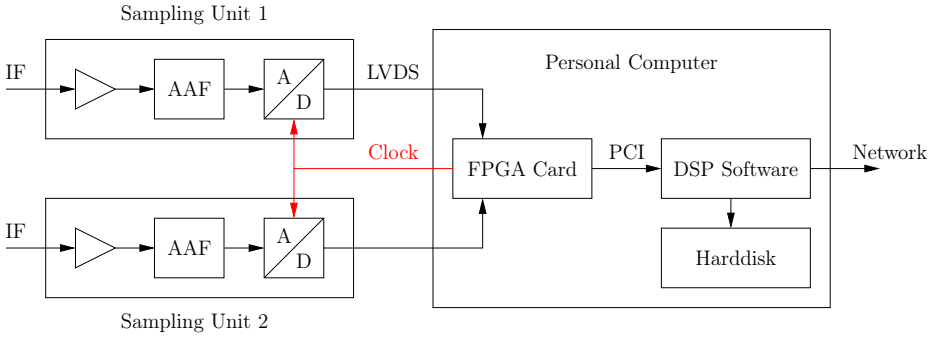


Figure 4.12: SDR demonstrator block diagram

of Chebyshev type, which exhibits considerable ripple in the passband. In the following, the affected noise floor will be applied for estimation of the scaling factor inside the blind demodulation framework and could lead to a constant offset. Keeping this in mind will help to interpret the results of the coarse method (C HW). For an SNR of $\gamma_{s,c} = 13.8\,\mathrm{dB}$, the coarse method shows outstanding performance. Nevertheless, it should be clear that the estimates deviate from the true values. The impact of the co-polar noise contribution decreases the mean estimator output, as can be seen more explicitly for an SNR of $\gamma_{s,c} = 10.1\,\mathrm{dB}$.

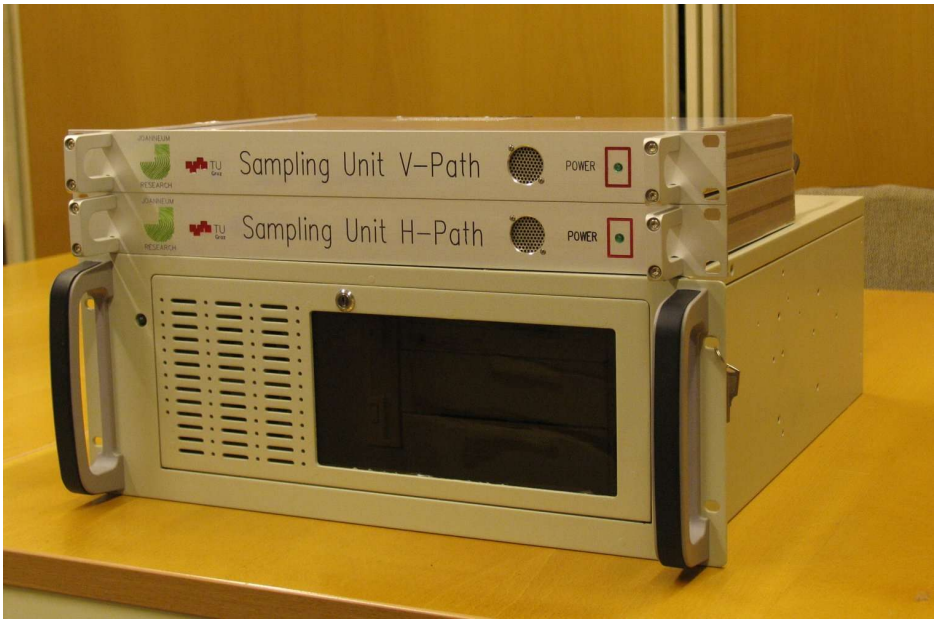


Figure 4.13: SDR demonstrator front side

In the second scenario, the cross-talk signal is influenced by opposite traffic, as depicted in Figure 4.18. The SNR of the interfering cross-polar 4 MSymb/s carrier was measured as $\gamma_{s,x} = 12.0\,\mathrm{dB}$. Previous to the fine method, no subtraction of the cross-polar traffic is applied, since the data generated by the signal generator is not continuous. The reason for this is that the cross-polar samples, stored on a hard disk inside the generator, are of finite length. Thus, the data record will repeat every N samples, e.g. $N = 1000000$. This will lead to non-continuous data and moreover the generator leaves a small gap before repeating the data stream. It should be clear that such a behavior would lead to problems for the blind demodulation process. Nevertheless such a scenario will not appear in an operational environment. The result for the second scenario is illustrated in Figure 4.19. The fine method shows the same solid behavior as for the first scenario. In contrast, the coarse scheme exhibits an increasing deviation from the ideal curve with decreasing co-polar attenuation ξ_c. It can be summarized that the results using the hardware environment verify the behavior, as expected from the pure software simulation.

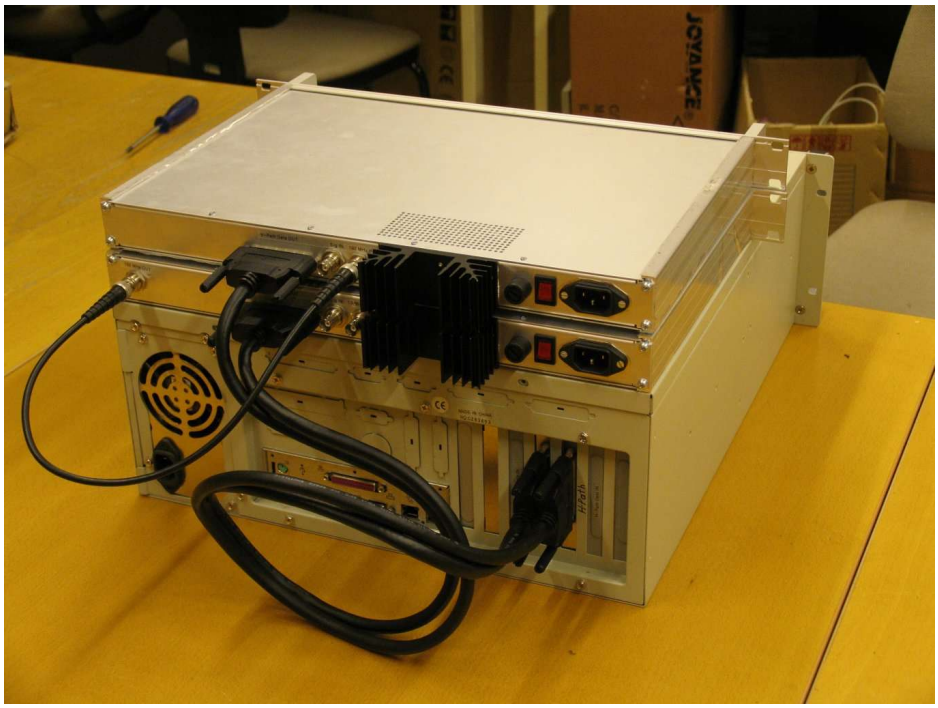


Figure 4.14: SDR demonstrator backside

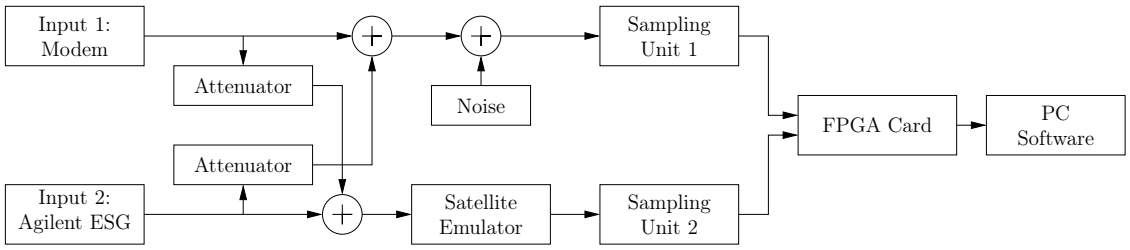


Figure 4.15: Overview of the testbed for XPD measurement

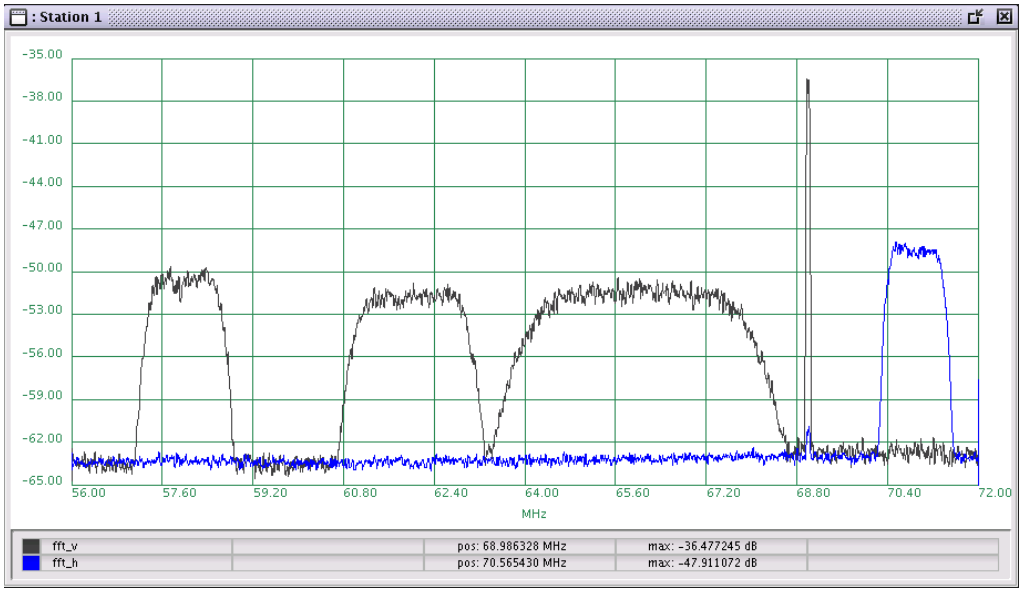


Figure 4.16: Testbed Scenario 1 (no cross-polar traffic)

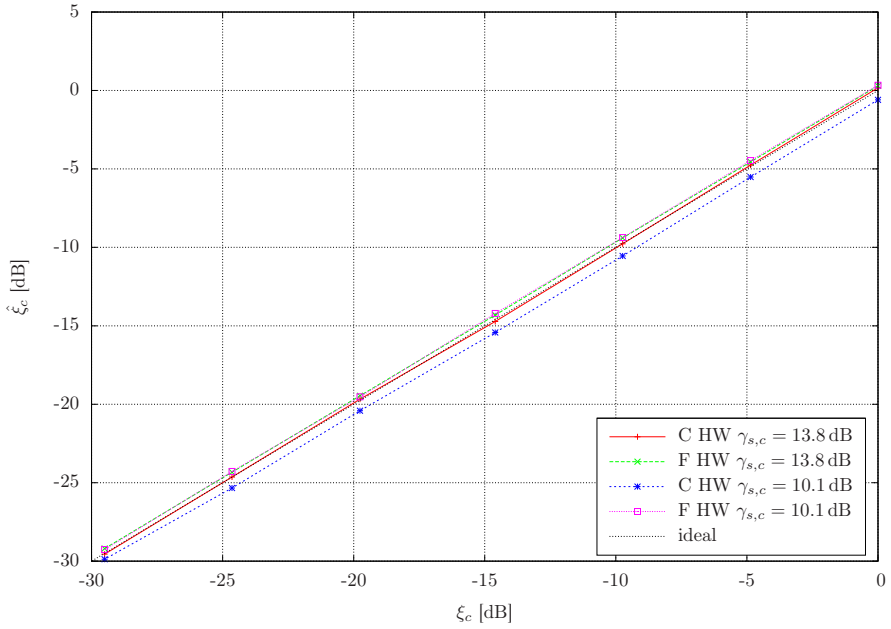


Figure 4.17: Testbed Scenario 1 (mean XPD estimator output)

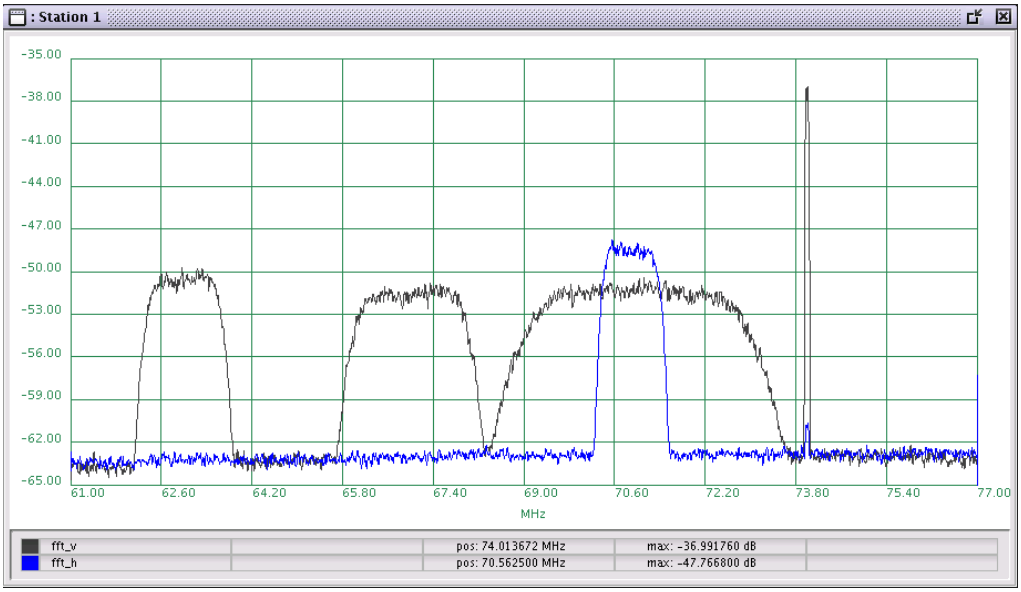


Figure 4.18: Testbed Scenario 2 (with cross-polar traffic)

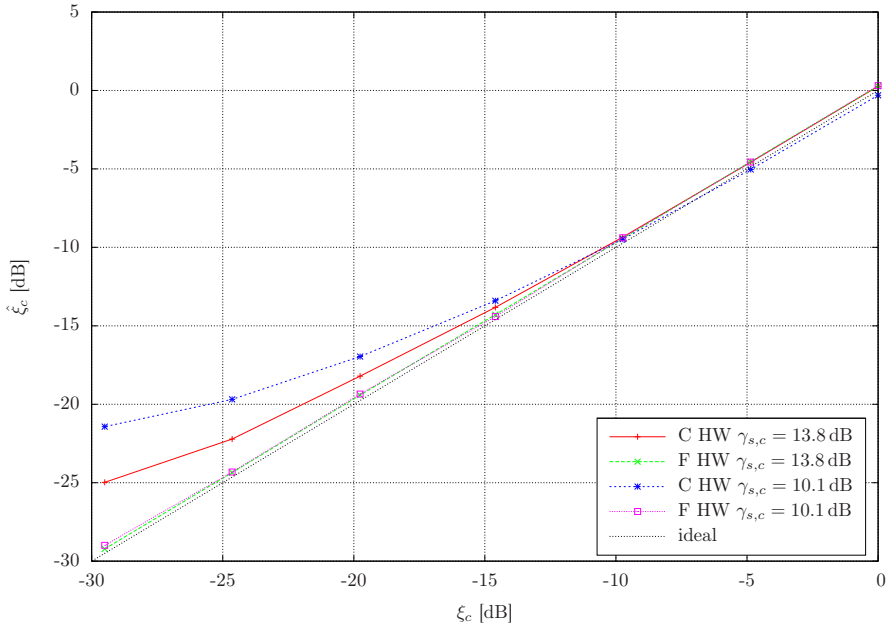


Figure 4.19: Testbed Scenario 2 (mean XPD estimator output)

4.2 Geolocation

On satellites, interference caused by misaligned antennas, satellite piracy or damaged oscillators, can affect the present satellite traffic seriously. Thus, the source of interference has to be identified as quickly as possible to be shut down. In case the perturbing carrier can not be determined from the provider database, the transmitting station has to be located by alternative means. Hence, with the increasing amount of satellite services and customers, the need for accurate passive geolocation systems arises.

4.2.1 Principle

The location of the interferer might be estimated by using both the time difference of arrival (TDOA) and the frequency difference of arrival (FDOA) of two adjacent satellites, as described in [70] and the according patent in [71]. In typical geolocation problems, the emitted signal is directly acquired by several receivers to compute position and velocity of the transmitter. However, this is not the case for the task to solve, since the signal is relayed by satellites and thus complicates the problem. The principle of the applied approach is sketched in Figure 4.20.

The signal $r_1(t)$ of the transmitter $\boldsymbol{r}_\mathrm{I}$ is relayed over the main satellite $\boldsymbol{r}_\mathrm{S1}$ to the monitoring site $\boldsymbol{r}_0$. Due to the finite transmitter antenna size, power will not only be radiated to the main satellite, but to others as well. The way how power is radiated can be characterized by the antenna pattern. From this it should be clear that an attenuated version $\tilde{r}_1(t)$ of the signal $r_1(t)$ might be relayed by an adjacent satellite $\boldsymbol{r}_\mathrm{S2}$. By deploying a second antenna at the monitoring station $\boldsymbol{r}_0$, the signal $r_2(t)$, which includes the overspilled signal $\tilde{r}_1(t)$, can be received. This assumption is only valid, if the adjacent satellite has a transponder in the same frequency band as the transmitted signal.

At the monitoring site $\boldsymbol{r}_0$, a TDOA between $r_1(t)$ and $\tilde{r}_1(t)$ can be observed. The latter is mainly caused by the different path lengths and additionally influenced by propagation effects and the differential transponder delay. Moreover, an FDOA can be observed, which stems from the differential Doppler and is directly affected by oscillator mismatches onboard the satellites. Any errors in the monitoring equipment are neglected, since proper calibration is assumed. By using a TDOA/FDOA pair, an estimate of the signal source may be obtained. Alternatively, by performing an additional TDOA measurement on a third satellite $\boldsymbol{r}_\mathrm{S3}$, the geolocation task can be solved by using a TDOA/TDOA pair as well.

It is obvious that the task to solve is a very complex one, such that in worst case an estimation of the transmitter location might be impossible. An example could be that the satellites exhibit no relative motion observable from the monitoring site leading to no measurable FDOA. On the other hand, for the TDOA/TDOA technique transmitters close to the equator are problematic as will be shown later on. Moreover, the measurement principle includes many input parameters that have a direct impact on the accuracy of the result, e.g. satellite ephemeris data and ground station coordinates. The main tasks [70]

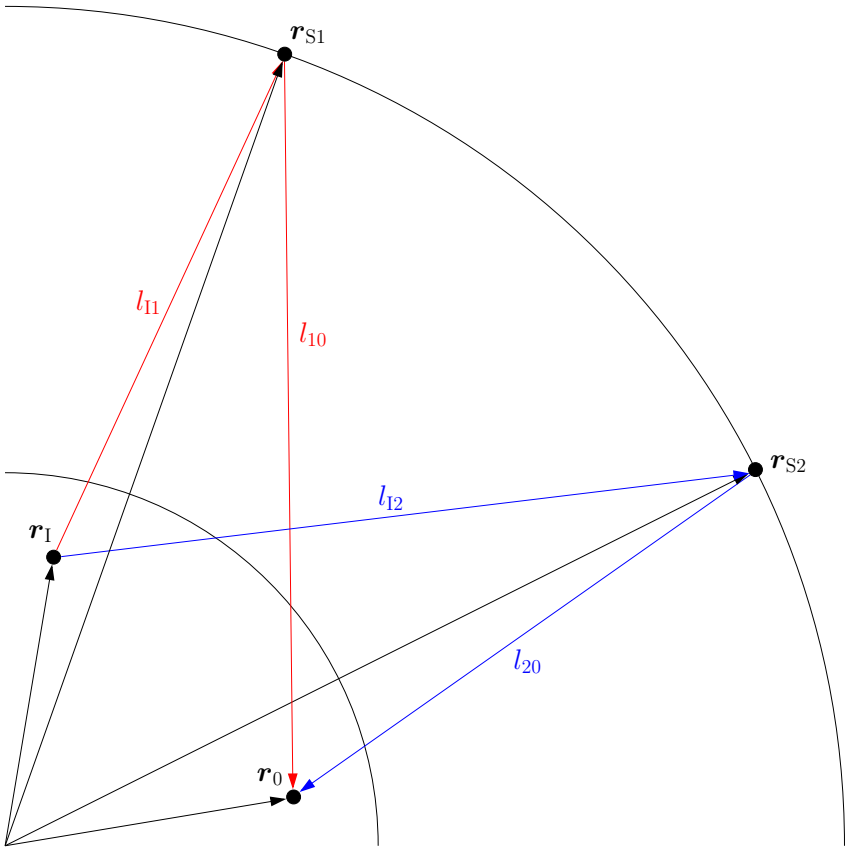


Figure 4.20: Basic configuration for TDOA/FDOA techniques

to solve are summarized in the following:

- At first, the CAF has to be evaluated for $r_1(t)$ and $r_2(t)$ to obtain a correlation peak and in the following an accurate TDOA/FDOA pair. The overspill from the main to the adjacent satellite may be very weak, depending on the actual geometrical constellation and the antenna characteristic. From the recommendations of the international radio consultative commitee (CCIR) for antenna patterns at 14.25 GHz, the attenuation at 3 degree from boresight is in the range of -30 dB to -42 dB, depending on the size of the transmitting antenna.
- The estimated TDOA/FDOA pair has to be converted to a transmitter location by applying suitable numerical methods.
- If input parameters, e.g. satellite positions or velocities, contain significant errors, the latter should be determined to get an improved geolocation accuracy.

4.2.2 Evaluation of the Cross-Ambiguity Function

To estimate TDOA/FDOA pairs, it is common practice to process the cross-ambiguity function (CAF)

$$A(\nu,\tau) = \int_0^{T_L} r_1(t)\, r_2^*(t+\tau) e^{-\jmath 2\pi\nu t} dt \tag{4.12}$$

The task is to find τ and ν so that $|A(\cdot)|$ becomes a maximum. The pair of values that satisfies this condition are the estimates for TDOA and FDOA. As already mentioned, the overspilled signal $\tilde{r}_1(t)$ embedded in $r_2(t)$ depends on the sidelobe characteristics of the interferer antenna and on the spacing between the two satellites. It is clear that sufficient processing gain is necessary to obtain a pronounced correlation peak, since the power level of $\tilde{r}_1(t)$ is usually far below the noise floor of the adjacent transponder. For this reason, long integration times T_L are required and thus a huge amount of data has to be processed. Similar as in the previous section, the post-correlation SNR can be obtained by

$$\gamma_{s,o} = 2BT_L \frac{\gamma_{s,1}\gamma_{s,2}}{\gamma_{s,1} + \gamma_{s,2} + 1} \tag{4.13}$$

Herein $\gamma_{s,1}$ denotes the SNR of $r_1(t)$ on the main satellite and $\gamma_{s,2}$ the SNR of $\tilde{r}_1(t)$ on the adjacent satellite.

Assuming that the signals are impaired by AWGN, lower bounds for the maximum achievable accuracy can be derived. The fundamental bound for TDOA can be assessed by [69]

$$\sigma_\tau^2 \geq \frac{3T^2}{\pi^2 \gamma_{s,o}} \tag{4.14}$$

It can be seen that $\sigma_\tau \sim T = 1/f_d$ and with $BT_L \sim L$, $\sigma_\tau \sim 1/\sqrt{L}$. The effective output SNR $\gamma_{s,o}$ includes the processing gain $2BT_L$ and as such is dependent on the number of samples L. The limit for FDOA can be stated as

$$\sigma_\nu^2 \geq \frac{3}{\pi^2 L^2 T_s^2 \gamma_{s,o}} \tag{4.15}$$

Thus, $\sigma_\nu \sim 1/T_s$ and $\sigma_\nu \sim 1/L^{3/2}$. It can be seen that the requirements for small σ_τ and σ_ν are in contrast to each other and thus an appropriate choice of integration time and bandwidth is necessary. The FDOA estimate benefits most from an increase of the observation time and a decrease of the sampling rate. More samples for processing do not increase the accuracy of the TDOA estimate that significantly, but broadband signals should be applied to the estimator.

Basically, a very similar task was solved for XPD estimation in Section 4.1. However, to achieve acceptable geolocation accuracy an increased amount of data has to be used for computation. Due to this data requirement, the library developed for XPD estimation is

not suited for this problem. For this reason, an optimized CAF library was built from scratch, based on the algorithm described in [69]. An example implementation of the latter can be found in [72] and a detailed comparison of efficient methods in [73].

The estimation of the accurate position of the correlation peak in the CAF surface in a single step is not possible from the computational point of view. Thus the search is split up in two steps: (i) a coarse step that uses small amounts of data and FFT correlation to detect the desired peak in the CAF plane; (ii) a fine step that uses huge amounts of data to track the peak until the required accuracy is achieved.

Coarse Method

The aim of the first step is to detect correlation peaks located in a defined time/frequency region. For this purpose, the ranges for time and frequency are set according to the actual scenario. Center and range of the time offset τ are selected by geometrical considerations and the maximum expected runtime differences. Similarly, the search region for the frequency offset ν is determined by the maximum Doppler and oscillator mismatch. It is obvious that the choice of the time/frequency search range determines the computational effort and the probability of detecting the appropriate correlation peak.

The CAF can be formulated in digitized form as [74]

$$A_{k,m} = \sum_{i=0}^{L-1} r_{1,i} r_{2,i+m}^{*} e^{-\jmath 2\pi ik/L} \tag{4.16}$$

This leaves us with a computational load in the order of $O(L^3)$. The correlation of the signals r_1 and r_2 is calculated using an FFT-based convolution which is suited for large search areas and small block lengths. Using the respective DFTs, denoted by $R_{1,l}$ and $R_{2,l}$, the CAF is equivalently given by

$$A_{k,m} = \frac{1}{L} \sum_{l=0}^{L-1} R_{1,l+k} R_{2,l}^{*} e^{-\jmath 2\pi ml/L} \tag{4.17}$$

Using the FFT for implementation of the DFTs, it can be shown that the computational effort is in the order of $O(L^2 \log_2 L)$. To search over a specified frequency search range, the DFT R_1 is shifted in frequency by k instead of multiplying r_1 with a complex exponential. If the time domain search range is set to the whole region of L samples, the outer sum can be implemented as an IFFT as well. Since an FFT-based convolution is circular, zero-padding has to be used to avoid aliasing in the time domain. The output of a linear convolution of two sequences of length L and M is of length $P = L + M - 1$; in contrast to $P = L$ of the circular convolution. Thus, it is obvious that in our case, where $M = L$, zero-padding of at least $k_{zp} = 2$ has to be applied.

After successful CAF computation, all peaks located within the specified search region are identified as trial peaks. Then the noise floor within the search area of the CAF plane

is calculated. In the following, only trial peaks that exhibit a significant SNR are treated as valid correlation peaks. The selection of the SNR threshold is a tradeoff between false detection and false rejection. For example, trial peaks with an SNR below 15 dB are rejected. Besides of spurious peaks caused by noise, additional peaks may occur because of multi-path effects.

If no peaks are found in the search region, the estimator length L is incremented to achieve a larger processing gain. However, it is clear that the estimator length L can not be incremented arbitrarily, since the computation time would raise out of bounds. Additionally, several CAFs can be combined non-coherently to obtain a flattening of the noise floor and thus avoid detection of spurious peaks. It is important to average the magnitude or squared magnitude of the CAFs and not the complex correlation samples. The magnitude is independent of phase, but complex correlation samples include phase information as well. Thus, depending on the FDOA it might be possible that each sample block starts at a different phase angle, resulting in a smaller correlation peak or even in disappearance of the peak.

In the following a simple scenario demonstrates the correlation peak detection in the CAF surface. For this reason, a single QPSK carrier oversampled at $N_s = 2$ and an SNR of $\gamma_s = 15\,\mathrm{dB}$ is generated. It is assumed that the overspilled signal $\tilde{r}_1$ is attenuated by 50 dB compared to signal r_1. Using a block length of $L = 262144$ samples for CAF computation, results in an effective output SNR of $\gamma_{s,o} = 19.05\,\mathrm{dB}$. The zero-padding factor is set to $k_{zp} = 2$. Finally, a time offset of $\tau = 0.3\,\mathrm{T_s}$ and a frequency offset of $\nu = 0.3\,\mathrm{bins}$ is adjusted. Figure 4.21 shows the resulting CAF without applying averaging. The correlation peak can be identified, however the fluctuations of the noise floor might lead to detection of spurious peaks. To reduce the probability of false detection, non-coherent averaging of 16 sample blocks is applied such that a total of 4194304 samples are processed. The obtained CAF plane is depicted in Figure 4.22. The correlation peak becomes very pronounced, since the jitter floor is flattened by the averaging process. The main drawback of averaging is the additional computational load by the increased amount of data.

Fine Method

After successful detection of a correlation peak using an estimator length L_d, the size of the time/frequency search area will be reduced dramatically to a 3×3 region. The CAF is evaluated according to Equation 4.16. The correlation of r_1 and r_2 for the 3 time delays τ of interest is calculated using the Intel performance primitives (IPP) library [1]. To evaluate the correlation at the designated frequency offset ν, the signal r_2 is multiplied by an appropriate complex exponential. This type of CAF evaluation is suited for very narrow search areas and huge block lengths.

The maximum of the correlation peak is set as the new center of the following CAF calculation. The initial estimator length of the fine method is set to $L_n = 2L_d$ and the

1 `http://software.intel.com/en-us/intel-ipp/`

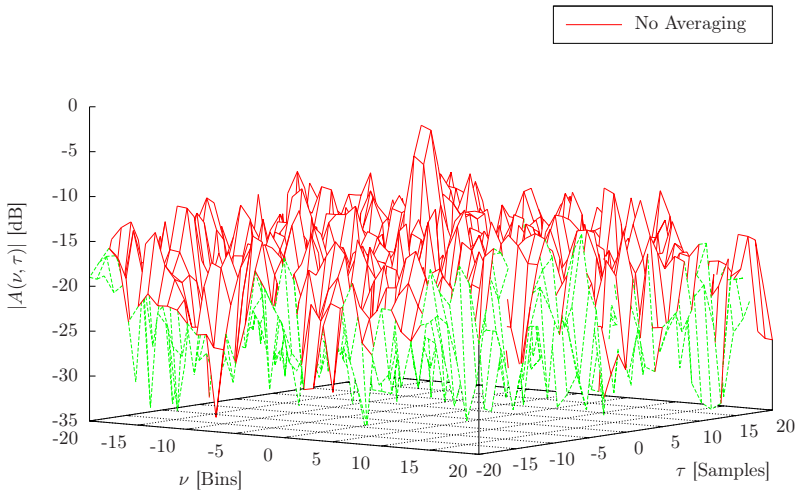


Figure 4.21: CAF without averaging

frequency resolution to $\Delta\nu = f_s/L_n$. Then the maximum of the CAF snippet is determined and set to the center of the next iteration using twice the input data $L_{n+1} = 2L_n$. This procedure is repeated until the requested accuracy or the number of all acquired samples is reached. Subsequently, parabolic interpolation (see Appendix A.1) is applied to get a refined value for time and frequency offset, since it is clear that the true values in general do not coincide with multiples of the sampling period or spectral resolution.

One major challenge is to handle the huge amount of data which is necessary to achieve the desired accuracy of the TDOA/FDOA pair. Depending on the specific scenario, the number of required input samples can be up to some hundred million samples. It is obvious that a powerful computing platform is necessary to process such an amount of data. Moreover, a 64-bit operating system is suggested to be able to allocate more than 2 GB of memory per application.

Simulation Results

In the following, the accuracy of the implemented CAF evaluation is compared to the theoretical limits, provided by the CRLBs in Equations 4.14 and 4.15. For this reason, it is assumed that the signal from the main satellite consists of a single QPSK carrier with an SNR of $\gamma_{s,1} = 15\,\text{dB}$ and a symbol rate of $f_d = 12\,\text{MBaud}$. The sample rate of the receiving hardware is assumed to be $f_s = 48\,\text{MHz}$, such that the oversampling factor

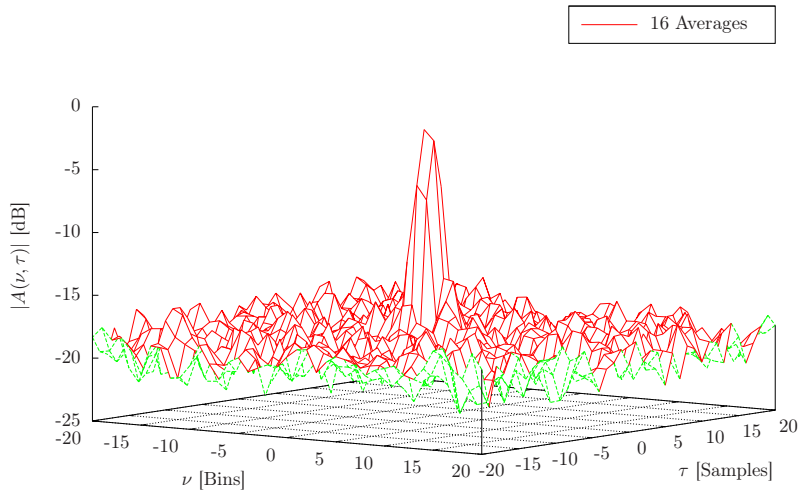


Figure 4.22: CAF with averaging

becomes $N_s = 4$. The number of samples used for CAF computation is set to $L = 2097152$. For each simulation point 100 iterations are performed with randomly selected values for TDOA and FDOA. The overspill to the adjacent satellite is varied in the range from -10 to $-50\,\mathrm{dB}$. The evolution of the standard deviation for the TDOA estimate $\hat{\tau}$ is plotted versus the overspill in Figure 4.23. For comparison reasons the according CRLB is illustrated as well. In the first simulation run no traffic is generated on the adjacent satellite (No TX). It can be seen that the standard deviation follows the CRLB except for large effective SNRs, for which a jitter floor can be observed. The latter stems mainly from the error introduced by interpolation. If the TDOA value coincides with a multiple of the sampling period T_s, the jitter floor will vanish completely. Different interpolation techniques or sampling periods could be used to reduce this sort of degradation. In the second simulation run, traffic with an SNR of $\gamma_s = 15\,\mathrm{dB}$ is present on the adjacent transponder (TX). Of course, this will lead to a reduction of the effective output SNR and thus to an increase of the standard deviation when not augmenting the estimator length appropriately. Using the subtraction capability of the blind demodulation framework, the traffic present on the adjacent satellite could be removed to increase the accuracy of the estimate. The performance evolution is shown for ideal (TXS Ideal) and real (TXS) carrier subtraction. It is obvious that the impact of the perturbing traffic can be removed entirely. For a strong overspill of -10 dB, the adjacent traffic is severely interfered by the main carrier. So, the adjacent traffic can not be subtracted successfully and an increase

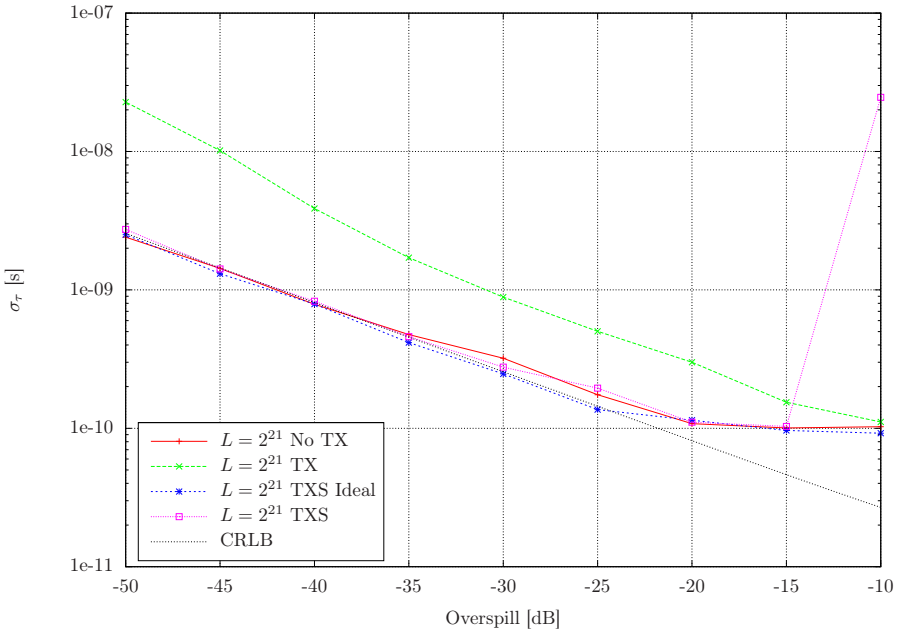


Figure 4.23: Evolution of the TDOA standard deviation

of the standard deviation can be observed. It should be mentioned that such a strong overspill should not occur in real life applications.

In Figure 4.24 the results for the FDOA estimate $\hat{\nu}$ are shown. It can be noticed that no jitter floor is present in this case, since frequency resolution is increased when augmenting the estimator length. As detailed for the TDOA estimate, the application of the subtraction capability can significantly improve the performance of the CAF evaluation in presence of adjacent traffic.

4.2.3 Analytical Model for Geolocation

Standard Approach

For both, TDOA and FDOA, a line of position (LOP) can be drawn on the surface of the Earth ellipsoid. The intersection of the LOPs corresponds to the estimate of the interferer location. Using the illustration in Figure 4.20, the LOP for the TDOA can be characterized by following nonlinear vector equation.

TDOA

$$f_1(\mathbf{r}_1) = l_{I1} + l_{10} - (l_{I2} + l_{20}) - c \cdot \mathrm{TDOA} + c \cdot \Delta t_{12} = 0 \tag{4.18}$$

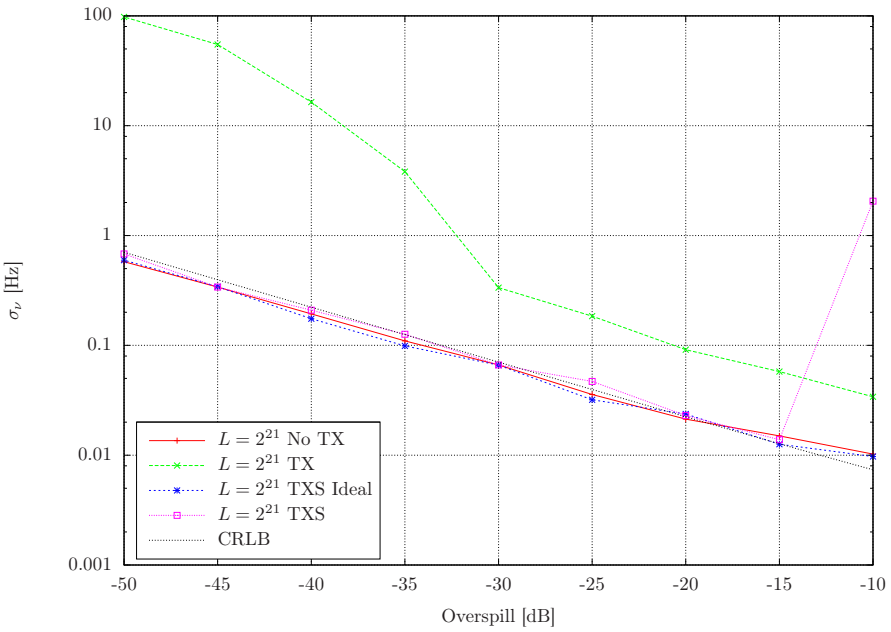


Figure 4.24: Evolution of the FDOA standard deviation

Commonly the uplink time difference $\text{TDOA}_{\text{up}} = \text{TDOA} - \text{TDOA}_{\text{down}}$ is normalized by the propagation speed c to arrive at the differential slant range (DSR). The term Δt_{12} considers a possible differential delay inside the satellites. Therefore, by rearranging the terms in Equation 4.18, we have that

$$\begin{aligned} f_1(\mathbf{r}_\text{I}) &= l_{I1} - l_{I2} - (c \cdot \text{TDOA} - (l_{10} - l_{20})) + \Delta s_{12} \\ &= l_{I1} - l_{I2} - c \cdot (\text{TDOA} - \text{TDOA}_{\text{down}}) + \Delta s_{12} \\ &= l_{I1} - l_{I2} - \text{DSR} + \Delta s_{12} = 0 \end{aligned} \tag{4.19}$$

In similar manner, an equation for the FDOA can be formulated.

FDOA

$$\begin{aligned} \tilde{f}_2(\mathbf{r}_\text{I}) = &\frac{f}{c}\,\mathbf{v}_{\text{S2}} \cdot \mathbf{e}_{\text{I2}} + \frac{f - f_{T2}}{c}\,\mathbf{v}_{\text{S2}} \cdot \mathbf{e}_{20} \\ &- \frac{f}{c}\,\mathbf{v}_{\text{S1}} \cdot \mathbf{e}_{\text{I1}} - \frac{f - f_{T1}}{c}\,\mathbf{v}_{\text{S1}} \cdot \mathbf{e}_{10} - \text{FDOA} + \Delta f_{12} = 0 \end{aligned} \tag{4.20}$$

Herein, $\mathbf{e}_{\text{I1}}$ is the unit vector pointing from the interferer to the main satellite. If $\mathbf{v}_{\text{S1}}$ is the velocity vector of the main satellite, then $\mathbf{v}_{\text{S1}} \cdot \mathbf{e}_{\text{I1}}$ considers the radial velocity of the main satellite observed from the interferer. Furthermore, f is the frequency of the

transmitted signal and the translation frequency of the main satellite is denoted by f_{T1}. The factor Δf_{12} accounts for the differential oscillator mismatch onboard the satellites. Finally, the normalized frequency difference in the uplink is designated as differential slant range rate (DSRR). The meaning of the remaining terms can be identified by the indices easily so that, by rearranging the terms in Equation 4.20, we obtain

$$\begin{aligned} f_2(\mathbf{r}_\mathrm{I}) &= \mathbf{v}_{\mathrm{S2}} \cdot \mathbf{e}_{\mathrm{I2}} - \mathbf{v}_{\mathrm{S1}} \cdot \mathbf{e}_{\mathrm{I1}} - \frac{c}{f} \cdot (\mathrm{FDOA} - \mathrm{FDOA}_{\mathrm{down}}) + \Delta\nu_{12} \\ &= \mathbf{v}_{\mathrm{S2}} \cdot \mathbf{e}_{\mathrm{I2}} - \mathbf{v}_{\mathrm{S1}} \cdot \mathbf{e}_{\mathrm{I1}} - \mathrm{DSRR} + \Delta\nu_{12} = 0 \end{aligned} \tag{4.21}$$

Since the location of the transmitter can only be described uniquely by three coordinates, an additional relationship is required. For this reason the surface of the Earth is approximated by an ellipsoid.

Ellipsoid

$$f_3(\mathbf{r}_\mathrm{I}) = \frac{x_I^2}{R_A^2} + \frac{y_I^2}{R_A^2} + \frac{z_I^2}{R_B^2} - 1 = 0 \tag{4.22}$$

whereby R_A is the major and R_B the minor semi-axis of the ellipsoid.

Assuming that Δs_{12} and $\Delta\nu_{12}$ are negligible, three nonlinear equations are available to estimate the location of the interferer. Since these equations are not solvable in closed form, an appropriate numerical approach has to be applied. The input parameters to the numerical method are the estimated TDOA/FDOA pairs, the geometrical parameters, i.e. position and velocity vectors of main and adjacent satellites, ground station coordinates and system-dependent parameters (oscillator frequencies, uplink frequency, differential time and frequency offset between both transponders). As already mentioned, the solutions of $f_1(\cdot)$ and $f_2(\cdot)$ produce LOPs that can be mapped on the Earth to find the transmitter location. A graphical representation of TDOA/FDOA LOPs in Europe can be seen in Figure 4.25. In this scenario, the transmitter is assumed to be in Lisbon and the monitoring site in Vienna. The first satellite is located at an azimuth of $\phi_{S1} = 10\,^\circ$, an elevation of $\beta_{S1} = 0\,^\circ$ and a radius of $r_{S1} = 42164.537\,\mathrm{km}$ when assuming a geocentric coordinate system. The coordinates of the adjacent satellite differ only in the azimuth by $\phi_{S2} = 7\,^\circ$.

In the second scenario, the monitoring site is located in Rio de Janeiro and the transmitter in Rio Branco. The main satellite exhibits an azimuth of $\phi_{S1} = -61\,^\circ$ and the adjacent satellite of $\phi_{S2} = -63\,^\circ$. The Brazilian scenario is illustrated in Figure 4.26.

The problem of this approach becomes obvious immediately. Up to now it was assumed that Δt_{12} as well as Δf_{12} are negligible. Usually, the differential transponder delay Δt_{12} is very small and thus does not degrade geolocation accuracy. However, this does not hold true for the differential oscillator mismatch Δf_{12}, which can be in the order of several kHz, leading to arbitrarily wrong results. The estimation of this mismatch can be carried out by a separate measurement or be incorporated into the geolocation process as demonstrated in the following.

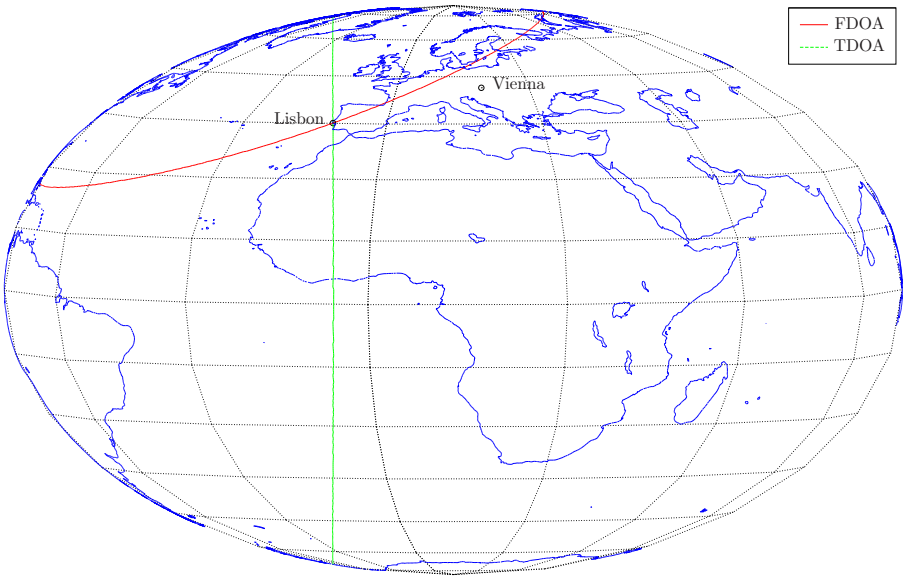


Figure 4.25: Evolution of TDOA/FDOA LOPs in Europe

Differential Approach

For determining the differential oscillator mismatch and transponder delay, which are not known in advance, additional to the monitoring station a reference station is needed [70]. In this case, a carrier with known origin, e.g. from the operator database, on the main transponder is used as reference signal. The differential method cancels out effects that are common to interferer and reference signal, i.e. satellite imperfections (differential oscillator mismatch and transponder delay), propagation effects on the downlink and impairments by the monitoring station equipment. Two additional benefits of the differential method are: (i) better numerical stability of the equation system; (ii) reduction of errors in satellite ephemeris data that have common impact on TDOA/FDOA for interferer and reference station. For the sake of completeness it should be mentioned that, if the interferer is located at the reference site, effects on the uplink will cancel out as well. The according TDOA equation for the reference station can be set up as follows:

$$f_1(\mathbf{r}_\mathrm{R}) = l_{R1} + l_{10} - (l_{R2} + l_{20}) - c \cdot \mathrm{TDOA}_\mathrm{R} + c \cdot \Delta t_{12} \tag{4.23}$$

Herein the subscript R denotes values belonging to the reference station. Subtraction of Equation 4.23 from 4.18 leads to the requested differential TDOA relationship.

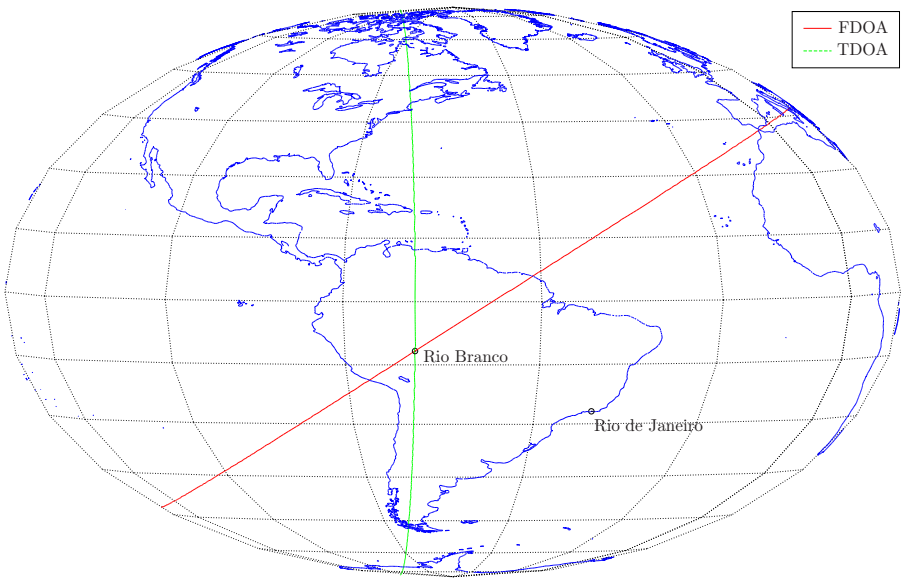


Figure 4.26: Evolution of TDOA/FDOA LOPs in Brazil

TDOA

$$\begin{aligned} f_1(\mathbf{r}_\mathrm{I}) &= l_{I1} - l_{I2} - (c \cdot (\mathrm{TDOA} - \mathrm{TDOA_R}) + l_{R1} - l_{R2}) \\ &= l_{I1} - l_{I2} - \mathrm{DSR} = 0 \end{aligned} \tag{4.24}$$

In a similar manner the differential approach can be formulated for the FDOA.

FDOA

$$\begin{aligned} f_2(\mathbf{r}_\mathrm{I}) = \mathbf{v}_{\mathrm{S2}} \cdot \mathbf{e}_{\mathrm{I2}} - \mathbf{v}_{\mathrm{S1}} \cdot \mathbf{e}_{\mathrm{I1}} - \Big(& \frac{c}{f} \cdot (\mathrm{FDOA} - \mathrm{FDOA_R}) \\ & + \frac{f_R}{f} \mathbf{v}_{\mathrm{S2}} \cdot \mathbf{e}_{\mathrm{R2}} - \frac{f_R}{f} \mathbf{v}_{\mathrm{S1}} \cdot \mathbf{e}_{\mathrm{R1}} \\ & - (\frac{f - f_R}{f} \mathbf{v}_{\mathrm{S2}} \cdot \mathbf{e}_{20} - \frac{f - f_R}{f} \mathbf{v}_{\mathrm{S1}} \cdot \mathbf{e}_{10}) \Big) \\ = \mathbf{v}_{\mathrm{S2}} \cdot \mathbf{e}_{\mathrm{I2}} - \mathbf{v}_{\mathrm{S1}} \cdot \mathbf{e}_{\mathrm{I1}} - & \mathrm{DSRR} \end{aligned} \tag{4.25}$$

Since basically two TDOA/FDOA measurements are necessary for differential processing, twice the jitter is introduced into the equation system. The main error contribution stems from uncertainty in the FDOA estimate. Another source of degradation are errors in the

available satellite ephemeris data, however, this issue will be discussed in more detail in a subsequent section.

4.2.4 Numerical Solution

The above equation system is solved by an iterative Gauss-Newton scheme, requiring the Jacobi matrix of the three nonlinear relationships. There are two reasons why not to use the Newton method directly [75]: (i) an exact solution may not exist, since the measured TDOA/FDOA pairs are superimposed by noise and the satellite ephemeris data may be erroneous as well; (ii) the Newton scheme is only suited for systems where the number of measurements coincides with the number of unknowns. Instead it is suggested to apply the Gauss-Newton method that uses a Newton scheme to minimize a nonlinear least squares problem $f(\mathbf{x})$ with m equations:

$$f(\mathbf{x}) = \frac{1}{2}\sum_{i=1}^{m}\left(f_i(\mathbf{x})\right)^2 \tag{4.26}$$

Herein $\mathbf{x} = \mathbf{r}_\mathrm{I} = (x_I,y_I,z_I)^T$ denotes the unknown vector holding the interferer position. When defining the equation vector $\mathbf{f}(\mathbf{x}) = (f_1,f_2,f_3)^T$, the (k+1)-th iteration step of the Gauss-Newton method can then be formulated as

$$\mathbf{x}_{k+1} = \mathbf{x}_k - (\mathbf{J}(\mathbf{x}_k)^T\mathbf{J}(\mathbf{x}_k))^{-1}\cdot\mathbf{J}(\mathbf{x}_k)^T\mathbf{f}(\mathbf{x}_k) \tag{4.27}$$

Herein $\mathbf{J}$ denotes the according Jacobi matrix. The definition of the latter can be found in Appendix A.3.1; the starting point $\mathbf{x}_0$ can be determined by a coarse search algorithm or is set to the location of the monitoring station $\mathbf{r}_0$ since the algorithm converges very fast. The termination of the iteration procedure relies on the following criteria

$$\left|1 - \frac{|\mathbf{x}_{k+1}|}{|\mathbf{x}_k|}\right| < \epsilon \tag{4.28}$$

with ϵ, for example, set to $\epsilon = 10^{-9}$.

Simulation Results

For the following simulation scenario it is assumed that the monitoring site is located in Vienna, the reference station in Paris and the interfering station in Lisbon. The SNR of interfering and adjacent traffic is set to $\gamma_s = 15\,\mathrm{dB}$. Oversampling the main carrier at a rate of $N_s = 4$ and using $L = 2^{25}$ samples results in a processing gain of 72.25 dB. The overspill from the main to the adjacent satellite is assumed to be $-50\,\mathrm{dB}$. The effective SNR is then $\gamma_{s,o} = 22.11\,\mathrm{dB}$. The two GEO satellites are located at an azimuth of $\phi_{S1} = 10\,^\circ$ and $\phi_{S2} = 7\,^\circ$. The satellite velocities and the Ku-band frequencies are set to values occurring in practice. For each of the 1000 geolocation attempts the true TDOA/FDOA pair is

superimposed by Gaussian noise $\mathcal{N}(0,\sigma_\tau^2)$ and $\mathcal{N}(0,\sigma_\nu^2)$ with standard deviations according to Equations 4.14 and 4.15. The geolocation errors in direction of azimuth and elevation are depicted in Figure 4.27 (red crosses). It is obvious that the jitter in North/South

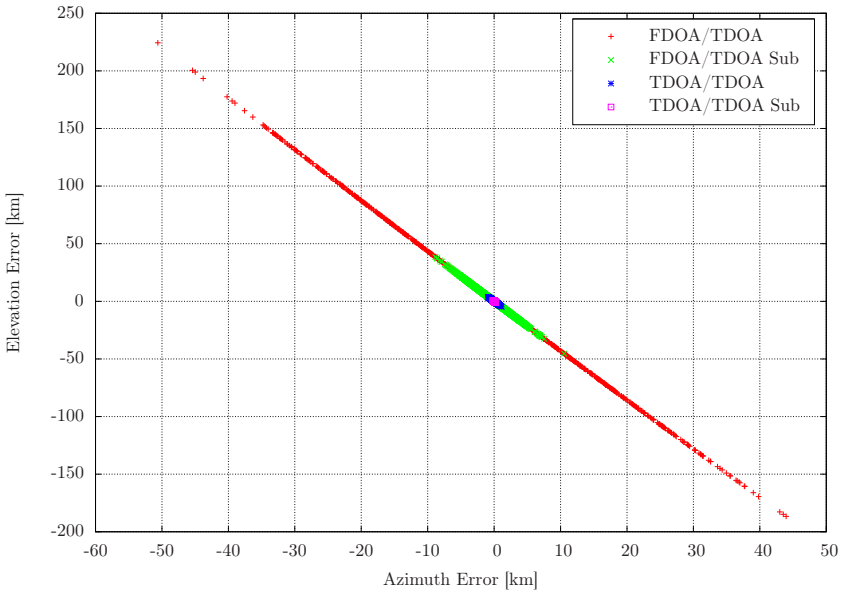


Figure 4.27: European scenario (accuracy improvement by blind demodulation)

direction is much larger than in East/West direction. This is mainly caused by the large jitter of the FDOA estimate, with a standard deviation in the order of $\sigma_\nu \approx 0.1\,\text{Hz}$. An estimation of the shape of the geolocation uncertainty can be given by the so-called error ellipse [70, 76].

The jitter in the FDOA estimate could be reduced by increasing the observation length L. However, further increasing the acquisition time will lead to stringent requirements on processing power and memory. Moreover, up to now it was assumed that the key parameters, i.e. satellite positions and velocities, are constant. With increasing observation time, this assumption might get more and more invalid, leading to a distortion of the CAF peak [77]. An enhancement of the geolocation accuracy using the same observation length could be possible by applying the subtraction capability of the blind demodulation framework on the adjacent traffic. In the above scenario, the transmitted signal $r_1(t)$ is buried 50 dB below signal-plus-noise on the adjacent signal path. The removal of the adjacent traffic could improve this relationship up to 35 dB, resulting in an effective SNR of $\gamma_{s,o} = 37.11\,\text{dB}$. The performance of the geolocation approach using adjacent traffic subtraction (green crosses) is shown in Figure 4.27 as well. It becomes obvious that successful carrier removal leads to a significant improvement of the geolocation accuracy.

Since the FDOA estimate seems to be a critical factor for the geolocation accuracy, an

approach using only TDOA estimates could be an attractive alternative [78, 79]. For this reason, a third satellite is required which relays the transmitted signal r_1 to the monitoring station. The above iterative procedure can be used with the Jacobi matrix $\mathbf{J}$ described in Appendix A.3.2. The results for the TDOA/TDOA approach with (magenta squares) and without (blue stars) adjacent traffic subtraction are also illustrated in Figure 4.27. The third GEO satellite is located at an azimuth of $\phi_{S3} = 13\,^\circ$ and the overspill from the main satellite is assumed to be attenuated by $-55\,\mathrm{dB}$. The TDOA/TDOA method, especially with additional carrier removal, exhibits outstanding performance. However, the latter approach features disadvantages as well: (i) a need of three satellites; (ii) additional costs at the monitoring site, including a third antenna; (iii) close to the equatorial region the TDOA LOPs are almost parallel, such that finding an accurate intersection might become a problem. An example for such a scenario is illustrated in Figure 4.28 with Libreville as transmitting station.

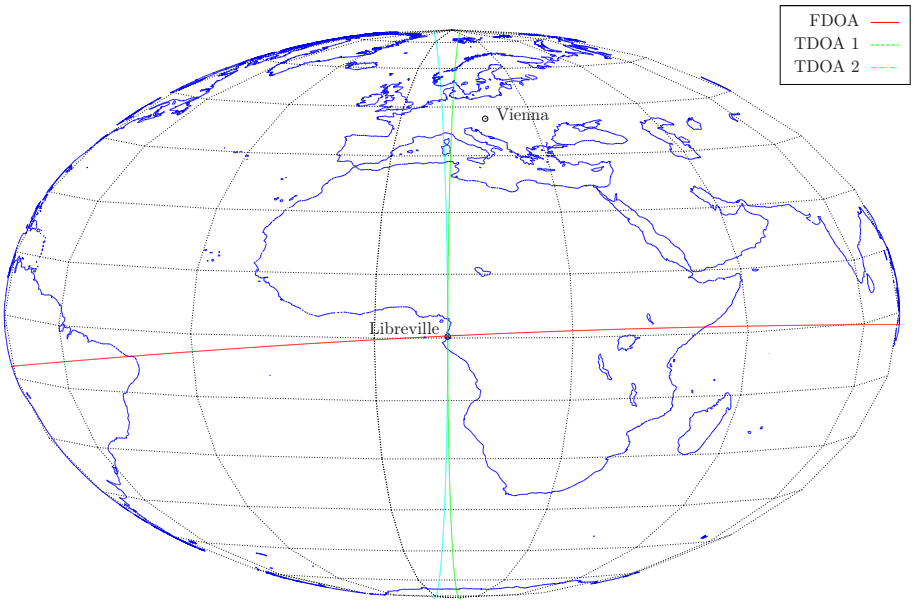


Figure 4.28: Evolution of TDOA/FDOA LOPs close to the equator

Numerical Aspects

A measure for the numerical stability of an equation system of the form $\mathbf{Ax} = \mathbf{b}$ is the condition number [80], which is defined as

$$\kappa = \mathrm{cond}\,(\mathbf{A}) = \left|\frac{s_{max}}{s_{min}}\right| \tag{4.29}$$

where s_{min} is the smallest and s_{max} the largest singular value of $\mathbf{A}$. For a condition number close to 1 the problem is well-conditioned, but with increasing condition numbers it gets more and more ill-conditioned. The latter leads to a behavior where a small perturbation of the measurement vector $\mathbf{b}$ or the coefficient matrix $\mathbf{A}$ results in a large variation of the solution vector $\mathbf{x}$ [81]. For the above scenario, the condition number is in the order of $\kappa_1 = 10^6$ for the TDOA/FDOA approach and of $\kappa_2 = 10^2$ for the TDOA/TDOA method. It is obvious that κ_2 is in a moderate range compared to the significant larger κ_1 which may lead to stability problems.

Satellite Ephemeris Error Up to now it was assumed that the available satellite ephemeris data, i.e. satellite position and velocity vectors, coincide with the true values. The accuracy of the ephemeris data is depending on the specific satellite operator. It may be available in data base form, relying on occasional measurements, orbit models or even both. Nevertheless, it is clear that any errors in satellite positions directly affect TDOA estimates and errors in satellite velocities influence FDOA estimates. In Figure 4.29 the geolocation accuracy is shown in presence of ephemeris errors. The position vectors

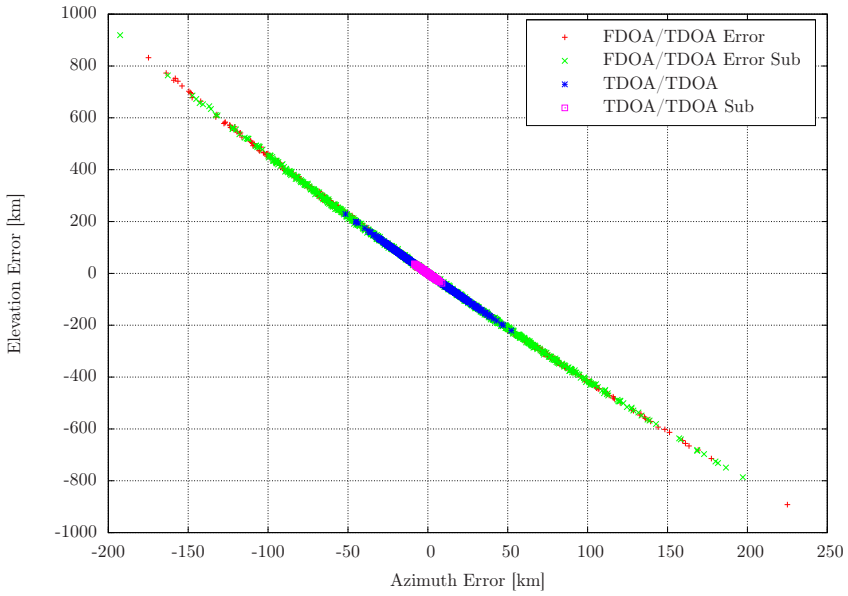


Figure 4.29: European scenario (geolocation accuracy impaired by ephemeris errors)

$\mathbf{r}_{S1}$, $\mathbf{r}_{S2}$ are superimposed by Gaussian noise with standard deviation $\sigma_r = 1\,\text{km}$ and the velocity vectors $\mathbf{v}_{S1}$, $\mathbf{v}_{S2}$ with standard deviation $\sigma_v = 0.1\,\text{m/s}$. It can be seen that TDOA/FDOA geolocation is dominated by ephemeris errors such that even subtraction of the adjacent traffic does not improve the catastrophic result. Moreover, it is obvious that satellite position errors have a significant impact on TDOA/TDOA as well. In this case,

removal of the adjacent traffic can reduce the geolocation jitter, however, it still remains at an unsatisfactory level. For this reason a scheme to refine the available ephemeris data would be favorable and is presented in the following.

4.3 Satellite Ranging

The refinement of the satellite ephemeris data was already requested in [70], later a procedure relying on TDOA/FDOA measurements was patented [82]. Figure 4.30 shows an example of the ephemeris data provided by a satellite operator. In the example, the

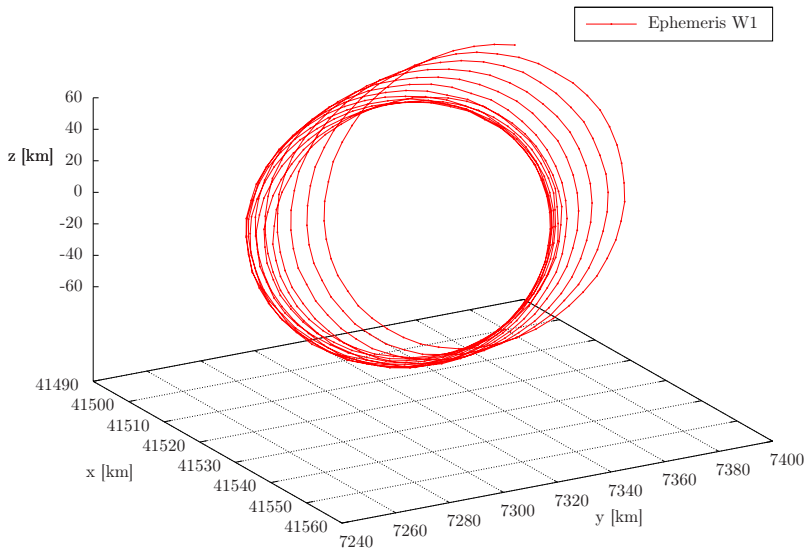


Figure 4.30: Satellite position ephemeris data (Eutelsat W1)

satellite position is predicted for two weeks in intervals of 30 minutes. In order to estimate position and velocity vectors unambiguously, seven TDOA/FDOA measurements are required. For this reason seven carriers with known origin have to be present on the main satellite. Moreover, an appropriate overspill to the second satellite is necessary for detection of a correlation peak within the CAF plane.

4.3.1 Standard Approach

TDOA The equation set for the differential time delay Δt_{12} and both satellite positions $\mathbf{r}_{\mathrm{S1}} = (x_{S1},y_{S1},z_{S1})^T$ and $\mathbf{r}_{\mathrm{S2}} = (x_{S2},y_{S2},z_{S2})^T$ can be stated as

$$f_{i,T}(\mathbf{r}_{S1},\mathbf{r}_{S2},\Delta t_{12}) = l_{i,1} + l_{10} - (l_{i,2} + l_{20}) - c \cdot \mathrm{TDOA}_i + c \cdot \Delta t_{12} \tag{4.30}$$

with i identifying the index of the corresponding base station. The unknown vector can be formulated as $\mathbf{x} = (x_{S1},y_{S1},z_{S1},x_{S2},y_{S2},z_{S2},\Delta t_{12})^T$ and the equation vector as $\mathbf{f} = (f_{1,T},\ldots,f_{7,T})^T$, resulting in the Jacobi matrix as $\mathbf{J} = \partial\mathbf{f}/\partial\mathbf{x}$. The definition of $\mathbf{J}$ can be found in Appendix A.3.3.

FDOA The equation set for the differential frequency offset Δf_{12} and both satellite velocities $\mathbf{v}_{S1} = (\dot{x}_{S1},\dot{y}_{S1},\dot{z}_{S1})^T$ and $\mathbf{v}_{S2} = (\dot{x}_{S2},\dot{y}_{S2},\dot{z}_{S2})^T$ is provided by

$$\begin{aligned} f_{i,F}(\mathbf{v}_{S1},\mathbf{v}_{S2},\Delta f_{12}) = {} & \mathbf{v}_{\mathrm{S2}} \cdot \mathbf{e}_{\mathrm{i,2}} + \frac{f_i - f_{T2}}{f_i}\mathbf{v}_{\mathrm{S2}} \cdot \mathbf{e}_{20} - \mathbf{v}_{\mathrm{S1}} \cdot \mathbf{e}_{\mathrm{i,1}} \\ & - \frac{f_i - f_{T1}}{f_i}\mathbf{v}_{\mathrm{S1}} \cdot \mathbf{e}_{10} + \frac{c}{f_i} \cdot \Delta f_{12} - \frac{c}{f_i} \cdot \mathrm{FDOA}_i \end{aligned} \tag{4.31}$$

The unknown vector can be formulated as $\dot{\mathbf{x}} = (\dot{x}_{S1},\dot{y}_{S1},\dot{z}_{S1},\dot{x}_{S2},\dot{y}_{S2},\dot{z}_{S2},\Delta f_{12})^T$ and the equation vector as $\dot{\mathbf{f}} = (f_{1,F},\ldots,f_{7,F})^T$, resulting in the Jacobi matrix as $\mathbf{J} = \partial\dot{\mathbf{f}}/\partial\dot{\mathbf{x}}$. More details on $\mathbf{J}$ are given in Appendix A.3.3.

4.3.2 Differential Approach

For reasons of stability it is suggested to use the differential approach for ranging as well.

TDOA For the positions of main and adjacent satellite six differential TDOA equations can be established according to

$$f_{i,T}(\mathbf{r}_{S1},\mathbf{r}_{S2}) = l_{i,1} - l_{i,2} - c \cdot \mathrm{TDOA}_i - (l_{R1} - l_{R2} - c \cdot \mathrm{TDOA}_R) \tag{4.32}$$

FDOA The differential FDOA equation system for both velocity vectors is given by

$$\begin{aligned} f_{i,F}(\mathbf{v}_{S1},\mathbf{v}_{S2}) = {} & \mathbf{v}_{\mathrm{S2}} \cdot \mathbf{e}_{\mathrm{i,2}} + \frac{f_i - f_{T2}}{f_i}\mathbf{v}_{\mathrm{S2}} \cdot \mathbf{e}_{20} - \mathbf{v}_{\mathrm{S1}} \cdot \mathbf{e}_{\mathrm{i,1}} \\ & - \frac{f_i - f_{T1}}{f_i}\mathbf{v}_{\mathrm{S1}} \cdot \mathbf{e}_{10} - \frac{c}{f_i} \cdot \mathrm{FDOA}_i \\ & - \frac{f_R}{f_i}[\mathbf{v}_{\mathrm{S2}} \cdot \mathbf{e}_{\mathrm{R2}} + \frac{f_R - f_{T2}}{f_R}\mathbf{v}_{\mathrm{S2}} \cdot \mathbf{e}_{20} - \mathbf{v}_{\mathrm{S1}} \cdot \mathbf{e}_{\mathrm{R1}} \\ & - \frac{f_R - f_{T1}}{f_R}\mathbf{v}_{\mathrm{S1}} \cdot \mathbf{e}_{10} - \frac{c}{f_R} \cdot \mathrm{FDOA}_R] \end{aligned} \tag{4.33}$$

In Appendix A.3.4 the Jacobi matrices **J** for TDOA and FDOA are listed.

4.3.3 Numerical Aspects

The standard approach tends to suffer from severe numerical problems caused by possible inversion of a singular matrix. The differential method reduces the large condition numbers by more than one order of magnitude to 10^6 or 10^7. However, still problems due to the mentioned inversion remain.

Singular Values Decomposition The singular value decomposition (SVD) is used to determine the numerical stability of the equation system [83]; but the approach might be applied as well for effective matrix inversion or the decision on equations to be taken into account for further computation. With SVD, the $M \times N$ matrix $\mathbf{A}$ can be decomposed as

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T \tag{4.34}$$

where $\mathbf{U}$ is an $M \times M$ and $\mathbf{V}$ an $N \times N$ orthogonal matrix; $\mathbf{S}$ is an $M \times N$ diagonal matrix with elements s_i denoted as singular values. The inversion is then given by

$$\mathbf{A}^{-1} = \mathbf{V}\mathbf{S}^{-1}\mathbf{U}^T \tag{4.35}$$

with

$$\mathbf{S}^{-1} = \begin{pmatrix} 1/s_1 & 0 & \cdots & 0 \\ 0 & 1/s_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/s_m \end{pmatrix} \tag{4.36}$$

The closer the singular values are to zero, the more linearly dependent the related equations are. If singular values are equal to zero (or close to machine precision), the matrix can not be inverted.

Regularization

Truncated Singular Values Decomposition A possible approach to avoid singularity of a matrix is to define a threshold for the smallest allowable singular value and set the corresponding inverted singular value $1/s_i$ in $\mathbf{S}^{-1}$ to zero, which will provide the closest solution in the least square sense. This regularization method is termed truncated singular value decomposition (TSVD). For satellite ranging using 6 base stations, the maximum number of reduced equations should not exceed 3. Based on numerous simulation runs, it turned out that a reduction number of 2 for satellite position estimation and of 1 for satellite velocity estimation is probably the best choice as trade-off between stability and offset.

Iterative Regularized Gauss-Newton The TSVD permits only a very coarse selection of the degree of regularization, i.e. taking an equation into account or not. A finer selection of regularization may be provided by the Levenberg-Marquardt or the iterative regularized Gauss-Newton (IRGN) method [81, 84]. The latter uses Tikhonov regularization for stabilization of the problem. Thus, an iteration step of the IRGN can be stated as

$$\mathbf{x}_{k+1} = \mathbf{x}_k - (\mathbf{J}(\mathbf{x}_k)^T\mathbf{J}(\mathbf{x}_k) + \lambda_k\mathbf{I})^{-1}[\mathbf{J}(\mathbf{x}_k)^T\mathbf{f}(\mathbf{x}_k) + \lambda_k(\mathbf{x}_k - \mathbf{x}^*)] \tag{4.37}$$

Herein λ_k denotes the regularization parameter which can be adjusted in each iteration step. Since for large values of λ_k the solution tends to stay close to $\mathbf{x}^*$, prior knowledge may be incorporated, which for satellite ranging would be the available ephemeris data.

The appropriate selection of the regularization parameter λ_k is a major challenge. In [80] several methods to choose λ_k are described. After numerous simulation runs and comparison of different schemes, it turned out that the TSVD performance is equal or even better, regardless the used selection scheme. For this reason, it is suggested to use the TSVD as long as no better parameter selection is found.

4.3.4 Simulation Results

For assessment of the impact of additional ranging on the geolocation accuracy proper scenarios have to be defined.

European Scenario The geolocation setup remains as introduced previously with monitoring in Vienna, reference in Paris and interferer in Lisbon. The geographic setup can be viewed in Figure 4.31. The six base stations are assumed to be located in Istanbul, Berlin, Stockholm, Moscow, Munich and Harlow, with precise coordinates available. For all base stations the SNR relationships are assumed to be equal to the previously defined interferer setup. In short, the SNR on the main satellite is $\gamma_{s,1} = 15\,\text{dB}$, the overspill from the main to the first adjacent satellite $-50\,\text{dB}$, to the second adjacent satellite $-55\,\text{dB}$ and a traffic signal with an SNR of 15 dB is assumed to be present on the adjacent transponder. Thus the overspilled signal $\tilde{r}_1(t)$ exhibits an SNR of $\gamma_{s,2} = -50\,\text{dB}$ on the first and of $\gamma_{s,2} = -55\,\text{dB}$ on the second adjacent satellite. As introduced in Section 4.2.4, the available ephemeris data correspond to the true values superposed by Gaussian noise, i.e. $\mathcal{N}(0,\sigma_r^2)$ with $\sigma_r = 1\,\text{km}$ for satellite positions and $\mathcal{N}(0,\sigma_v^2)$ with $\sigma_v = 0.1\,\text{m/s}$ for satellite velocities. The ephemeris data are updated using the above differential TSVD approach before solving the geolocation problem. In Figure 4.32 the geolocation improvement using satellite ranging and adjacent carrier subtraction is depicted. It can be seen that for both, TDOA/FDOA and TDOA/TDOA, the increase in geolocation accuracy is almost one order of magnitude. In contrast to TDOA/TDOA, TDOA/FDOA shows still a significant jitter around the true transmitter position.

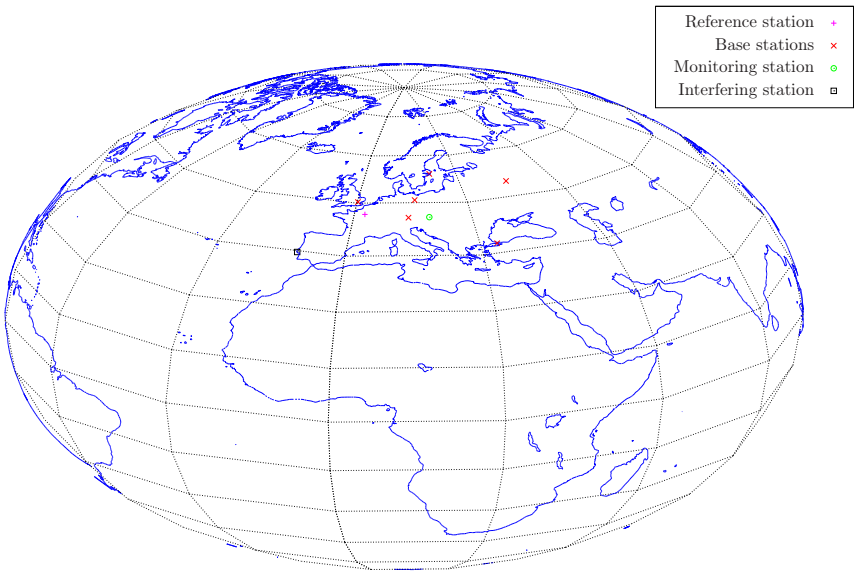


Figure 4.31: European ranging and geolocation scenario

Brazilian Scenario The second scenario is located on the southern hemisphere with monitoring in Rio de Janeiro, reference in Manaus and interferer in Sao Paulo. The geographic setup is illustrated in Figure 4.33. The six required base stations are Brasilia, Recife, Porto Alegre, Fortaleza, Belo Horizonte and Rio Branco. Again it is assumed that the exact location of these base stations is known. The main satellite is located at an azimuth of $\phi_{S1} = -61\,^\circ$, the first adjacent satellite at $\phi_{S2} = -63\,^\circ$ and the second at $\phi_{S3} = -59\,^\circ$. The geolocation results for the second scenario are shown in Figure 4.34. It can be seen that without ranging and adjacent traffic subtraction TDOA/FDOA suffers from severe localization errors. This is mainly due to the changed setup including geometry and velocity vectors. Nevertheless, when updating ephemeris data and removing perturbing adjacent carriers, a considerable increase in accuracy can be achieved. The performance of TDOA/TDOA geolocation shows expected results, since the distance of the interferer from the equator is sufficient.

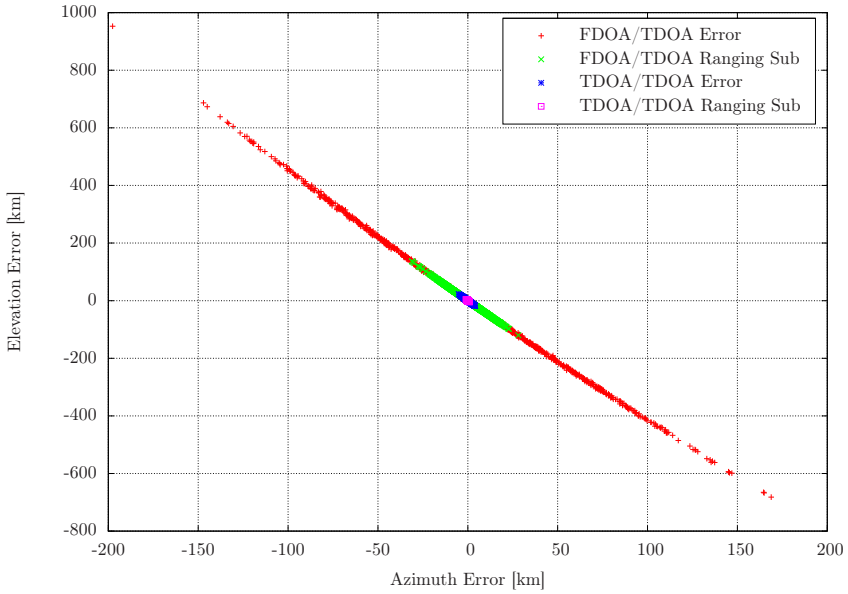


Figure 4.32: European scenario (overall geolocation performance)

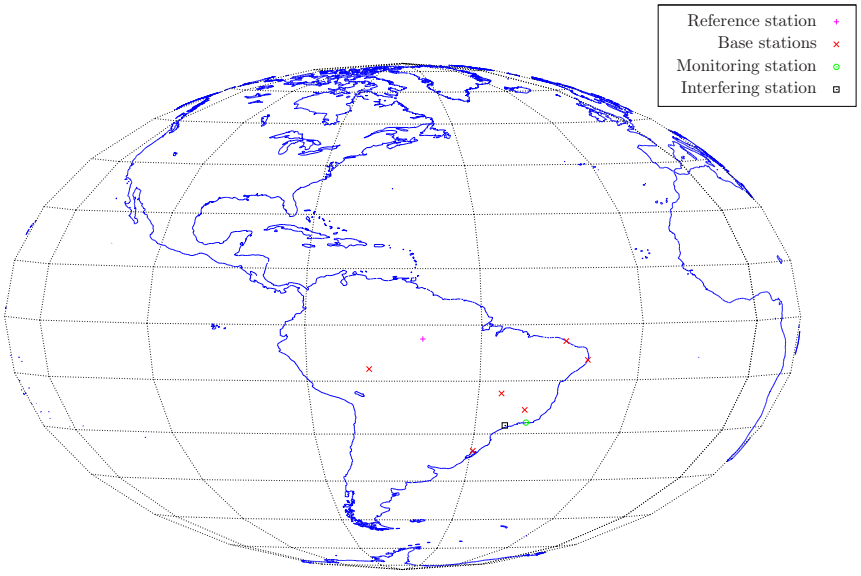


Figure 4.33: Brazilian ranging and geolocation scenario

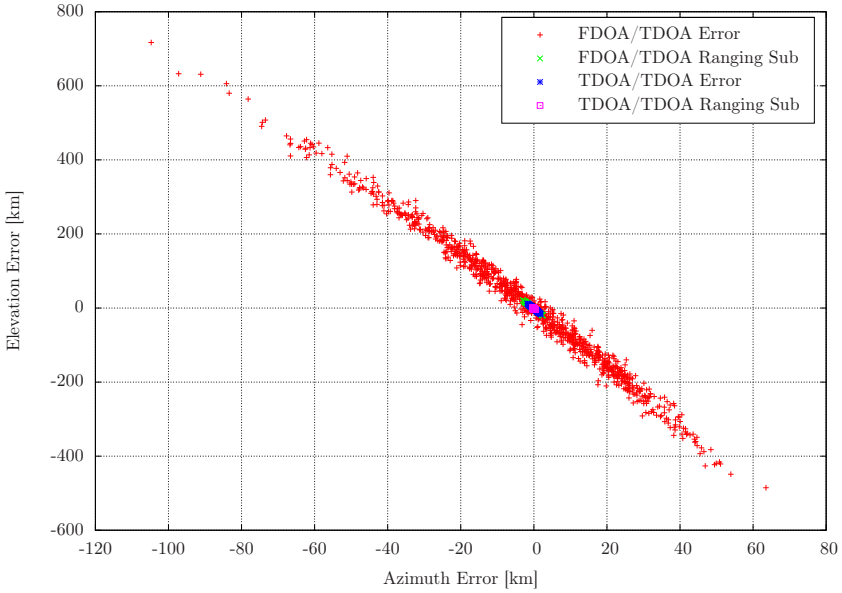


Figure 4.34: Brazilian scenario (overall geolocation performance)

CHAPTER 5

Conclusion and Future Work

5.1 Conclusion

A main objective of satellite operators is an efficient usage of the available resources. Appropriate monitoring is inevitable to allow reliable operation of satellite networks. For supporting the daily work of the providers, an innovative blind demodulation framework is suggested. To arrive at a flexible and low-cost solution, the algorithms should be suited for implementation on a software-defined radio (SDR) platform. The whole framework was integrated and tested on a versatile in-house developed SDR platform. This allows the received signal to be sampled at the intermediate frequency (IF), so that all further signal processing can be performed in the digital domain. Therefore, this work does not only treat the blind demodulation task from a theoretical perspective, but incorporates advices from a practical point of view as well.

After setting the stage by definition of the related signal model, an in-depth overview of the proposed blind demodulation platform is given. For each of the successive demodulation stages, possible procedures are introduced, including adaptation of existing and design of new algorithms. The latter are compared in terms of accuracy and complexity. Additionally, pitfalls that might occur during implementation of the suggested framework are pointed out as well. Topics that were given special attention are summarized in the following.

Only few algorithms on blind estimation of the symbol rate are available in the open literature. Moreover, no performance results exist on non-integer oversampling ratios, which are the case for monitoring systems. It is shown that the performance of the algorithm in [7] can be enhanced significantly by applying narrow band pre-filtering and a symbol rate hint available from the carrier detection stage. Additionally, a timing tracking algorithm might be applied for refinement [8], if the used symbol rate estimator exhibits poor performance.

Established algorithms for non-data-aided (NDA) recovery of the symbol timing are thoroughly compared to recent extensions and a self-developed method in terms of jitter

variance and acquisition performance, since such a comprehensive roundup is not available from the open literature. Representatives of feedforward schemes are the Oerder and Meyr (OM) [25] method and the scheme proposed by the author in [11]. The examined feedback methods include the Gardner synchronizer [26], an extended version [37] and the reinvestigated algorithm by Moeneclaey and Batsele (MB) using only one sample per symbol [36, 10]. It turns out that the OM method delivers the best overall performance, independent of modulation scheme and roll-off factor. When both are available, the most appropriate estimation scheme can be selected to reduce the contained jitter. For signal constellations exhibiting a constant envelope the algorithm proposed by the author outperforms the established methods in the medium-to-high SNR range. Moreover the author verified recently that the S-curve of the MB detector exhibits a sinusoidal shape [11]. Therefore the acquisition trajectories of the MB as well as of the extended Gardner detector, can be described by the analytical model introduced by the author in [9]. This statement is verified by numerous simulation runs.

After successful recovery of the optimum timing instant, the knowledge of the used modulation scheme is necessary to determine the SNR, frequency/phase offset and finally the transmitted data symbols. Since no modulation classification scheme from the open literature supports all the constellations of present and upcoming satellite networks or relies on assumptions that do not hold for non-cooperative demodulation, a classifier [14] is proposed by the author for this purpose. It is demonstrated that the introduced feature-based method provides sufficient performance and the required computational effort remains at an acceptable level.

Although a reliable SNR estimate is available from the spectral pre-estimation procedure, an additional SNR estimation exhibits advantages for several reasons. Three moment-based schemes [47, 50, 51] are combined to obtain reasonable performance for all modulation schemes over the whole SNR range. In this context, multi-level constellations need special treatment by using the SNR pre-estimate to select the appropriate estimator according to SNR level and signal constellation. The performance of the combined estimator is demonstrated by simulations in terms of estimator mean output and normalized mean square error (NMSE).

Surprisingly less publications have been found on NDA synchronization of APSK constellations in the open literature, although the latter exhibit considerable advantages compared to QAM on satellite channels. For this reason an in-depth roundup on this topic is provided within this work.

To assess the performance of the proposed blind demodulation framework, an appropriate measure has to be defined, since evaluation of each stage separately is not suitable due to the large number of parameters. Therefore, the demodulated symbols are remodulated and compared to the original transmitted signal in terms of error vector magnitude (EVM). It is shown that the performance of the framework coincides with the theoretical curve in the medium SNR range. Apart from a jitter floor in the high SNR range, the performance deviates for small SNR values due to remodulation failures. This effect is most pronounced for amplitude phase shift keying (APSK) constellations because of the poor frequency

estimation performance. This motivates further research on this topic.

Two applications are described where the capabilities of blind demodulation techniques can be exploited.

Insufficient cross-polarization discrimination (XPD) may lead to impairments of satellite traffic and thus has to be monitored by appropriate means. Instead of the standard method, which requires stop of service and transmission of a continuous wave signal, an online scheme relying only on passive measurements is introduced. It could be shown that application of blind demodulation techniques leads to a very elegant estimation method with a significant increase in accuracy. Moreover, the proposed algorithms are verified on a demonstrator SDR platform using an elaborate testbed.

Interference may also stem from a ground terminal due to malfunction or by intention. Geolocation schemes relying on TDOA/TDOA [78] or TDOA/FDOA [70] techniques exist for estimation of the source of interference. First of all, the accuracy of the geolocation result depends on TDOA and FDOA estimates. It is shown, that the huge amount of data required for acceptable geolocation accuracy can be decreased by applying blind demodulation techniques. Secondly, errors in the available ephemeris data directly affect the geolocation result. Using a satellite ranging approach and blind demodulation in addition, these errors might be mitigated significantly.

Up to now no statement regarding the execution time of the blind demodulation procedure was made. All the signal processing was performed using a personal computer (PC). Scanning the whole transponder for carriers and performing extraction of the according signal parameters can be accomplished within several seconds. Of course, the execution time depends on the number of present carriers. The above timing requirement is valid for approximately up to 10 carriers. Additional evaluation of the current XPD relationships would raise the necessary time up to few minutes. It is clear that these time requirements are considerably superior to state-of-the-art techniques. The time required for geolocation including satellite ranging should be below 30 minutes. Herein the cross-ambiguity function (CAF) computation is the most time consuming part and depends on the observation length L. A reduction of the execution time could be achieved by using chips dedicated for digital down-conversion (DDC), e.g. GC4016 or GC5016 by Texas Instruments[1]. The latter are available as application-specific integrated circuits (ASICs) on various acquisition cards. Moreover, time consuming tasks might be performed in FPGAs on the digitizer board after DDC. The main drawback of the suggested accelerations is the increasing dependence on a specific platform. Relying on the PC as processing platform remains the most flexible solution. With the increasing number of processing cores, allowing simple parallel computing, the latter becomes more and more attractive from the computational point of view.

1 `http://www.ti.com`

5.2 Future Work

It is obvious that the large number of possible input signal scenarios makes automated blind demodulation a very challenging task. Several simplifications were introduced in this work to arrive at a tractable problem. Since modularity is a key point of the developed framework, new features might be implemented in an easy way allowing simplifications to be revoked.

In a first approach, it was considered that signals are continuous and not of the time-division multiple access (TDMA) type. By appropriate power measurements, TDMA signals may be identified. For bursts of short duration, synchronization using NDA algorithms might get problematic. However, if it is detected that signals rely on a usual communication standard, appropriate synchronization algorithms can be implemented including data-aided (DA) methods as well. Furthermore, suggested forward error correction (FEC) schemes could be added to allow successful remodulation also at very low SNR values.

Although root-raised cosines (RRCOSs) are well established for baseband pulse shaping, other pulses may exhibit interesting properties too [12]. Anyhow, to achieve the most general solution, a stage for estimation of the baseband impulse response could be added.

In Section 3.12 it turned out that the performance of the Rife-Boorstyn algorithm delivers poor results for APSK constellations. Basically, the algorithm only delivers reasonable results above a certain threshold SNR value, which tends to be quite large for APSK signals. An alternative algorithm with satisfactory performance would alleviate the synchronization process significantly.

The accuracy of the geolocation problem directly depends on the quality of the ephemeris data. Thus refining the latter before geolocation might be necessary. It was shown in Section 4.3 that the satellite ranging approach may suffer from stability problems, especially when applying FDOA techniques. Regularization by using truncated singular value decomposition (TSVD) is required to mitigate numerical problems. Since reliable initial values for satellite positions and velocities are available by the operators ephemeris data, this knowledge should be used in some way. By using a Baysian approach for satellite ranging, this prior information could be incorporated in a very elegant manner. Moreover, no error due to linearization would be introduced as well. With the increasing computational power, Baysian inference could be an attractive alternative to the implemented iterative procedure.

APPENDIX A

Implementation Issues

A.1 Polynomial Maximum Interpolation

With data analysis problems, it is often required to determine the position of a maximum of a continuous function available only at discrete points. Since the maximum may be situated between the available data points, interpolation has to be applied to estimate the location and the according data value. For this reason a simple polynomial interpolation algorithm is used [83]. The maximum y_0 is located at $x = 0$ with the adjacent data points y_1 and y_{-1} at $x = 1$ and $x = -1$ respectively. A second-order polynomial

$$y = cx^2 + bx + a \tag{A.1}$$

can be fit to the data points with coefficients according to

$$a = y_0 \tag{A.2}$$

$$b = \frac{y_1 - y_{-1}}{2} \tag{A.3}$$

$$c = \frac{y_1 + y_{-1}}{2} - y_0 \tag{A.4}$$

Setting the first derivative of Equation A.1 to zero delivers then the required location

$$x_{max} = -\frac{b}{2c} \tag{A.5}$$

Inserting this result into Equation A.1 yields the according data value y_{max}.

A.2 Random Numbers

A.2.1 Generation of Uniform Random Numbers

A uniformly distributed random variable denotes numbers which occur with equal probability between two limits, e.g. 0 and 1. The so-called Wichman-Hill algorithm [85] relies on a vector x_0, y_0 and z_0 for initialization, usually termed as seed. When using the same seed for repeated calls of the random generator, the same pseudo-random sequence is generated. It is common practice to derive the seed from the local PC clock to avoid successive use of the same pseudo-random sequence. The seed is chosen according to

$$x_0, y_0, z_0 \in \{1, 2, \ldots, 3 \cdot 10^4\} \tag{A.6}$$

The integers x_i, y_i and z_i are generated by the following recursive procedure:

$$\begin{aligned} x_i &\equiv 171 \cdot x_{i-1} \bmod(30269) \\ y_i &\equiv 172 \cdot y_{i-1} \bmod(30307) \\ z_i &\equiv 170 \cdot z_{i-1} \bmod(30323) \end{aligned} \tag{A.7}$$

With the above generated values the uniformly distributed random variable U_i is calculated according to

$$U_i \equiv \left(\frac{x_i}{30269.0} + \frac{y_i}{30307.0} + \frac{z_i}{30323.0} \right) \bmod(1.0) \tag{A.8}$$

Since uniform random number generation experienced a considerable advance in the last decade, newer schemes with longer periods should be used in future implementations [86].

A.2.2 Generation of Gaussian Random Numbers

The prominent Box-Muller transform [85, 83] requires two uniformly distributed random variables $0 \leq U_1$,$U_2 < 1$ as input. The transform of the two uniformly distributed random variables U_1 and U_2 to two independent random variables X and Y with normal

distribution is performed by

$$\begin{aligned} X &= \mu_x + \sigma_n \sqrt{-2\,\ln(U_1)}\cos(2\,\pi\,U_2) \\ Y &= \mu_y + \sigma_n \sqrt{-2\,\ln(U_1)}\sin(2\,\pi\,U_2) \end{aligned} \tag{A.9}$$

Herein μ_x and μ_y denote the means and σ_n the standard deviation which is set to $\sigma_n = \sigma_x = \sigma_y$ for convenience. A detailed overview of Gaussian random number generators is provided in [87].

A.3 Jacobi Matrices

A.3.1 Geolocation TDOA/FDOA

The equation vector is defined as $\mathbf{f}(\mathbf{x}) = (f_1,f_2,f_3)^T$. Herein f_1 is the TDOA Equation 4.18 for the standard and 4.24 for the differential approach, f_2 the FDOA Equation 4.21 for the standard and 4.25 for the differential approach, respectively, and f_3 the Earth ellipsoid given by Equation 4.22. The Jacobi matrix can then be set up as

$$\mathbf{J} = \begin{pmatrix} \frac{\partial f_1}{\partial x_I} & \frac{\partial f_1}{\partial y_I} & \frac{\partial f_1}{\partial z_I} \\ \frac{\partial f_2}{\partial x_I} & \frac{\partial f_2}{\partial y_I} & \frac{\partial f_2}{\partial z_I} \\ \frac{\partial f_3}{\partial x_I} & \frac{\partial f_3}{\partial y_I} & \frac{\partial f_3}{\partial z_I} \end{pmatrix} = \begin{pmatrix} \mathbf{g}_1^T \\ \mathbf{g}_2^T \\ \mathbf{g}_3^T \end{pmatrix} \tag{A.10}$$

with the vectors $\mathbf{g}_1$, $\mathbf{g}_2$ and $\mathbf{g}_3$ defined as

$$\mathbf{g}_1 = \begin{pmatrix} \frac{x_{S2}-x_I}{l_{I2}} - \frac{x_{S1}-x_I}{l_{I1}} \\ \frac{y_{S2}-y_I}{l_{I2}} - \frac{y_{S1}-y_I}{l_{I1}} \\ \frac{z_{S2}-z_I}{l_{I2}} - \frac{z_{S1}-z_I}{l_{I1}} \end{pmatrix} \tag{A.11}$$

$$\mathbf{g}_2 = \begin{pmatrix} \frac{\dot{x}_{S1}}{l_{I1}} - \frac{(x_{S1}-x_I)\,\mathbf{v}_{\mathrm{S1}}\cdot\mathbf{e}_{\mathrm{I1}}}{l_{I1}^2} - \left(\frac{\dot{x}_{S2}}{l_{I2}} - \frac{(x_{S2}-x_I)\,\mathbf{v}_{\mathrm{S2}}\cdot\mathbf{e}_{\mathrm{I2}}}{l_{I2}^2} \right) \\ \frac{\dot{y}_{S1}}{l_{I1}} - \frac{(y_{S1}-y_I)\,\mathbf{v}_{\mathrm{S1}}\cdot\mathbf{e}_{\mathrm{I1}}}{l_{I1}^2} - \left(\frac{\dot{y}_{S2}}{l_{I2}} - \frac{(y_{S2}-y_I)\,\mathbf{v}_{\mathrm{S2}}\cdot\mathbf{e}_{\mathrm{I2}}}{l_{I2}^2} \right) \\ \frac{\dot{z}_{S1}}{l_{I1}} - \frac{(z_{S1}-z_I)\,\mathbf{v}_{\mathrm{S1}}\cdot\mathbf{e}_{\mathrm{I1}}}{l_{I1}^2} - \left(\frac{\dot{z}_{S2}}{l_{I2}} - \frac{(z_{S2}-z_I)\,\mathbf{v}_{\mathrm{S2}}\cdot\mathbf{e}_{\mathrm{I2}}}{l_{I2}^2} \right) \end{pmatrix} \tag{A.12}$$

$$\mathbf{g}_3 = \begin{pmatrix} \frac{2x_I}{R_A^2} \\ \frac{2y_I}{R_A^2} \\ \frac{2z_I}{R_B^2} \end{pmatrix} \tag{A.13}$$

Herein $\mathbf{r}_{\mathrm{S1}} = (x_{S1},y_{S1},z_{S1})^T$ and $\mathbf{r}_{\mathrm{S2}} = (x_{S2},y_{S2},z_{S2})^T$ correspond to the position vectors of main and adjacent satellite. The velocity vectors are given by $\mathbf{v}_{\mathrm{S1}} = (\dot{x}_{S1},\dot{y}_{S1},\dot{z}_{S1})^T$ and $\mathbf{v}_{\mathrm{S2}} = (\dot{x}_{S2},\dot{y}_{S2},\dot{z}_{S2})^T$ respectively.

A.3.2 Geolocation TDOA/TDOA

The Jacobi Matrix of the TDOA/TDOA approach is very similar to the one of the TDOA/FDOA method. The difference is that the second FDOA line is replaced by an additional TDOA line considering a third satellite. Thus, the vector $\mathbf{g}_2$ can be formulated as

$$\mathbf{g}_2 = \begin{pmatrix} \frac{x_{S3}-x_I}{l_{I3}} - \frac{x_{S1}-x_I}{l_{I1}} \\ \frac{y_{S3}-y_I}{l_{I3}} - \frac{y_{S1}-y_I}{l_{I1}} \\ \frac{z_{S3}-z_I}{l_{I3}} - \frac{z_{S1}-z_I}{l_{I1}} \end{pmatrix} \tag{A.14}$$

A.3.3 Ranging TDOA/FDOA Standard

Satellite Position According to the Equation Set 4.30 the 7×7 Jacobi matrix can be established by

$$\mathbf{J} = \partial\mathbf{f}/\partial\mathbf{x} = \begin{pmatrix} \mathbf{g}_1^T \\ \mathbf{g}_2^T \\ \vdots \\ \mathbf{g}_7^T \end{pmatrix} \tag{A.15}$$

with the vectors $\mathbf{g}_i$ for each base station defined as

$$\mathbf{g}_i = \begin{pmatrix} \frac{x_{S1}-x_{B,i}}{l_{i,1}} + \frac{x_{S1}-x_0}{l_{10}} \\ \frac{y_{S1}-y_{B,i}}{l_{i,1}} + \frac{y_{S1}-y_0}{l_{10}} \\ \frac{z_{S1}-z_{B,i}}{l_{i,1}} + \frac{z_{S1}-z_0}{l_{10}} \\ \frac{x_{S2}-x_{B,i}}{l_{i,2}} + \frac{x_{S2}-x_0}{l_{20}} \\ \frac{y_{S2}-y_{B,i}}{l_{i,2}} + \frac{y_{S2}-y_0}{l_{20}} \\ \frac{z_{S2}-z_{B,i}}{l_{i,2}} + \frac{z_{S2}-z_0}{l_{20}} \\ 1 \end{pmatrix} \tag{A.16}$$

Satellite Velocity From the Equation Set 4.31 the 7×7 Jacobi matrix can be established by

$$\mathbf{J} = \partial\dot{\mathbf{f}}/\partial\dot{\mathbf{x}} = \begin{pmatrix} \dot{\mathbf{g}}_1^T \\ \dot{\mathbf{g}}_2^T \\ \vdots \\ \dot{\mathbf{g}}_7^T \end{pmatrix} \tag{A.17}$$

with the vectors $\dot{\mathbf{g}}_i$ defined as

$$\dot{\mathbf{g}}_i = \begin{pmatrix} -\frac{x_{S1}-x_{B,i}}{l_{i,1}} - \frac{f_i-f_{T1}}{f_i}\frac{x_{S1}-x_0}{l_{10}} \\ -\frac{y_{S1}-y_{B,i}}{l_{i,1}} - \frac{f_i-f_{T1}}{f_i}\frac{y_{S1}-y_0}{l_{10}} \\ -\frac{z_{S1}-z_{B,i}}{l_{i,1}} - \frac{f_i-f_{T1}}{f_i}\frac{z_{S1}-z_0}{l_{10}} \\ \frac{x_{S2}-x_{B,i}}{l_{i,2}} + \frac{f_i-f_{T2}}{f_i}\frac{x_{S2}-x_0}{l_{20}} \\ \frac{y_{S2}-y_{B,i}}{l_{i,2}} + \frac{f_i-f_{T2}}{f_i}\frac{y_{S2}-y_0}{l_{20}} \\ \frac{z_{S2}-z_{B,i}}{l_{i,2}} + \frac{f_i-f_{T2}}{f_i}\frac{z_{S2}-z_0}{l_{20}} \\ \frac{c}{f_i} \end{pmatrix} \tag{A.18}$$

A.3.4 Ranging TDOA/FDOA Differential

Satellite Position According to the Equation Set 4.32 the 6 × 6 Jacobi matrix can be established by

$$\mathbf{J} = \partial\mathbf{f}/\partial\mathbf{x} = \begin{pmatrix} \mathbf{g}_1^T \\ \mathbf{g}_2^T \\ \vdots \\ \mathbf{g}_6^T \end{pmatrix} \tag{A.19}$$

with the vectors $\mathbf{g}_i$ defined as

$$\mathbf{g}_i = \begin{pmatrix} \frac{x_{S1}-x_{B,i}}{l_{i,1}} - \frac{x_{S1}-x_R}{l_{R1}} \\ \frac{y_{S1}-y_{B,i}}{l_{i,1}} - \frac{y_{S1}-y_R}{l_{R1}} \\ \frac{z_{S1}-z_{B,i}}{l_{i,1}} - \frac{z_{S1}-z_R}{l_{R1}} \\ \frac{x_{S2}-x_{B,i}}{l_{i,2}} - \frac{x_{S2}-x_R}{l_{R2}} \\ \frac{y_{S2}-y_{B,i}}{l_{i,2}} - \frac{y_{S2}-y_R}{l_{R2}} \\ \frac{z_{S2}-z_{B,i}}{l_{i,2}} - \frac{z_{S2}-z_R}{l_{R2}} \end{pmatrix} \tag{A.20}$$

Satellite Velocity From the Equation Set 4.33 the 6×6 Jacobi matrix can be established by

$$\mathbf{J} = \partial\dot{\mathbf{f}}/\partial\dot{\mathbf{x}} = \begin{pmatrix} \dot{\mathbf{g}}_1^T \\ \dot{\mathbf{g}}_2^T \\ \vdots \\ \dot{\mathbf{g}}_6^T \end{pmatrix} \tag{A.21}$$

with the vectors $\dot{\mathbf{g}}_i$ defined as

$$\dot{\mathbf{g}}_i = \begin{pmatrix} -\frac{x_{S1}-x_{B,i}}{l_{i,1}} - \frac{f_i-f_{T1}}{f_i}\frac{x_{S1}-x_0}{l_{10}} + \frac{f_R}{f_i}\frac{x_{S1}-x_R}{l_{R1}} + \frac{f_R-f_{T1}}{f_i}\frac{x_{S1}-x_0}{l_{10}} \\ -\frac{y_{S1}-y_{B,i}}{l_{i,1}} - \frac{f_i-f_{T1}}{f_i}\frac{y_{S1}-y_0}{l_{10}} + \frac{f_R}{f_i}\frac{y_{S1}-y_R}{l_{R1}} + \frac{f_R-f_{T1}}{f_i}\frac{y_{S1}-y_0}{l_{10}} \\ -\frac{z_{S1}-z_{B,i}}{l_{i,1}} - \frac{f_i-f_{T1}}{f_i}\frac{z_{S1}-z_0}{l_{10}} + \frac{f_R}{f_i}\frac{z_{S1}-z_R}{l_{R1}} + \frac{f_R-f_{T1}}{f_i}\frac{z_{S1}-z_0}{l_{10}} \\ \frac{x_{S2}-x_{B,i}}{l_{i,2}} + \frac{f_i-f_{T2}}{f_i}\frac{x_{S2}-x_0}{l_{20}} - \frac{f_R}{f_i}\frac{x_{S2}-x_R}{l_{R2}} - \frac{f_R-f_{T2}}{f_i}\frac{x_{S2}-x_0}{l_{20}} \\ \frac{y_{S2}-y_{B,i}}{l_{i,2}} + \frac{f_i-f_{T2}}{f_i}\frac{y_{S2}-y_0}{l_{20}} - \frac{f_R}{f_i}\frac{y_{S2}-y_R}{l_{R2}} - \frac{f_R-f_{T2}}{f_i}\frac{y_{S2}-y_0}{l_{20}} \\ \frac{z_{S2}-z_{B,i}}{l_{i,2}} + \frac{f_i-f_{T2}}{f_i}\frac{z_{S2}-z_0}{l_{20}} - \frac{f_R}{f_i}\frac{z_{S2}-z_R}{l_{R2}} - \frac{f_R-f_{T2}}{f_i}\frac{z_{S2}-z_0}{l_{20}} \end{pmatrix} \tag{A.22}$$

Author's Publication List

1. M. Flohberger, W. Gappmair, and O. Koudelka, "Blind Demodulation Framework for Satellite Signals," in *Proc. 3rd Int. ICST Conf. on Personal Satellite Services (PSATS)*, Malaga, Spain, Feb. 2011, submitted.

2. ——, "Modulation Classifier for Signals Used in Satellite Communications," in *Proc. IEEE 5th Advanced Satellite Mobile Systems (ASMS) Conf.*, Cagliari, Italy, Sep. 2010, pp. 198–202.

3. M. Flohberger, W. Gappmair, and S. Cioni, "Two Iterative Algorithms for Blind Symbol Timing Estimation of M-PSK Signals," in *Proc. IEEE 5th Int. Workshop on Satellite and Space Communications (IWSSC)*, Siena, Italy, Sep. 2009, pp. 8–12.

4. M. Flohberger, W. Gappmair, and O. Koudelka, "Open-Loop Analysis of an Error Detector for Blind Symbol Timing Recovery Using Baud-Rate Samples," in *Proc. IEEE 4th Int. Workshop on Satellite and Space Communications (IWSSC)*, Toulouse, France, Oct. 2008, pp. 176–180.

5. ——, "Improved Self-Noise Performance of Gardner Synchronizers for Bandwidth-Efficient M-PSK Signals," in *Proc. IEEE 3rd Int. Workshop on Satellite and Space Communications (IWSSC)*, Salzburg, Austria, Sep. 2007, pp. 89–91.

6. M. Flohberger, W. Kogler, W. Gappmair, and O. Koudelka, "Symbol Rate Estimation with Inverse Fourier Transforms," in *Proc. IEEE 2nd Int. Workshop on Satellite and Space Communications (IWSSC)*, Madrid, Spain, Sep. 2006, pp. 110–113.

7. W. Gappmair and M. Flohberger, "Error Performance of Coded FSO Links in Turbulent Atmosphere Modeled by Gamma-Gamma Distributions," *IEEE Trans. Wireless. Comm.*, vol. 8, no. 5, pp. 2209–2213, May 2009.

8. W. Gappmair, M. Flohberger, and O. Koudelka, "Moment-Based Estimation of the Signal-to-Noise Ratio for Oversampled Narrowband Signals," in *Advances in Mobile and Wireless Communications*, ser. Lecture Notes in Electrical Engineering, F. István, J. Bitó, and P. Bakki, Eds. Springer Berlin Heidelberg, 2008, vol. 16, ch. 7, pp. 137–148.

9. ——, "Moment-Based Estimation of the Signal-to-Noise Ratio for Oversampled Narrowband Signals," in *Proc. of the 16th IST Mobile and Wireless Communications Summit*, 2007, pp. C4.1–1 – C4.1–4.

10. ——, "Acquisition Performance for NDA Symbol-Timing Recovery with Simple Adaptive First-Order Loops," *Euro. Trans. Telecomms.*, vol. 17, no. 5, pp. 599–605, Apr. 2006.

References

[1] G. Maral, M. Bousquet, and Z. Sun, *Satellite Communications Systems*, 5th ed. New York: Wiley, 2009.

[2] G. Maral, *VSAT Networks*, 2nd ed. West Sussex, England: Wiley, 2003.

[3] B. Evans, M. Werner, E. Lutz, M. Bousquet, G. E. Corazza, G. Maral, R. Rumeau, and E. Ferro, "Integration of Satellite and Terrestrial Systems in Future Multimedia Communications," *IEEE Wireless Commun. Mag.*, vol. 12, no. 4, pp. 72–80, Oct. 2005.

[4] J. Mitola, *Software Radio Architecture: Object-Oriented Approaches to Wireless Systems Engineering*, 1st ed. New York: Wiley, 2000.

[5] B. Evans, *Satellite Communication Systems*, 3rd ed. London, U.K.: Institution of Engineering and Technology, 1999.

[6] M. Flohberger, W. Gappmair, and O. Koudelka, "Blind Demodulation Framework for Satellite Signals," in *Proc. 3rd Int. ICST Conf. on Personal Satellite Services (PSATS)*, Malaga, Spain, Feb. 2011, submitted.

[7] M. Kueckenwaitz, F. Quint, and J. Reichert, "A Robust Baud Rate Estimator for Noncooperative Demodulation," in *Proc. 21st Century Military Communications (MILCOM) Conf.*, vol. 2, Los Angeles, USA, Aug. 2000, pp. 971–975.

[8] M. Flohberger, W. Kogler, W. Gappmair, and O. Koudelka, "Symbol Rate Estimation with Inverse Fourier Transforms," in *Proc. IEEE 2nd Int. Workshop on Satellite and Space Communications (IWSSC)*, Madrid, Spain, Sep. 2006, pp. 110–113.

[9] W. Gappmair, M. Flohberger, and O. Koudelka, "Acquisition Performance for NDA Symbol-Timing Recovery with Simple Adaptive First-Order Loops," *Euro. Trans. Telecomms.*, vol. 17, no. 5, pp. 599–605, Apr. 2006.

[10] M. Flohberger, W. Gappmair, and O. Koudelka, "Open-Loop Analysis of an Error Detector for Blind Symbol Timing Recovery Using Baud-Rate Samples," in *Proc.*

IEEE 4th Int. Workshop on Satellite and Space Communications (IWSSC), Toulouse, France, Oct. 2008, pp. 176–180.

[11] M. Flohberger, W. Gappmair, and S. Cioni, "Two Iterative Algorithms for Blind Symbol Timing Estimation of M-PSK Signals," in *Proc. IEEE 5th Int. Workshop on Satellite and Space Communications (IWSSC)*, Siena, Italy, Sep. 2009, pp. 8–12.

[12] M. Flohberger, W. Gappmair, and O. Koudelka, "Improved Self-Noise Performance of Gardner Synchronizers for Bandwidth-Efficient M-PSK Signals," in *Proc. IEEE 3rd Int. Workshop on Satellite and Space Communications (IWSSC)*, Salzburg, Austria, Sep. 2007, pp. 89–91.

[13] ETSI EN 302 307, "Second Generation Framing Structure, Channel Coding and Modulation Systems for Broadcasting, Interactive Services, News Gathering and other Broadband Satellite Applications," Mar. 2005.

[14] M. Flohberger, W. Gappmair, and O. Koudelka, "Modulation Classifier for Signals Used in Satellite Communications," in *Proc. IEEE 5th Advanced Satellite Mobile Systems (ASMS) Conf.*, Cagliari, Italy, Sep. 2010, pp. 198–202.

[15] W. Gappmair, M. Flohberger, and O. Koudelka, "Moment-Based Estimation of the Signal-to-Noise Ratio for Oversampled Narrowband Signals," in *Advances in Mobile and Wireless Communications*, ser. Lecture Notes in Electrical Engineering, F. István, J. Bitó, and P. Bakki, Eds. Springer Berlin Heidelberg, 2008, vol. 16, ch. 7, pp. 137–148.

[16] ——, "Moment-Based Estimation of the Signal-to-Noise Ratio for Oversampled Narrowband Signals," in *Proc. of the 16th IST Mobile and Wireless Communications Summit*, 2007, pp. C4.1–1 – C4.1–4.

[17] S. Cheung, "A (4,12) Circular Signal Set for Satellite Transmission," in *Proc. Second IEE National Conf. on Telecommunications*, York, UK, Apr. 1989, pp. 122–127.

[18] F. M. Gardner, "Demodulator Reference Recovery Techniques Suited for Digital Implementation," *ESA Final Report: Contract No. 6847/86/NL/DG*, Aug. 1988.

[19] U. Mengali and A. N. D'Andrea, *Synchronization Techniques for Digital Receivers*. New York: Plenum Press, 1997.

[20] J. Treichler, M. Larimore, and J. Harp, "Practical Blind Demodulators for High-Order QAM Signals," *Proc. IEEE*, vol. 86, no. 10, pp. 1907–1926, Oct. 1998.

[21] P. D. Welch, "The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging Over Short, Modified Periodograms," *IEEE Trans. Audio Electroacoust.*, vol. 15, no. 2, pp. 70–73, Jun. 1967.

[22] E. B. Hogenauer, "An Economical Class of Digital Filters for Decimation and Interpolation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 29, no. 2, pp. 155–162, Apr. 1981.

[23] Z. Yu, Y. Q. Shi, and W. Su, "Symbol-Rate Estimation Based on Filter Bank," in *Proc. IEEE Int. Symposium on Circuits and Systems (ISCAS)*, Kobe, Japan, May 2005, pp. 1437–1440.

[24] P. Ciblat, P. Loubaton, E. Serpedin, and G. Giannakis, "Asymptotic Analysis of Blind Cyclic Correlation-Based Symbol-Rate Estimators," *IEEE Trans. Inform. Theory*, vol. 48, no. 7, pp. 1922–1934, Jul. 2002.

[25] M. Oerder and H. Meyr, "Digital Filter and Square Timing Recovery," *IEEE Trans. Commun.*, vol. 36, pp. 604–612, May 1988.

[26] F. M. Gardner, "A BPSK/QPSK Timing-Error Detector for Sampled Receivers," *IEEE Trans. Commun.*, vol. 34, pp. 423–429, May 1986.

[27] W. Gappmair, S. Cioni, G. E. Corazza, and O. Koudelka, "Symbol-Timing Recovery with Modified Gardner Detectors," in *Proc. Int. Symposium on Wireless Communication Systems (ISWCS)*, Siena, Italy, Sep. 2005, pp. 831–834.

[28] R. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing.* New Jersey: Englewood Cliffs, 1983.

[29] F. M. Gardner, "Interpolation in Digital Modems - Part I: Fundamentals," *IEEE Trans. Commun.*, vol. 41, pp. 501–507, Mar. 1993.

[30] J. O. Smith. (2002, Jan.) Digital Audio Resampling Home Page. [Online]. Available: https://ccrma.stanford.edu/~jos/resample/

[31] L. Erup, F. M. Gardner, and R. A. Harris, "Interpolation in Digital Modems - Part II: Implementation and Performance," *IEEE Trans. Commun.*, vol. 41, pp. 998–1008, Jun. 1993.

[32] H. Xu, Y. Zhou, and Z. Huang, "Blind Roll-Off Factor and Symbol Rate Estimation Using IFFT and Least Squares Estimator," in *Int. Conf. on Wireless Communications, Networking and Mobile Computing (WICOM)*, Shanghai, China, Sep. 2007, pp. 1052–1055.

[33] B. Sklar, *Digital Communications*, 2nd ed. New Jersey: Prentice Hall, 2001.

[34] W. Gappmair, M. Flohberger, and O. Koudelka, "Moment-Based Estimation of the Signal-to-Noise Ratio for Oversampled Narrowband Signals," in *Proc. 16th IST Mobile and Wireless Communications Summit*, Budapest, Hungary, Jul. 2007, pp. C4.1–1 – C4.1–4.

[35] M. Flohberger, "Acquisition of Symbol Timing in TDMA Satellite Modems," Master's thesis, Graz University of Technology, 2004, in German.

[36] M. Moeneclaey and T. Batsele, "Carrier-independent NDA Symbol Synchronization for M-PSK, Operating at Only One Sample Per Symbol," in *Proc. IEEE Global Telecommunications (GLOBECOM) Conf.*, vol. 1, San Diego, USA, Dec. 1990, pp. 594–598.

[37] W. Gappmair, S. Cioni, G. E. Corazza, and O. Koudelka, "Extended Gardner Detector for Improved Symbol-Timing Recovery of M-PSK Signals," *IEEE Trans. Commun.*, vol. 54, no. 11, pp. 1923–1927, Nov. 2006.

[38] W. Gappmair, "Open-Loop Analysis and Self-Noise Performance of an NDA Synchroniser for Carrier-Independent Symbol Timing Recovery at Baud Rate," *Euro. Trans. Telecomms.*, vol. 21, no. 1, pp. 73–78, Jan. 2010.

[39] O. A. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, "Blind Modulation Classification: A Concept Whose Time Has Come," in *Proc. IEEE Sarnoff Symposium on Advances in Wired and Wireless Communication*, Apr. 2005, pp. 223–228.

[40] ——, "Survey of Automatic Modulation Classification Techniques: Classical Approaches and New Trends," *IET Commun.*, vol. 1, no. 2, pp. 137–156, Apr. 2007.

[41] Z. Yu, Y. Shi, and W. Su, "A Blind Carrier Frequency Estimation Algorithm for Digitally Modulated Signals," in *Proc. IEEE Military Communications (MILCOM) Conf.*, vol. 1, Monterey, USA, Oct. 2004, pp. 48–53.

[42] A. Swami and B. M. Sadler, "Hierarchical Digital Modulation Classification Using Cumulants," *IEEE Trans. Commun.*, vol. 48, pp. 416–429, Mar. 2000.

[43] J. M. Mendel, "Tutorial on Higher-Order Statistics (Spectra) in Signal Processing and System Theory: Theoretical Results and Some Applications," *Proc. IEEE*, vol. 79, no. 3, pp. 278–305, Mar. 1991.

[44] P. Marchand, "Detection et Reconnaissance de Modulations Numeriques a l'Aide des Statistiques Cycliques d'Ordre Superieur," Ph.D. dissertation, Institut National Polytechnique, Grenoble, France, 1992.

[45] Q. Shi, Y. Gong, and Y. L. Guan, "Asynchronous Classification of High-Order QAMs," in *Proc. IEEE Wireless Communications and Networking Conf.*, Las Vegas, USA, Apr. 2008, pp. 1188–1193.

[46] M. R. Mirarab and M. A. Sobhani, "Robust Modulation Classification for PSK/QAM/ASK Using Higher-Order Cumulants," in *Proc. 6th Int. Conf. on Information, Communications and Signal Processing*, Singapore, Dec. 2007, pp. 1–4.

[47] D. R. Pauluzzi and N. C. Beaulieu, "A Comparison of SNR Estimation Techniques for the AWGN Channel," *IEEE Trans. Commun.*, vol. 48, pp. 1681–1691, Oct. 2000.

[48] W. Dai, Y. Wang, and J. Wang, "Joint Power Estimation and Modulation Classification Using Second- and Higher Statistics," in *Proc. IEEE Wireless Communications and Networking Conf.*, vol. 1, Orlando, USA, Mar. 2002, pp. 155–158.

[49] R. López-Valcarce and C. Mosquera, "Sixth-Order Statistics-Based Non-Data-Aided SNR Estimation," *IEEE Commun. Lett.*, vol. 11, no. 4, pp. 351–353, Apr. 2007.

[50] W. Gappmair and O. Koudelka, "Moment-based SNR Estimation of Signals with Non-Constant Envelope," in *Proc. 3rd Advanced Satellite Mobile Systems (ASMS) Conf.*, Hersching, Germany, May 2006, pp. 301–304.

[51] M. Álvarez Díaz, R. López-Valcarce, and C. Mosquera, "SNR Estimation for Multilevel Constellations Using Higher-Order Moments," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1515–1526, Mar. 2010.

[52] López-Valcarce. (2007) Mathematica and MATLAB Code Supporting Some of the Derivations in the Paper. University of Vigo. [Online]. Available: http://www.gts.tsc.uvigo.es/~valcarce/snr_hos.html

[53] M. Morelli and U. Mengali, "Feedforward Frequency Estimation for PSK: A Tutorial Review," *Euro. Trans. Telecomms.*, vol. 9, no. 2, pp. 103–116, Sep. 1998.

[54] D. C. Rife and R. R. Boorstyn, "Single-Tone Parameter Estimation from Discrete-Time Observations," *IEEE Trans. Inform. Theory*, vol. 20, pp. 591–598, Sep. 1974.

[55] M. Luise and R. Reggiannini, "Carrier Frequency Recovery in All-Digital Modems for Burst-Mode Transmissions," *IEEE Trans. Commun.*, vol. 43, pp. 1169–1178, Feb./Mar./Apr. 1995.

[56] M. Morelli, A. N. D'Andrea, and U. Mengali, *Broadband Wireless Communications.* Springer-Verlag, 1998, ch. Feedforward Estimation Techniques for Carrier Recovery in 16-QAM Modulation.

[57] A. J. Viterbi and A. M. Viterbi, "Nonlinear Estimation of PSK-Modulated Carrier Phase with Application to Burst Digital Transmission," *IEEE Trans. Inform. Theory*, vol. 29, pp. 543–551, Jul. 1983.

[58] H. Sari and S. Moridi, "New Phase and Frequency Detectors for Carrier Recovery in PSK and QAM Systems," *IEEE Trans. Commun.*, vol. 36, no. 9, pp. 1035–1043, Sep. 1988.

[59] T. Duryea, I. Sari, and E. Serpedin, "Blind Carrier Recovery for Circular QAM Using Nonlinear Least-squares Estimation," *Digital Signal Processing*, vol. 16, no. 4, pp. 358–368, Jul. 2006.

[60] Y. Wang, E. Serpedin, and P. Ciblat, "Optimal Blind Nonlinear Least-Squares Carrier Phase and Frequency Offset Estimation for General QAM Modulations," *IEEE Trans. Wireless Commun.*, vol. 2, no. 5, pp. 1040–1054, Sep. 2003.

[61] F. Rice, M. Rice, and B. Cowley, "A New Algorithm for 16-QAM Carrier Phase Estimation Using QPSK Partitioning," *Digital Signal Processing*, vol. 12, no. 1, pp. 77–86, Jan. 2002.

[62] E. Casini, R. De Gaudenzi, and A. Ginesi, "DVB-S2 Modem Algorithms Design and Performance Over Typical Satellite Channels," *Int. Journal of Satellite Communications and Networking*, vol. 22, no. 3, pp. 281–318, Sep. 2004.

[63] R. De Gaudenzi, A. Guillen, and A. Martinez, "Performance Analysis of Turbo-Coded APSK Modulations Over Nonlinear Satellite Channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 9, pp. 2396–2407, Sep. 2006.

[64] W. Gappmair, S. Cioni, A. Vanelli-Coralli, and O. Koudelka, "Exact Analysis of Different Detector Algorithms for NDA Carrier Phase Recovery of 16/32-APSK Signals," *Euro. Trans. Telecomms.*, vol. 21, no. 2, pp. 154–166, Mar. 2010.

[65] W. Gappmair, "Extended Analysis of Viterbi-Viterbi Synchronizers," *Euro. Trans. Telecomms.*, vol. 16, no. 2, pp. 151–155, Feb./Mar. 2005.

[66] M. McKinley, K. Remley, M. T. Myslinski, J. Kenney, D. Schreurs, B. Nauwelaers, and S. Cheung, "EVM Calculation for Broadband Modulated Signals," in *Proc. Automatic RF Techniques Group Conf.*, Orlando, USA, Nov. 2004, pp. 45–52.

[67] R. A. Shafik, M. S. Rahman, and A. H. M. R. Islam, "On the Extended Relationships Among EVM, BER and SNR as Performance Metrics," in *Proc. 4th International Conf. on Electrical and Computer Engineering*, Dhaka, Bangladesh, Dec. 2006, pp. 408–411.

[68] A. Papoulis, *Probability, Random Variables and Stochastic Processes*, 3rd ed. New York: McGraw-Hill Companies, 1991.

[69] S. Stein, "Algorithms for Ambiguity Function Processing," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 29, no. 3, pp. 588–599, Jun. 1981.

[70] D. P. Haworth, N. G. Smith, R. Bardelli, and T. Clement, "Interference Localisation for Eutelsat Satellites - The First European Transmitter Location System," *Int. Journal of Satellite Communications*, vol. 15, no. 4, pp. 155–183, Jul./Aug. 1997.

[71] D. P. Haworth, "Locating the Source of an Unknown Signal," Patent 6018312, Jan. 2000.

[72] J. J. Johnson, "Implementing the Cross-Ambiguity Function and Generating Geometry-Specific Signals," Master's thesis, Naval postgraduate school, Sep. 2001.

[73] C. L. Yatrakis, "Computing the Cross Ambiguity Function - A Review," Master's thesis, Binghamton University, State University of New York, Jan. 2005.

[74] W. Gappmair, "Satellite-Based Interferer Localization Using TDOA and FDOA Techniques," ESA, Feasibility Study, 2006.

[75] R. Aster, B. Borchers, and C. Thurber, *Parameter Estimation and Inverse Problems (International Geophysics)*. Academic Press, Jan. 2005.

[76] P. C. Chestnut, "Emitter Location Accuracy Using TDOA and Differential Doppler," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 18, no. 2, pp. 214–78, Mar. 1982.

[77] R. J. Ulman and E. Geraniotis, "Wideband TDOA/FDOA Processing Using Summation of Short-Time CAF's," *IEEE Trans. Signal Process.*, vol. 47, no. 12, pp. 3193–3200, Dec. 1999.

[78] W. W. Smith and P. G. Steffes, "Time Delay Techniques for Satellite Interference Location System," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 25, no. 2, pp. 224–231, Mar. 1989.

[79] K. C. Ho and Y. T. Chan, "Solution and Performance Analysis of Geolocation by TDOA," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 29, no. 4, pp. 1311–1322, Oct. 1993.

[80] P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 1998.

[81] J. Nocedal and S. J. Wright, *Numerical Optimization*. New York, USA: Springer, 1999.

[82] D. P. Haworth, "Satellite Ephemeris Error," Patent WO2007/113475 A1, Oct. 2007.

[83] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, 2nd ed. Cambridge: Cambridge University Press, 1992.

[84] M. Meyer, "Iterative Lösung nichtlinearer inverser Probleme," Master's thesis, Technische Universität Chemnitz, 2006, in German.

[85] M. C. Jeruchim, P. Balaban, and K. S. Shanmugan, *Simulation of Communication Systems*. New York: Plenum Press, 1992.

[86] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, 3rd ed. Cambridge: Cambridge University Press, 2007.

[87] D. B. Thomas, W. Luk, P. H. Leong, and J. D. Villasenor, "Gaussian Random Number Generators," *ACM Comput. Surv.*, vol. 39, no. 4, pp. 11:1–11:38, 2007.