On dual-polarization signalling techniques in satellite communications

Lionel Arend

Édition Lionel Arend

UNIVERSITÉ DU LUXEMBOURG

PhD-FSTC-2015-19
The Faculty of Sciences, Technology and Communication

DISSERTATION

Defence held on 27/04/2015 in Luxembourg

to obtain the degree of

DOCTEUR DE L'UNIVERSITÉ DU LUXEMBOURG EN SCIENCES DE L'INGÉNIEUR

by

Lionel Arend

Born on 18 September 1986 in Luxembourg (Luxembourg)

ON DUAL-POLARIZATION SIGNALLING TECHNIQUES IN SATELLITE COMMUNICATIONS

Dissertation defence committee

Dr. Thomas Engel, chairman
Professor, Université du Luxembourg

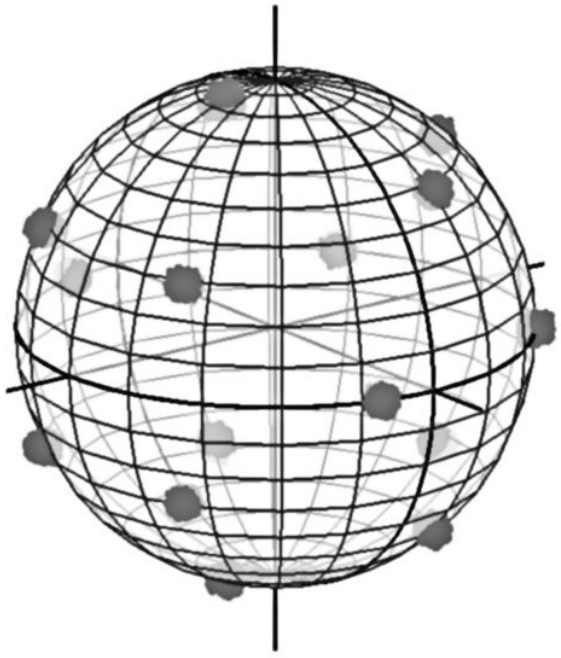
Dr. Claude Oestges, vice-chairman
Professor, Université Catholique de Louvain

Dr.-Ing. Michel Marso, dissertation supervisor
Professor, Université du Luxembourg

Dr.-Ing. Jens Krause
Senior Manager, Systems Engineering, SES S.A.

Dr. François Horlin
Professor, Université Libre de Bruxelles

Mr. Ray Sperber, advisor
Senior Manager, Systems Engineering, SES S.A.



Université du Luxembourg
Faculté des Sciences, de la Technologie et de la Communication

Doctoral dissertation

On dual-polarization signalling techniques in satellite communications

Lionel Arend

Édition Lionel Arend

The polarization-signalling project was funded by the

Fonds National de la Recherche (FNR)
6, rue Antoine de Saint-Exupéry
L-1017 Luxembourg,

under the AFR grant scheme (grant number 3063264). The following organizations were involved in the project's execution:

Université du Luxembourg
6, rue Richard Coudenhove-Kalergi
L-1359 Luxembourg

and

SES S.A.
Château de Betzdorf
L-6815 Betzdorf.

The author can be contacted via electronic mail at
lionel.arend@gmx.net

ISBN 978-99959-0-159-2

Imprimé au Luxembourg en cinquante exemplaires

Mengen Elteren Claude, Jean an Catherine,
menge Bridder Olivier a Florian,
a menger léiwer Frëndin Anastasia

In satellite broadcasting and broadband communications, two orthogonal polarizations can be employed for signalling. Typically, two independent carrier signals are transmitted on the horizontal and vertical linear polarizations. In this way the available frequency spectrum is used twice. Contrary to this independent use, the present work investigates the combined use of both polarizations for signalling and the possibly associated benefits.

Two different modulation techniques are investigated. The first, polarization shift keying, modulates the state of polarization of the electromagnetic wave to convey information. It is based on the phase and amplitude relationship between both carriers and is advantageous for channels strongly impaired by phase noise. The second technique defines a four-dimensional signal space for the combined modulation of amplitude and phase on both carriers. It was initially introduced as 4-quadrature signalling and permits transmission with improved power efficiency. Both techniques come from optical fibre communications, where similar propagation conditions exist.

The conducted investigation is based on an analysis, simulations and the experimental verification using a specially developed demonstration transmitter and receiver. The simulations compare the power efficiency of the proposed modulation techniques with that of alternatives currently in use and analyse their behaviour upon disturbance by phase noise or non-linear distortion. The demonstrator is used to validate the findings in experimental measurements and to confirm the feasibility of both proposed techniques through a live satellite transmission.

En télédiffusion et dans les télécommunications à haut débit par satellite, deux polarisations orthogonales peuvent être employées pour la transmission. Habituellement, deux porteuses indépendantes sont transmises sur les polarisations linéaires horizontale et verticale. Cela rend possible une double exploitation des gammes de fréquences disponibles. Contrairement à une telle utilisation indépendante, le présent travail étudie l'utilisation combinée de ces deux polarisations pour les télécommunications ainsi que les éventuels avantages associés.

Deux techniques de modulations sont soumises à l'examen : la première est la modulation de polarisation. Elle se sert de l'état de polarisation de l'onde électromagnétique pour transporter de l'information. Elle se base ainsi sur les relations de phase et d'amplitude entre les deux porteuses et se montre particulièrement avantageuse pour les canaux fortement dégradés par du bruit de phase. La seconde technique définit un espace quadridimensionnel dans lequel la modulation d'amplitude et de phase des deux porteuses est effectuée de façon conjointe. Nommée modulation 4-quadratures, elle permet une transmission avec une efficacité énergétique élevée. Ces deux techniques ont pour origine le domaine des télécommunications par fibre optique, pour lequel les conditions de propagation sont similaires.

L'étude s'appuie sur une analyse préalable, des simulations et leur vérification expérimentale à l'aide de dispositifs de transmission et de réception développés spécialement à cet effet. Les simulations comparent l'efficacité énergétique des techniques proposées à celle d'alternatives utilisées actuellement et analysent leur comportement lors de perturbations par du bruit de phase et de la distorsion non-linéaire. Les résultats obtenus sont validés au moyen de mesures expérimentales avec les modèles d'exposition. Finalement, la faisabilité de ces deux techniques est démontrée lors d'une transmission en direct par satellite.

Im Satellitenrundfunk und in der satellitengestützten Breitbandkommunikation können zwei orthogonal zueinanderstehende Polarisationen zur Übertragung genutzt werden. Im Allgemeinen werden dazu zwei unabhängige Trägersignale auf der linear waagerechten und senkrechten Polarisation gesendet. Die verfügbaren Frequenzbereiche werden somit doppelt genutzt. Im Gegensatz zu einer solchen unabhängigen Nutzung untersucht die vorliegende Arbeit die gemeinsame Nutzung beider Polarisationen zur Nachrichtenübertragung sowie der damit einhergehenden, möglichen Vorteile.

Es werden zwei unterschiedliche Modulationsverfahren erforscht. Ersteres, die Polarisationsumtastung, verwendet den Polarisationszustand der elektromagnetischen Welle als eigentlichen Informationsträger. Sie beruht somit auf dem Phasen- und Amplitudenverhältnis beider Träger und ist besonders vorteilhaft in stark von Phasenrauschen gestörten Kanälen. Das zweite Verfahren definiert einen vierdimensionalen Signalraum in welchem die Phase und Amplitude beider Träger im Verbund moduliert werden. Die ursprünglich als 4-Quadratur Signalgebung eingeführte Technik erlaubt eine höhere Leistungseffizienz bei der Übertragung. Beide Verfahren kommen aus der optischen Übertragungstechnik, wo vergleichbare Ausbreitungsbedingungen vorherrschen.

Die Untersuchung beruht auf einer vorhergehenden Analyse, Simulationen und der experimentellen Überprüfung mithilfe einer eigens zu diesem Zweck entwickelten Sende- und Empfangsvorrichtung. Die Simulationen vergleichen die Leistungseffizienz der vorgeschlagenen Verfahren mit der gegenwärtig genutzter Alternativen und analysieren das Verhalten im Fall von Beeinträchtigungen durch Phasenrauschen und nichtlineare Verzerrung. Die Simulationsresultate werden durch experimentelle Messungen am Vorführmodell bestätigt. Zusätzlich wird für beide Verfahren mittels einer Satelliten-Direktübertragung die Machbarkeit demonstriert.

Для вещания и широкополосной передачи через спутники могут быть использованы две ортогональные поляризации. В общей практике передача двух независимых несущих сигналов осуществляется по горизонтальной и вертикальной линейным поляризациям. Таким образом, становится возможным двойное использование доступных диапазонов частот. В противоположность такому независимому применению, настоящее исследование посвящено изучению совместного применения обоих поляризаций в электросвязи, а так же его возможных преимуществ.

Две техники модуляции вынесены на рассмотрение: первой является модуляция поляризации. В качестве носителя информации она использует состояние поляризации электромагнитной волны. Базируясь на отношении фаз и амплитуд двух несущих, она показывает себя особенно выгодной при наличии в канале фазового шума. Вторая техника определяет четырёхмерное пространство, в котором модуляция амплитуды и фазы двух несущих совершается комбинировано. Введённая под названием 4-квадратурной модуляции, она делает возможной передачу с повышенной энергетической эффективностью. Две описанные выше техники принадлежат области волоконно-оптической связи, условия распространения в которой являются сопоставимыми.

Исследование основывается на предварительном анализе, симуляциях и их экспериментальной апробации при помощи специально разработанных приборов передачи и приёма. Симуляции служат для сравнения энергетической эффективности предложенных техник с используемыми в настоящее время альтернативами и анализируют их поведение во время нарушений фазовым шумом и нелинейным искажением. Результаты подтверждаются путём использования образцов в ходе экспериментальных измерений. Кроме этого, реализуемость двух предложенных техник продемонстрирована посредством прямой трансляции через спутник.

Preface

Ray Sperber, an engineer working in the technical department for the satellite operator SES, developed the initiative to employ polarization shift keying in satellite communication links. Through a previous work experience, where I had established good contacts with the company, I heard about the project and enthusiasticly offered my participation. The University of Luxembourg, to which SES entertains good relationships, was a natural academic partner for this kind of research project and Michel Marso, professor in telecommunication technology, showed great interest and agreed to join in as academic counsel and as my thesis supervisor.

The project was set up as a public-private partnership between the University and SES, with responsibilities equally attributed. The laboratories of both institutions provided equipment and helped with the development of the prototypes. SES committed to provide a transponder simulator and satellite capacity for my experiments. Financial support came from the Luxembourg National Research Fund in the form of an AFR (aide à la formation recherche) grant.

Although this dissertation bears my name alone, the project would not have been possible without the contributions of colleagues and friends to the planning and execution of the work. Throughout the three years the project lasted, I could count on the help and advice of many.

First of all, I would like to thank my parents, who work hard and have unfailingly supported me during my university curriculum, permitting me to pursue the studies of my choice, without any financial worries.

I would like to thank Ray Sperber and Michel Marso for having offered to me the possibility to be part in this project and for agreeing to arrange it as a doctoral thesis. I thank them as well for granting me a lot of independence in my work, while having kept a benevolent eye on my progress to keep me well on track and on time. By their great support, they managed to keep me going even at times when my courage failed a little.

A special thank goes to Jens Krause. During many long and interesting discussions, Jens shared with me his practical knowledge on receiver implementations and transmission theory and helped me with the accurate interpretation of the phenomena and problems I encountered.

Ray Sperber and Achim Kleine have acquainted me with the satellite transponder and the behaviour of travelling wave tube amplifiers. Achim has also per-

mitted me to use his amplifier model for linearized tubes.

The experiments would not have been possible without the help of Edward Cardew. Edward introduced me to the transponder simulator and suggested practical solutions for setting up the measurements. He also helped me to select and assemble the equipment I required.

The satellite experiments were prepared with the support of Ian Hesketh and his team, notably Aiden O'Dwyer, Pascal Stemper and Stefan Nalbach. Many others as well have added their contribution to this work, which would have been a lot harder without the valuable engineering knowledge I could benefit from.

I also would like to thank Thomas Engel for sharing his strategic views on the project and its completion. Our section head, Carlos Vieira, deserves the credit for keeping my back free of all corporate matters, while at the same time allowing my personal integration into the systems engineering team at SES.

An important matter during these years was my project to learn the Russian language. In this context I thank the Russian Cultural and Scientific Centre in Luxembourg for the opportunity to two unforgettable linguistic stays in Russia, during which, among others, I was able to visit the A. S. Popov radio museum in Ekaterinbourg. I also thank Ray, Michel and Tom Jakobs from the FNR for their unproblematic acquiescence to these stays. My learning efforts were honoured when, thanks to them, I met my wonderful young lady, Anastasia Belotserkovskaia. It is also with her help only that I was able to translate the abstract of this dissertation into Russian.

Finally, my thanks appertain to Etienne Bischops, Nathalie and Ian Hamill, Pierre Schumann, Ray Sperber and Jens Krause, for they have accepted the tedious task of proofreading my final manuscript.

Contents

1. Introduction

> [...] I can express the hope, that my instrument, after further refinement, can be employed for the transmission of signals over distances with the help of fast electric oscillations [...]
>
> *(Alexander Stepanovich Popov, 1896, [62])*

1.1. Background

In a bandpass communications system, the output signal of a high-frequency oscillator is modulated to carry information. Modulation, i.e. the impression of an information message, is accomplished by varying the oscillator's phase, frequency or amplitude. This modulated carrier signal is then transmitted, either over an antenna or through a cable, to a receiver. The receiver performs a demodulation and recovers the initial message. Different modulation techniques are used in various frequency bands of the electromagnetic spectrum, depending on the technology and the application. In broadband data transmission and digital broadcasting over geostationary satellites, the phase and amplitude are modulated in predefined discrete levels [27, 28]. The respective modulation techniques are called phase shift keying or amplitude and phase shift keying.

The emergence of optical communications, modulating light-waves in fibre cables, provided a formidable breakthrough. Due to the extremely high frequency of light, the bandwidth available for information transmission rose by orders of magnitude. In the early days of fibre communications very simple modulation techniques were used, like two-level intensity modulation [12]. Technical challenges and the high signalling speed prevented the use of the spectrally more efficient modulation techniques used in the radio-frequency regions of the spectrum.

In optical communications, multilevel modulation, conveying more than a single bit per transmitted information symbol, came within reach after the discovery that polarization, a property of light, is conserved in certain fibres (specifically in single-mode fibres); when injecting polarized light into such a fibre, polarized light is detected at the other end. More than that, when light of two different polarizations is injected, these are clearly discernible at the other end of the fibre. The idea to modulate the state of polarization, instead of phase or amplitude, was

born. The modulation technique that emanated from the rigorous development of this idea is called *Polarization Shift Keying* (PolSK, [6]). It elegantly solves technical challenges associated with the fibre and allows the transmission of several bits per symbol.

One of these challenges is the random phase variation of the light carrier in an emitting laser diode. This *phase noise* renders coherent, i.e. phase accurate, detection very difficult. Polarization shift keying circumvents this by modulating the phase difference between two components suffering from equal phase noise. This method provides not only complete immunity to phase noise, but also allows direct detection of the signal, without the precise knowledge of the carrier frequency.

A second technical problem is also solved: An optical fibre shows non-linear behaviour when carrying an intensity-modulated signal, leading to signal distortion. For PolSK, the power is split variably between two polarizations, according to a predefined symbol constellation. In this way, the fibre carries constant power while the modulation only modifies its allocation to either one or the other polarization. The non-linearity is thereby avoided [4].

As well as for light, the property of polarization also exists for other types of electromagnetic waves. In terrestrial microwave and in satellite communications from geostationary orbit, polarization is employed as well, although in a different way. By transmitting two signals on *orthogonal* polarizations, for instance on the linear horizontal and the linear vertical, these signals can be separated again at the reception site. This technique is called *polarization division multiplexing* and allows the available frequency spectrum to be used twice.

Today, advances in stabilizing laser diodes and improved processing possibilities allow coherent signalling for optical communications as well. The use of pertinent modulation types, like phase shift keying, has become possible together with the polarization division multiplex already applied in satellite systems. A further step is to consider both orthogonal polarizations as components of the same signal and to modulate them in a combined manner. This technique, called four-dimensional or *4-quadrature* signalling [8], is a second approach to dual-polarization modulation.

1.2. Presentation and objectives

At the foundations of the present study lies the hypothesis that using polarization as a signalling quantity, as is done in polarization shift keying, should be feasible in satellite communication systems as well. Even though the technology and propagation environment are entirely different – guided light-waves in a fibre versus free-space microwaves – both settings permit the transmission and recep-

tion of different states of polarization.

The study investigates the use of polarization shift keying in the channel conditions of a satellite link. A first step shall transpose the modulation technique from the fibre optical to the microwave domain and analyse its applicability to the given satellite setting. A principal objective is to experimentally demonstrate the feasibility by a PolSK satellite transmission. Further investigations shall be conducted to determine to what extent the benefits of PolSK, identified above, can be preserved and whether other benefits become available.

A couple of requirements for a successful application on the satellite link can be set forth in advance. The satellite must support reception and retransmission of two orthogonal polarizations. This is generally fulfilled on many communication satellites with fixed positions on the geostationary arc. Channel definitions on the satellite should be aligned on both polarizations and the frequency conversion from the reception towards the transmission bands should be coherent such that the state of polarizations is preserved. Those prerequisites are available on today's modern communication satellites, operating usually in the C-, K_u- and K_a-bands[1].

Several publications report on the advantageous power efficiency of multilevel polarization shift keying with respect to comparable established modulation techniques [9, 5]. In the satellite setting, PolSK would supplant the previously used polarization division multiplexing. In that case, a fair comparison between the proposed and the currently established alternatives requires *dual* (i.e. on each polarization) operation of the latter. Taking this into consideration, the power efficiency of PolSK shall be re-evaluated in the present context. Should the power efficiency of PolSK reveal itself insufficient, the possibility exists to increase the spectral efficiency by embedding an additional phase modulation at practically no additional power cost.

Unlike polarization shift keying, four-dimensional signalling consists of a direct and conjoint modulation of the amplitudes and phases on two orthogonal polarizations. By fully using the available signal space, it could provide higher power efficiency than PolSK and is a valid alternative to simply adding a supplementary phase modulation. The implications of using such power-efficient techniques shall be determined as well. They shall be compared with polarization shift keying with the aim of defining preferred application scenarios.

For optimal detection performance, receivers require previous *synchronization* to an incoming signal before demodulation can start. Polarization shift keying does not require carrier synchronization, but symbol timing still needs to be

[1] Approximate uplink and downlink frequencies for the different bands:
C-band: ~ 4 GHz downlink and ~ 6 GHz uplink.
K_u-band: ~ 12 GHz downlink and ~ 14 GHz uplink.
K_a-band: ~ 20 GHz downlink and ~ 30 GHz uplink.

estimated. Four-dimensional signalling will in general require phase, frequency and symbol clock synchronization. Although some generic algorithms might be used, the literature generally remains silent on the synchronization of these signals. One objective of the project is therefore to define and implement adequate acquisition and tracking methods.

Finally, receiving modulated signals on both polarizations can provide other, additional benefits. A digital receiver demodulating both polarizations is able to remove crosstalk between both components by appropriate signal processing. In addition, reception can be accomplished independently of the antenna design as a horizontal/vertical linear or left/right circular reception unit. This potential should be kept in mind, but its exploitation is not planned for this project.

1.2.1. Methodology

To achieve the aforementioned goals, the proposed polarization shift keying technique shall in the first instance be analysed using a simulation program. This software shall be designed such as to emulate a dual-polarization channel in the satellite setting, including the expected disturbances from signal processing and noise. It will allow the questions of phase noise and the behaviour upon non-linear amplification to be addressed. A modular structure is targeted, such that the software will be able to analyse other single- and dual-polarization modulation techniques as well. This will allow the comparison with prevalent satellite modulation techniques, such as those proposed by the DVB-S2 standard[2] [28].

The simulations should provide initial insights into the proposed modulation techniques and prepare the design and implementation of a transmitter and receiver pair for the subsequent experiments and the planned demonstration. All practical questions of constellation design and synchronization shall be addressed at this stage. A satellite transponder simulator shall be used to validate the findings from the simulations and to explore the behaviour under realistic conditions.

In order to focus the analysis on the actual, physical modulation, the impact of error correction codes will not be considered. This should keep the designs simple and easily reproducible.

Once a sufficient degree of confidence is achieved with the transponder simulator, the experimental verification by a satellite demonstration shall be attempted.

[2] *Digital video broadcasting over satellite, second generation* is a popular standard for providing, amongst others, television broadcasting over satellite.

1.2.2. Evolution

During the course of the project it was found that augmenting a polarization shift keyed signal with an additional phase modulation, as considered in the initial project definition, was conceptually very close to the aforementioned four-dimensional signalling method. As the latter allows a more direct and organized design, and at the same time seems more promising in terms of achievable power efficiency, the decision was taken not to consider a PolSK augmentation but to focus the efforts on these four-dimensional signals alone.

For this second investigation, an approach similar to that for PolSK was selected, starting with simulations and proceeding subsequently with the implementation of a demonstrator. Due to the conceptual proximity, lessons and parts from the PolSK demonstrator have been reused to ease this task.

1.3. Structure of the dissertation

The second chapter introduces the general background to the work and refreshes some notions on communication systems. It presents the different modulation techniques under investigation and the current state of the research. Some of the propositions formulated superficially in this introduction are reformulated with the scientifically demanded rigour. The third chapter is dedicated to polarization shift keying and contains the first core element of the study. The second dual-polarization signalling technique, four-dimensional modulation, is presented and analysed in Chapter 4. The developments in Chapters 3 and 4 lead to a discussion on synchronization in dual-polarization signalling, which is covered in Chapter 5. The final chapter provides a review of the material which has been developed and draws the conclusions. Some thoughts and open trails will be summarized in the hope they can guide future research.

Four appendices complement the findings presented in the main document. They contain essential information and documentation that would have severed the thread of thoughts if included within the principal chapters. The inner workings of the simulation software developed for the analysis of the various modulation techniques are dissected in Appendix A. The implementation of the demonstration system represented a large and important part of the work; the details of the implementation are exposed in Appendix B. Appendix C describes some rarely used modulation techniques used for the comparison with the presented dual-polarization technique. Finally, Appendix D lists the mathematical notation and provides an explanatory table of all the acronyms which are employed.

2. State of the art

The work presented in the following chapters spans several disciplines, from optics to radio-frequency engineering, communication systems, digital signal processing and digital circuit design. Before plunging into the project's core subject-matter – the implementation and analysis of dual-polarization signalling techniques – the context is set by retracing some historical developments and by refreshing pertinent textbook knowledge. The definitions required for later use are provided and a choice of the employed models and tools is presented.

The phenomenon of optical polarization is a cornerstone of this work and is presented first.

It is followed by the definition of the satellite communication channel and the signal-to-noise ratio. The particularities of the satellite transponder and the main impairments in the channel are presented.

Two different modulation techniques have been analysed in the project, reflecting different views on the representation of polarization. These techniques, *polarization shift keying* and *4-quadrature signalling* were developed mainly during the 1980s and 1990s. They are introduced in the Sections 2.3 and 2.4, by briefly retracing their history. The relevant signal models and essential previous findings are restated.

Synchronization of dual-polarization modulation techniques has not been discussed in the literature so far. To prepare the associated developments in the later chapters, the synchronization problem is briefly introduced. A few selected, classical synchronization methods are shown, together with the tools for their analysis.

At the end, the basic structures of a typical modulator and a demodulator are revisited in view of the implementation of a demonstrator modem.

As a note, throughout the entire dissertation complex notation is used to describe all phenomena of narrowband bandpass nature, like field strength in waves or modulated bandpass *signals*. This is common practice in the field and eases representation and calculation.

2.1. Polarization

In microwave communications a propagating electromagnetic field is commonly modelled by a plane wave. In a plane wave, all field values are constant in any

plane perpendicular to the direction of propagation. As a consequence, the plane wave is also a transversal wave. In such a transversal electromagnetic wave (TEM wave) there are no field components in the direction of propagation [74, p. 645]. The TEM wave is an accurate and popular model for the propagation of short-wave beams or light in free-space. It is also a prerequisite for the study of optical polarization.

Polarization is a property of a monochrome (i.e. containing one single frequency) TEM wave and was discovered, initially as an optical phenomenon, in the 17$^{\text{th}}$ century by Christiaan Huygens [43, p. 3]. As a wave propagates, the endpoint of its electrical field vector moves in the plane perpendicular to the propagation direction. The nature of this movement – in general elliptical with linear and circular boundary cases – is called *polarization* [43, pp. 319ff.].

This movement takes place in a two-dimensional plane, such that the electrical field can be decomposed and described by two orthogonal components. In a Cartesian coordinate system the decomposition will be along x and y axes, for a wave propagating in the z-direction. Assuming such a monochrome TEM wave, the electrical field can be described as follows [74, pp. 646f.]:

$$\vec{E}(t,z) = \begin{pmatrix} E_{\mathrm{x}}(t,z) \\ E_{\mathrm{y}}(t,z) \end{pmatrix} = \begin{pmatrix} A_{\mathrm{x}} e^{\mathrm{j}(\omega t - k_z z + \phi_{\mathrm{x}})} \\ A_{\mathrm{y}} e^{\mathrm{j}(\omega t - k_z z + \phi_{\mathrm{y}})} \end{pmatrix} \tag{2.1}$$

In the time interval of one oscillation period $T = 2\pi/\omega$, or along one wavelength $\lambda = 2\pi/k_{\mathrm{z}}$ into the direction of propagation, $\vec{E}$ traces a figure that determines the polarization: The figure will be an ellipse, line or circle, depending on the amplitude ratio $A_{\mathrm{x}}/A_{\mathrm{y}}$ and the phase difference $\phi_{\mathrm{x}} - \phi_{\mathrm{y}}$ between both wave components.

In summary, the state of polarization (SoP) of a TEM wave can be determined by looking at the amplitudes and phases of two orthogonal field components. This is commonly done in optics, where the SoP is described using the Jones-vector.

2.1.1. Jones-vectors

The Jones-vector $\vec{J}$ [43, pp. 368ff.] is one way to describe the state of polarization of a plane wave. This two-dimensional vector contains the complex amplitude and phase terms of two linear orthogonal electrical field components perpendicular to the propagation direction (e.g. along the x and y axes as above). This allows the Jones-vector representation using either polar or Cartesian coordinates. Per analogy to the respective modulation quantities, the real part can be called the *in-phase* (I) term and the imaginary part the *quadrature* (Q) term (cf. Section

2.2.1). Both variants are illustrated in the equation below:

$$\vec{J} = \begin{pmatrix} A_x e^{j\phi_x} \\ A_y e^{j\phi_y} \end{pmatrix} = \begin{pmatrix} X_I + jX_Q \\ Y_I + jY_Q \end{pmatrix} \tag{2.2}$$

The Jones-vector is obtained by removing the time- and space-periodic terms in the exponential from equation (2.1). States of polarization are said to be orthogonal when their Jones-vectors are orthogonal[1]. This is the case for horizontal and vertical, or left- and right-handed circular polarization for instance. Jones-vectors are limited in that they can represent only fully polarized light. Transformations of the SoP can be described by Jones-matrices, which upon left-sided multiplication to an initial Jones-vector will yield the description of the newly created SoP [43].

2.1.2. Stokes-vectors

The phenomenon of polarization was known before the confirmation of the undulatory nature of light by Heinrich Hertz in 1888. Already in 1852, Georges Gabriel Stokes developed a description of polarization based on irradiance measurements along well-defined states of polarization: the so-called *Stokes*-parameters [43, pp. 366ff.]. This empirical approach, oblivious of the actual nature of light, allows a direct measurement of the SoP without requiring the concept of phases. There are four Stokes-parameters, enumerated below. They express a wave's degree of polarization and its tendency to resemble three predefined states of polarization:

S_0**:** Is proportional to the power flux density of the wave, including non-polarized radiation.

S_1**:** Resemblance to linear polarization along the principal reception axes x (horizontal) and y (vertical) ($S_1 > 0$: horizontal, $S_1 < 0$: vertical)

S_2**:** Resemblance to linear polarization diagonal to the principal reception axes x and y ($S_2 > 0$: linear +45°, $S_2 < 0$: linear −45°)

S_3**:** Resemblance to circular polarization ($S_3 > 0$: right circular, $S_3 < 0$: left circular)

The S_0-parameter is proportional to the total irradiance of the incoming wave and the following relation holds [43]:

$$S_1^2 + S_2^2 + S_3^2 \leq S_0^2 \tag{2.3}$$

Only fully polarized light is considered in the present work, for which equation (2.3) turns into an equality. In this case, S_0 is fully determined by the three other Stokes-parameters.

[1] i.e. when the vectors' inner product is zero.

The four Stokes-parameters, joined together, form the *Stokes-vector* $\vec{S}$. The Stokes-vector is applicable to any polarized transversal-electromagnetic wave – light or microwaves. Its components can be derived from the Jones-vector using the following expressions:

$$\vec{S} = \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} = \begin{pmatrix} A_x^2 + A_y^2 \\ A_x^2 - A_y^2 \\ 2A_xA_y\cos(\delta) \\ 2A_xA_y\sin(\delta) \end{pmatrix} = \begin{pmatrix} X_I^2 + X_Q^2 + Y_I^2 + Y_Q^2 \\ X_I^2 + X_Q^2 - Y_I^2 - Y_Q^2 \\ 2\left(X_IY_I + X_QY_Q\right) \\ 2\left(X_QY_I - X_IY_Q\right) \end{pmatrix} \quad \text{with } \delta = \phi_x - \phi_y \tag{2.4}$$

The unit-vectors along S_1, S_2, S_3 are orthogonal and span a three-dimensional *polarization*-space. Every physical state of polarization can be expressed as a linear combination of these three quantities[2]. In this Stokes-space, a constant-power wave will have all SoPs lying on a spherical surface of radius S_0. This sphere is called the *Poincaré-sphere*. Orthogonal states of polarization lie on diametrically opposed sides of the Poincaré-sphere. A similar calculus as for Jones-vectors exists to describe transformations of the SoP. For Stokes-vectors the transformation matrices are called *Müller-matrices* [43].

The information on a wave's state of polarization, be it in the form of Jones- or Stokes-vectors, can be used in communications. Modulating Stokes-vectors is called polarization shift keying, while modulating the Jones-vectors can be interpreted as 4-quadrature or four-dimensional modulation.

2.2. Satellite communications

Before passing on to the description of these modulation techniques, it is necessary to establish the setting in which they shall be employed. After a few remarks on the link scenario, basic knowledge on bandpass communication systems is restated to provide the essential definitions. The section ends with a description of a satellite's communication payload.

The project's focus is on satellite communication applications with highly directive, fixed antennas. These antennas have good polarization discrimination and allow the separate transmission and reception of two orthogonal states of polarization. Many antennas isolate the horizontal from the vertical polarization (along the x and y axes as defined in Section 2.1) or the right- from the left-handed circular polarization. This straightforward physical separation is used to transmit

[2] The definition of δ corresponds to the one in [6] and is opposite to that from [43]. This is due to the conflicting definitions of right- and left-handed polarization. This dissertation follows the *engineering* viewpoint, where right-handed circular polarization corresponds to the circular movement of the field along the fingers of the right hand, when the thumb is pointing *into* the direction of propagation [74, p. 649].

two independent carrier signals and allows the available frequency spectrum to be exploited twice. In satellite broadcasting and broadband data communication for example, this technique is common practice and called *polarization division multiplexing* (PDM).

To be able to process dual-polarization signals, the hardware involved needs to be able to do PDM. All analyses conducted in the framework of the project assume this capability. As a consequence, the presented findings are valid more specifically for broadcasting, broadband and direct-to-home applications over geostationary satellites. Without explicit limitation, the investigation focusses on the pertinent frequency bands stated in Section 1.2. Experiments are conducted in the K_u-band.

The reference scenario for the following work is a point-to-point, unidirectional link relayed over a geostationary satellite. The transmitter and receiver as well as the satellite have large high-gain antennas supporting good (i.e. 20-30 dB) polarization discrimination (Y as defined by the ITU[3] in [47]). Transmission and reception sites are able to operate on both polarizations simultaneously.

The essential impairments in this scenario are exposed below, after reviewing the bandpass channel model.

2.2.1. Bandpass signalling and signal-to-noise ratio

In typical digital bandpass signalling, the amplitude and phase of a carrier signal at a frequency ω_c are modulated. The information to be transmitted is first coded into a sequence of symbols, each representing a number, each having a specific amplitude and phase value identifying it unambiguously. The symbol with the position i in the sequence will be denoted c_i:

$$c_i = A_i e^{j\phi_i} = c_{I,i} + j c_{Q,i} \tag{2.5}$$

A polar and a Cartesian representation is possible. In the latter case the real part is the symbol's *in-phase* component, the imaginary part the *quadrature* component. The specific attribution of amplitude and phase to each symbol is governed by a *symbol constellation*. Many different constellations exist, popular examples are shown in [63, pp. 101-107] and include phase shift keying (PSK) and quadrature amplitude modulation (QAM).

With a given *symbol frequency* $1/T$, the carrier phase and amplitude are modified sequentially according to the values of these data-carrying information symbols. This is carried out by a device called a *modulator* (cf. Section 2.6). To ensure smooth transitions between symbols and thereby limit the bandwidth of the so-created bandpass signal, the symbol sequence is *pulse-shaped* prior to the

[3] International Telecommunication Union

actual modulation. A *root-raised cosine filter* (RRCS) filters the sampled symbol sequence to achieve this spectral shaping [63, pp. 607f.]. The same filter is also used in the receiver, where it is called the *matched filter*. The combined impulse response of both filters is denoted by $g(t)$ and corresponds to a *raised-cosine*.

To describe a modulated carrier signal $s(t)$, an expression similar to a TEM wave component is used:

$$s(t) = A(t)e^{\mathrm{j}(\omega_\mathrm{c} t+\phi(t))} \tag{2.6}$$

$$= e^{\mathrm{j}\omega_\mathrm{c} t} \sum_i c_\mathrm{i} \cdot g\,(t - iT) \tag{2.7}$$

Equation (2.6) describes a generic variation of amplitude and phase, whereas (2.7) accounts for the pulse-shaped discrete sequence of symbols c_i. For convenience, the full raised-cosine $g(t)$ is used in the expression of the transmitted signal[4].

In the literature, the unit of a signal $s(t)$ is often omitted. It may be a voltage at the antenna ports for instance. As digital signal processing works with numbers without direct physical meaning, the unit is lost after the conversion into the digital domain (after sampling and quantization). Still, it is preferable that a connection (e.g. proportionality) is maintained between physical values and their digital representation [48]. This is especially valid for a signal's *power*, which is usually defined as follows:

$$P_\mathrm{s} = \frac{1}{2T_\mathrm{m}} \int_{T_\mathrm{m}} s^*(t)s(t)\,dt \tag{2.8}$$

Equation (2.8) describes average power, measured over the arbitrary time interval T_m. The star * denotes complex conjugation. For P_s to represent a physical power, two conjugated physical values are required, like voltage and current. This is arranged for by assuming an implicit normative coefficient with the value of 1 Ω or 1 S, dependent on the hypothetical unit of $s(t)$. Inspired by a derivation proposed in [57], P_s is evaluated by combining (2.7) with (2.8):

$$P_\mathrm{s} = \frac{1}{2T_\mathrm{m}} \int_{T_\mathrm{m}} \sum_i c_\mathrm{i}^* \cdot g\,(t - iT) \sum_k c_\mathrm{k} \cdot g\,(t - kT)\;dt \tag{2.9}$$

The complex exponential dissolves into 1. The data symbols c_i are modelled as taken independently and with equal probability from a symbol constellation[5]:

$$\mathrm{E}\big[c_\mathrm{i}^* c_\mathrm{k}\big] = \begin{cases} C = 2E_\mathrm{s} & \text{if } i = k \\ 0 & \text{else} \end{cases} \tag{2.10}$$

[4] This leads to a negligible imprecision in the description of the noise statistics.

[5] For most real signals, source- and channel-coded, scrambled and interleaved, this is a perfectly valid assumption.

Assuming the measurement time T_m long enough, the statistical knowledge on c_i is used to collapse the product into a single sum. Upon distribution, only terms with $i = k$ remain:

$$P_\mathrm{s} = \frac{C}{2T_\mathrm{m}} \int_{T_\mathrm{m}} \sum_i g^2 (t - iT)\ dt \tag{2.11}$$

The expression can now be transformed into the frequency domain using the Poisson-Sum formula [40, p. 344], restated hereunder:

$$\sum_i f (t - iT) = \frac{1}{T} \sum_n F \left(\frac{2\pi n}{T} \right) e^{\mathrm{j} \frac{2\pi n}{T} t} \tag{2.12}$$

where $F(\omega)$ is the Fourier-transform of $f(t)$:

$$F(\omega) = \mathcal{F} \{f(t)\} = \int_{-\infty}^{\infty} f(t) e^{-\mathrm{j}\omega t}\ dt \tag{2.13}$$

Performing the transformation (2.12) on equation (2.11) and introducing $G_2(\omega)$ as the Fourier-transform of $g^2(t)$ yields:

$$P_\mathrm{s} = \frac{C}{2TT_\mathrm{m}} \int_{T_\mathrm{m}} \sum_n G_2 \left(\frac{2\pi n}{T} \right) e^{\mathrm{j} \frac{2\pi n}{T} t}\ dt \tag{2.14}$$

It is reasonable to assume the measurement duration T_m to be an integer multiple of the symbol clock period T [6]. In this way, only the term with $n = 0$ is non-zero and the calculation of P_s using (2.14) simplifies to:

$$P_\mathrm{s} = \frac{C}{2TT_\mathrm{m}} T_\mathrm{m} G_2(0) = \frac{E_\mathrm{s}}{2\pi T} \int_{-\infty}^{\infty} |G(\omega)|^2\ d\omega = \frac{E_\mathrm{s}}{T} \int_{-\infty}^{\infty} |g(t)|^2\ dt = \frac{E_\mathrm{s}}{T} \tag{2.15}$$

In the last equality it is convenient to assume that the *energy* of the pulse-shape $g(t)$ is normalized to 1.

Noise

The receiver gets a noisy copy $r(t)$ of the transmitted signal $s(t)$:

$$r(t) = s(t) + n(t) \tag{2.16}$$

[6] In other cases, the contribution added by the additional *fractional* symbol is dependent on the pulse-shape, which complicates the expression without adding real insight.

$n(t)$ is the noise captured by the receiver. In the satellite channel, it is usually a combination of thermal noise in the receiver's low-noise amplifier (LNA) and interference from neighbouring satellites or transmissions on adjacent frequencies. It is modelled as additive white Gaussian noise (AWGN) of power spectral density $N_0/2$ [63, p. 69]. As the signal is expressed in complex form, noise is defined as a complex-valued, circular white Gaussian noise process with:

$$\mathrm{E}\left[n^*(t)n(t+\Delta t)\right] = 2N_0\delta(\Delta t) \tag{2.17}$$

In practice, the noise power in the receiver is determined by the equivalent noise bandwidth B of the reception filter (i.e. the matched filter). This reception bandwidth is well approximated by the symbol clock frequency, such that:

$$B = \frac{1}{T} \tag{2.18}$$

The noise power P_n is:

$$P_\mathrm{n} = \frac{1}{2}B\,(2N_0) = \frac{N_0}{T} \tag{2.19}$$

Finally, the signal-to-noise ratio (SNR), an essential gauge for quantifying signal quality, can be deduced:

$$\mathrm{SNR} = \frac{P_\mathrm{s}}{P_\mathrm{n}} = \frac{E_\mathrm{s}}{T}\frac{T}{N_0} = \frac{E_\mathrm{s}}{N_0} \tag{2.20}$$

2.2.2. The satellite transponder

The task of the satellite in satellite communications is to receive a signal transmitted from Earth, to process it and to send a response back. The advantage of using satellites as relays is the exceptional coverage they can achieve and the possibility for broadcasting into vast areas. A signal transmitted to a satellite is called an *uplink* signal, the reverse direction is the *downlink*.

In the context of the present work, a basic platform, called a *repeater* or colloquially *bent pipe*, is considered; the satellite will transmit the same signal it receives back to Earth. This is accomplished by a satellite transponder. The processing steps in a transponder are schematically illustrated in Figure 2.1: First, an LNA amplifies the signal captured by the antenna. It is then frequency-converted into the downlink band. An input-multiplexing (IMUX) filter splits the signal according to predefined frequency channels and distributes them to different processing chains. In each chain the channels are amplified in two steps; the final high-power amplifier (HPA) is usually preceded by a pre-amplifier. An output-multiplexing (OMUX) filter joins the different chains again, before the combined signal is fed to the transmission antenna. Often, the same antenna is used for

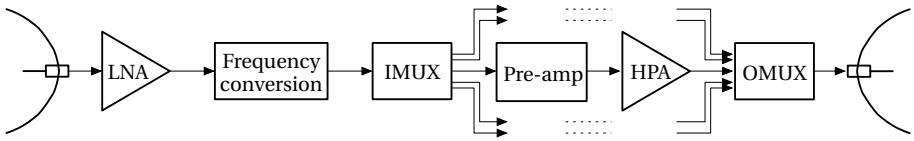


Figure 2.1.: Basic block diagram of a *bent pipe* satellite transponder

transmission and reception. In that case, a front-end filter before the LNA separates transmission and reception bands to prevent oscillations.

As the available electrical power from solar arrays is limited, the amplifiers (usually travelling wave tubes (TWTs), but semiconductor amplifiers also exist) have to operate as efficiently as possible, i.e. close to their maximal output power. The amplification characteristic deviates from an ideal linear curve as one approaches this so-called saturation point. This non-linearity is typically quantified by amplitude-to-amplitude modulation (AM-AM) and amplitude-to-phase modulation (AM-PM) curves, giving the amount of amplitude compression and phase rotation for each amplifier operating point (cf. Figure A.2 in Appendix A for an example).

The non-linearity distorts the signal and impedes successful reception. The amount of distortion introduced depends also on the signal's envelope variations. Therefore, the signals to be transmitted over such transponders should ideally be limited in their peak-to-average power ratio (PAPR). The PAPR is a simple and popular measure for envelope variations, the definition being obvious by its denomination:

$$\mathrm{PAPR} = \frac{P_{\mathrm{peak}}}{P_{\mathrm{average}}} \tag{2.21}$$

The higher the PAPR, the more severely a signal will suffer from distortion upon non-linear amplification.

A satellite in geostationary orbit usually has transponders for two polarizations, either left- and right-handed circular, or linear horizontal and vertical. It is common practice to signal on both polarizations; this is the *polarization division multiplex* mentioned above. For modularity or redundancy reasons the transponders on each polarization can be separate units, each having their own local oscillator. This leads to an incoherency in frequency conversion between the polarizations, which can pose a challenge for the considered dual-polarization modulation techniques.

Due to the analogue legacy in the broadcast channel definitions, transponder channels are not always aligned between polarizations. That is, the centre frequencies of the individual channels are positioned in echelon over both polarizations to reduce polarization cross-talk. In such a case, dual-polarization mod-

ulation can only be envisaged narrow-banded and in a scientifically motivated context. For commercial exploitation of such payloads, initiatives like the variable polarization/frequency multiplexing proposed in [97] seem more appropriate.

2.2.3. Propagation and polarization crosstalk

In free-space and clear sky conditions, all states of polarization propagate equally well. During rain, the horizontal polarization is more strongly attenuated than the vertical. Also, rain drops will lead to crosstalk and a rotation of the polarization plane. Antennas on both sides of the link are usually large and discriminate polarizations very well. However, this discrimination is finite and a certain level of crosstalk is unavoidable.

As long as the antennas are properly aligned, the level of crosstalk lies typically an order of magnitude below the noise and can be ignored in a first approximation. This is the approach chosen here.

2.3. Polarization shift keying

Polarization shift keying is the first modulation type proposed for use in satellite communications. It uses the state of polarization of a carrier wave as the information-bearing quantity. To this end, the Stokes-vector is used as a modulation parameter.

2.3.1. Presentation of previous work

Modulation using the state of polarization was explored in the early 1960s for binary signalling over microwaves [34] and continuous and multi-state transmission over optical fibre [59]. This is possible in media allowing two orthogonal polarizations to propagate, such as single-mode fibres or free-space. Roughly twenty years later, these ideas are used to bypass problems with phase noise in optical channels by modulating the phase *difference* between orthogonally polarized carrier signals [20]. An early demonstration of the technology is described in [23] and the name *Polarization Shift Keying* is introduced. A second experimental demonstration was performed by a research group in Italy [15]. [10] first describes the use of Stokes-vectors to describe the modulation and performs detection in Stokes-space using antipodal states of polarization.

The main motivation behind PolSK in fibre optics is that the broad laser line-width, i.e. the width of the phase noise spectrum of the emitting laser diodes, makes coherent reception more or less impossible (a line-width of 20 MHz is put forward as an example in [23]). Polarization shift keying encodes the phase

difference between orthogonally polarized components of the same laser source. The carrier variations are equal on both components and can be neutralized by mixing. In this way, phase modulation becomes possible in an otherwise *non-coherent* channel.

This property allows multi-level (i.e. more than 1 bit per symbol) modulation with constant transmission power. [9] addresses the state of polarization fluctuations occurring in an optical fibre and demonstrates how a Stokes-receiver can overcome them using electronic polarization tracking. Also, first multi-level constellations are presented and their error-sensitivity is estimated. A sensitivity improvement is shown for PolSK with respect to classical phase shift keying and differential phase shift keying (DPSK) signalling schemes.

A Stokes-receiver with a direct-detection front-end and an electronic back-end is proposed in [7]. The system performance is analytically derived for channels dominated by Gaussian noise and for *Amplified Spontaneous Emission* [3] noise.

The combined findings of the research community on polarization shift keying are well summarized in the papers [6] and [5]. Furthermore, [6] gives exact sensitivity calculations for regular constellations, along with results for selected four- and eight-point constellations. A spectral analysis and a proposal for differential PolSK, without the need for polarization tracking follows. Finally, [5] proposes a decision rule for the optimal maximum-likelihood PolSK receiver and computes the exact symbol-error rate of constant-power constellations. A solution to the problem of channel dichroism is proposed by expanding the receiver to all four Stokes-parameters. These last two papers present PolSK in a way very close to radio-frequency engineering, including the possibility of matched filtering and a discussion on white noise. In this sense, they provide very solid foundations for the work conducted in the framework of the present thesis.

From the information in [9, 6, 5], it can be concluded that constant-power PolSK schemes outperform comparable PSK and DPSK schemes in power efficiency for higher-order constellations. Thus for more than eight symbols in the constellation, PolSK has better sensitivity than PSK [9]. Concerning DPSK, PolSK presents a superior sensitivity already from four symbols on [6]. This advantage increases with the number of constellation points. This observation is explained by the fact that on the two-dimensional surface, the mutual distance decreases more slowly with an increasing number of points than on a planar circle [9].

Recently, polarization shift keying has regained some popularity as shown by an increasing number of publications. [42] presents the demonstration of quaternary, differential PolSK over fibre channel. Results on another demonstration are published in [78], this time in a free-space optical channel. Published theses on PolSK include [24] and [77].

2.3.2. PolSK signal model

For PolSK, symbol constellations are designed in the three-dimensional Stokes-space, in the same way as is done in the I/Q plane for classic digital modulation. The literature proposes a certain number of constant-power constellations, $N = 4, 8$ up to $N = 64$ which are laid out on the two-dimensional Poincaré-sphere given by $S_0 = \text{const}$.

Different transmitter models have been proposed, which all start from a light source linearly polarized in x-direction:

$$\vec{E}_{\text{x,in}} = Ae^{\text{j}\omega_c t}\vec{e}_{\text{x}} \tag{2.22}$$

The modulator described in [5] splits the power of the light-wave to both polarizations according to an angle θ. In addition, an angle ϕ determines the phase difference between both polarizations, by advancing the x and retarding the y component for instance. These operations can be expressed in Jones-calculus by two matrices $\mathbf{T}_\theta$ and $\mathbf{T}_\phi$.

$$\vec{E}_{\text{out}} = \mathbf{T}_\theta \mathbf{T}_\phi \vec{E}_{\text{x,in}} \tag{2.23}$$

$$\begin{pmatrix} E_{\text{x,out}} \\ E_{\text{y,out}} \end{pmatrix} = \begin{pmatrix} e^{\text{j}\frac{\phi}{2}} & 0 \\ 0 & e^{-\text{j}\frac{\phi}{2}} \end{pmatrix} \begin{pmatrix} \cos(\frac{\theta}{2}) & -\sin(\frac{\theta}{2}) \\ \sin(\frac{\theta}{2}) & \cos(\frac{\theta}{2}) \end{pmatrix} \vec{E}_{\text{x,in}} \tag{2.24}$$

The result is a polarization modulated wave with the expression:

$$\vec{E}_{\text{out}} = A \begin{pmatrix} e^{\text{j}\frac{\phi}{2}} \cos(\frac{\theta}{2}) \\ e^{-\text{j}\frac{\phi}{2}} \sin(\frac{\theta}{2}) \end{pmatrix} e^{\text{j}\omega_c t} \tag{2.25}$$

The state of polarization is thus realized by the superimposition of two linearly polarized components. By modulating the angles θ and ϕ, the state of polarization can be continuously changed. Another transmitter described in [9] is slightly more involved; it seems to account for more realistic implementation limitations.

Figure 2.2 shows the receiver front-end from [6]. The incoming light-wave is split along two orthogonal linear polarizations. Both polarizations are mixed with a common local oscillator signal and bandpass filtered. Subsequent squaring, phase-shifting and mixing extracts the Stokes-parameters which are passed on to the electronic part of the receiver where polarization de-rotation and the final detection occur.

The operations extracting the Stokes-parameters will also transform the statistics of the received noise; the Gaussian property is lost. The probability density functions of a noisy PolSK signal are derived in [6]. Fortunately, [6] shows as well that a demodulator for *constant-envelope* PolSK can still use the minimum distance criterion for optimal demapping.

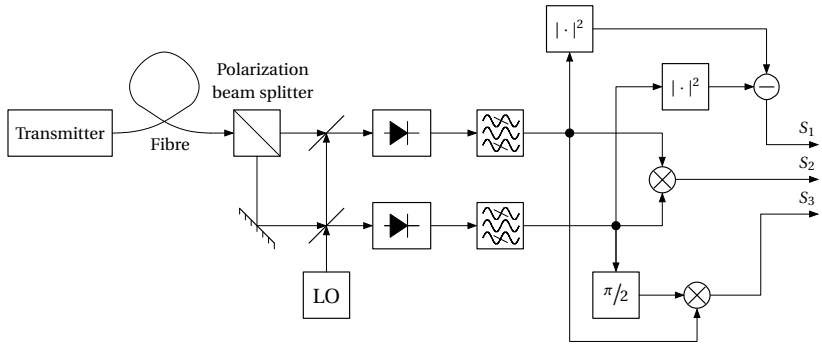


Figure 2.2.: Basic PolSK receiver front-end. Drawing from [6].

2.4. Four-dimensional signalling

Four-dimensional (4-D) signalling is structurally closer to standard, linear modulation. It uses all four orthogonal carriers available in a dual-polarization scenario (i.e. in-phase and quadrature carriers on each polarization) for signalling. In this way, the coherent channels on both polarizations are combined into one.

2.4.1. Presentation of previous work

An early publication on four-dimensional modulation is [86]. It applies the general considerations on codes from [75] to waveform transmission in four dimensions under peak and average power constraints. Hints for the concrete implementation of the theoretical results are given as using two bandpass channels, possibly as orthogonal waveguide modes. A series of *codes*[7] with spectral efficiencies from 8 up to 256 points are presented.

Another significant publication on four-dimensional signalling is [8]. Written at the time when coherent optical modulation started to become popular, it departs from the premise that in the coherent fibre-channel four orthogonal signalling dimensions are available. It is a very complete publication, presenting a transmitter, modelling the signal propagation and describing several reception methods using a baseband processor to equalize polarization fluctuation and to generate control signals for synchronization. The technique is called *4-quadrature* signalling.

A little later, a first article on four-dimensional modulation specific to satellite communications was published, [79]. The four dimensions are defined using

[7] read: constellations

different basis functions in the same bandpass channel, following a concept introduced in [66]. It uses different methods from sphere-packing, lattices and coded modulation to modify power and spectral efficiency of four-dimensional constellations and presents simulation results for the satellite channel.

With the recent advances in high-speed digital circuits and fast analogue-to-digital converters, four-dimensional coherent optical signalling has been revived [2, 101]. A couple of publications discuss the use of four-dimensional modulation in a generalized multi-dimensional signalling approach. This allows the available bandwidth to be managed more flexibly. Other advantages include the increased minimum distance between symbol points with respect to two-dimensional constellations [25], the ability to detect cycle slips [100] or the ability to equalize nonlinear phenomena in the fibres [25]. [11] contains proposals for coding schemes.

In [51] a comparison between different signalling constellations in two, three and four dimensions is shown, concluding that the bi-orthogonal constellation in four-dimensions is the most power efficient among the candidates. This modulation is called polarization-switched quaternary PSK or PS-QPSK in the optical domain. The findings from [51] are generalized to a large selection of 4-D constellations in [2]. Another publication from the same authors, [50], focusses on large constellations, targeting spectral rather than power efficiency.

2.4.2. Multi-dimensional signalling

The increase from two to four dimensions re-opens the question of good symbol constellations for such a channel. Useful hints on how to approach constellation design come from the coding literature, where a high number of dimensions is not uncommon. The concepts of *coding gain* and *shaping gain* can orient the design and permit improvements with respect to classic I/Q modulations to be quantified. Good overviews are provided in [32] or [31] and, especially for the topic of shaping, [54]. The pertinent ideas can be summarized as follows:

An important figure for the noise sensitivity of a constellation is the minimum distance between signal points. It is given by the constellation packing and largely determines the symbol-error probability when the signal-to-noise ratio is high. In a standard 2^{2n}-point quadrature amplitude modulation (2^{2n}-QAM) constellation for instance, the packing is a rectangular grid. In four dimensions, a more dense packing is possible, allowing more constellation points at equal mutual distance for the same average symbol energy. The densest packing in four dimensions is given by the so-called chequerboard, or D_4, lattice [19]. The increase in power efficiency that can be achieved by choosing a more dense packing is called *coding gain*.

Modulation schemes like 2^{2n}-QAM have a rectangular shape, while the average symbol energy boundary, given by a radius $\sqrt{E_s}$ in the I/Q-plane, is *circu-*

lar. For this reason cross-shaped constellations are proposed in [32] to fill the signalling space more efficiently. This principle can be extended to a higher number of dimensions, where it is then called *constellation shaping*. In four dimensions the average symbol energy is bounded by a three-dimensional spherical surface, according to which the constellation space can be filled. For equiprobable symbols such constellations, when projected into the respective I/Q-planes, approach the Gaussian bell curve, which is optimal for the linear AWGN-channel [73]. Shaping gain in four dimensions can increase by up to 0,46 dB the relationship between minimum distance and average power [32].

In short, the increase in dimensions allows constellation points to be packed more densely, which in practice means that more information can be transmitted using the same power at the same symbol-error rate (SER). In addition, the signal space defined by a power constraint can be filled up more efficiently. These principles from coding theory can be used in physical modulation as well to increase the power efficiency in uncoded modulation.

The achievable gains and the techniques' practical applicability are limited. In an increasingly dense lattice, each node has a higher number of neighbours, leading to a *redundancy loss*. This can be quantified using the lattice's number of nearest neighbours (also *Kissing-number* [31]). In addition, the presence of many neighbours hampers a *good* attribution of bit-values to each node. Finally, the application of coding and shaping for constellation design leads to the expansion of the underlying, lower-dimensional constellations [32]. As a consequence, the peak-to-average power ratio of such constellations increases.

2.4.3. Four-dimensional signal model

In order to describe 4-D modulation, the common channel model for bandpass signalling from equation (2.7) is expanded to two conjointly modulated carrier signals. This leads to a vector-like signal description:

$$\mathbf{s}(t) = \begin{pmatrix} A_{\mathrm{x}}(t)e^{\mathrm{j}(\omega_c t+\phi_{\mathrm{x}}(t))} \\ A_{\mathrm{y}}(t)e^{\mathrm{j}(\omega_c t+\phi_{\mathrm{y}}(t))} \end{pmatrix} \tag{2.26}$$

In coherent modulation on a single polarization, a particular symbol, numbered i, is defined by a complex scalar c_{i} containing the carrier's amplitude and phase information. As in 4-D modulation carriers on both polarizations are used, a pair of complex numbers ($c_{\mathrm{x,i}}$, $c_{\mathrm{y,i}}$) is required to completely define a symbol:

$$\mathbf{s}(t) = e^{\mathrm{j}\omega_c t} \sum_i \begin{pmatrix} c_{\mathrm{x,i}} \\ c_{\mathrm{y,i}} \end{pmatrix} \cdot g\,(t - iT) \tag{2.27}$$

A 4-D symbol has thus four orthogonal parameters: two in-phase, and two quadrature components, one on each polarization. The quantities $c_{\mathrm{x,i}}$ and $c_{\mathrm{y,i}}$, named

symbol-elements, are the projections of the symbol vector $\mathbf{c}_i$ onto a single polarization. The *apparent* 2-D symbol constellations formed by these projections are called *constituent 2-D constellations* [32].

2.4.4. Bi-orthogonal signalling

The bi-orthogonal signal, identified in [2] as the most power efficient modulation technique in four-dimensions, is particularly interesting. In bi-orthogonal modulation, constellation points are positioned on opposite sides of each axis defining the orthogonal dimensions [63, pp. 111f.].

In a one-dimensional transmission channel, bi-orthogonal modulation has the constellation points $\{-1;1\}$, corresponding to binary PSK (BPSK). In two dimensions, bi-orthogonal signalling is equivalent to QPSK while in four dimensions, eight constellation points are available, listed in the table below:

$$\left\{\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}-1\\0\end{pmatrix},\begin{pmatrix}\mathrm{j}\\0\end{pmatrix},\begin{pmatrix}-\mathrm{j}\\0\end{pmatrix},\begin{pmatrix}0\\1\end{pmatrix},\begin{pmatrix}0\\-1\end{pmatrix},\begin{pmatrix}0\\\mathrm{j}\end{pmatrix},\begin{pmatrix}0\\-\mathrm{j}\end{pmatrix}\right\} \tag{2.28}$$

This constellation conveys 3 bits per symbol. It contains symbol-elements with a zero-value and is consequently not particularly suited for transmission over satellite transponders. This can be remedied by a rotation in 4-D space; the consequence is the constellation called PS-QPSK:

$$\begin{aligned}\frac{1}{2}\cdot\Bigg\{&\begin{pmatrix}-1+\mathrm{j}\\-1-\mathrm{j}\end{pmatrix},\begin{pmatrix}1-\mathrm{j}\\1+\mathrm{j}\end{pmatrix},\begin{pmatrix}-1-\mathrm{j}\\1-\mathrm{j}\end{pmatrix},\begin{pmatrix}1+\mathrm{j}\\-1+\mathrm{j}\end{pmatrix},\\&\begin{pmatrix}1-\mathrm{j}\\-1-\mathrm{j}\end{pmatrix},\begin{pmatrix}-1+\mathrm{j}\\1+\mathrm{j}\end{pmatrix},\begin{pmatrix}1+\mathrm{j}\\1-\mathrm{j}\end{pmatrix},\begin{pmatrix}-1-\mathrm{j}\\-1+\mathrm{j}\end{pmatrix}\Bigg\}\end{aligned} \tag{2.29}$$

The rotation does not modify the mutual distances between the constellation points. Simple geometry shows that the distances are larger by $\sqrt{2}$ with respect to QPSK at the same power expense. The rotated constellation has constant-power carriers on both polarizations and is thus more suitable for amplifiers operating close to saturation.

2.5. Receiver synchronization

A prerequisite for successful reception and demapping in a receiver is the proper synchronization to the incoming signal. Literature seems to provide no guidance on how to synchronize the aforementioned modulation techniques. One objective of the present work is to identify and implement adequate synchronization for PolSK and 4-D signalling over satellite. After a brief overview of the synchronization problem, a selection of popular detectors are presented and the *Modified*

Cramer-Rao Bound (MCRB) is introduced, a tool to estimate a detector's performance.

Synchronization is the alignment of internal timing references in a receiver to timing properties of the received signal. More specifically, a receiver is required to estimate the incoming signal's symbol clock, carrier phases and frequencies as a prerequisite for successful reception[8] [57]. This is usually accomplished using feedback control systems called phase-locked loops (PLLs) [39].

In a satellite receiver, a signal's time of arrival is not known beforehand: It arrives with an unknown time delay τ and unknown phase ϕ_0. Also the exact carrier frequency is not known precisely, which is expressed by including an offset $\Delta\omega$. In order to take these uncertainties into account, equation (2.7) is expanded [57, p. 17]:

$$s(t) = e^{\mathrm{j}[(\omega_\mathrm{c}+\Delta\omega)t+\phi_0]} \sum_i A_\mathrm{i} e^{\mathrm{j}\phi_\mathrm{i}} \cdot g\left(t - iT - \tau\right) \tag{2.30}$$

Synchronization consists thus in the estimation and neutralization of these uncertainties and needs to be accomplished prior to demodulation. This is usually achieved by using a detector that provides an estimate of the error between the current value of the receiver's internal clock and the real value in the signal. This estimate is fed to a PLL, which is capable of adjusting the internal clock to the target value until the error vanishes.

In the following, a couple of popular detectors are proposed and the functionality of a phase-locked loop is described. The section closes with the introduction of the MCRB and its derivation for timing estimates.

2.5.1. Detectors

Different detectors exist to recover the phase, timing or frequency information in a signal. Modern recovery systems can estimate several of these quantities simultaneously, although a separate estimation is considered more robust in general [57, p. 40]. Detectors can be classified in two broad categories. Non-data-aided and data-aided. A non-data-aided recovery method tries to recover the desired information from the waveform directly, whereas data-aided methods rely either on pilot symbols (i.e. known sequences in the incoming signal that are transmitted solely for the purpose of synchronization) or on previous receiver decisions (the decisions are used to *wipe off* the modulation from the carrier). Data-aided systems are more generic but at the same time more delicate. Therefore, only non-data-aided methods are considered here.

[8] For higher-layer communication protocols, the start and end of frames need to be estimated as well.

Square timing-error detector

As a consequence of pulse-forming, a digitally modulated signal has slight power fluctuations, synchronous to the symbol clock. Upon squaring the signal, these power fluctuations express themselves by a spectral component at the symbol clock frequency. This is used by the simple but powerful timing-error detector proposed in [36]; it takes the downconverted baseband components of the received signal $r(t)$ and extracts a timing wave $z(t)$ by squaring:

$$z(t) = r_{\mathrm{bb}}^2(t) \tag{2.31}$$

$r_{\mathrm{bb}}(t)$ is an in-phase or quadrature baseband component of $r(t)$. The squared signal $z(t)$ can be fed to a narrow-band filter at the expected clock frequency to extract the timing wave. A more advantageous method is to get a timing error $\Delta\varepsilon(t)$ by multiplying $z(t)$ with a reference clock from an oscillator and use it to control a PLL. It is preferable to sum the timing waves from both in-phase and quadrature components to make the detector work without requiring phase-lock.

The detector's precision depends largely on the pulse-shape and on the data sequence. It is very convenient for analogue timing recovery.

Gardner timing-error detector

A second timing error detector is presented in [37], suitable for digital systems operating on a sampling frequency of twice the symbol frequency. As it is derived from the squaring detector introduced above, it operates on the baseband signal as well. It directly computes the timing-error $\Delta\varepsilon[n]$ according to the following formula:

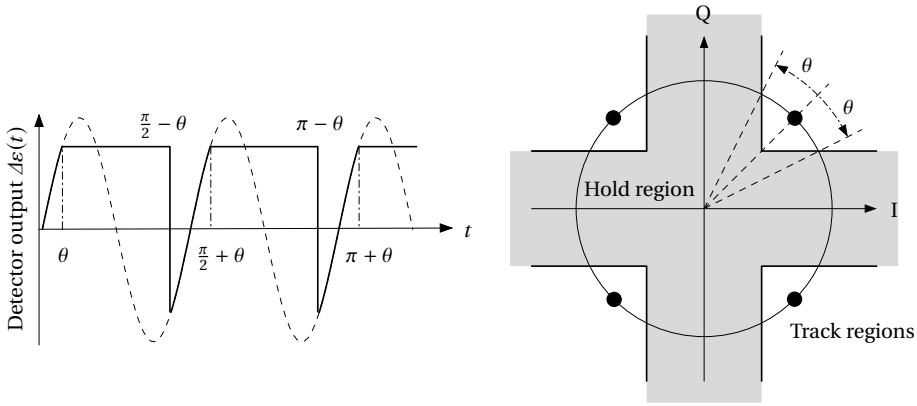
$$\Delta\varepsilon[n] = r_{\mathrm{bb}}[n - 1/2]\,(r_{\mathrm{bb}}[n] - r_{\mathrm{bb}}[n-1]) \tag{2.32}$$

The Gardner timing-error detector (GED) operates in discrete-time, on samples at twice the symbol clock: $r_{\mathrm{bb}}[n]$ is the n^{th} symbol of a baseband component (I or Q). The brackets, as opposed to round braces, emphasize the discrete-time nature of the expressions. The GED is based on the observation that the timing error can be measured by taking the difference between the squared samples a quarter symbol before and after the symbol strobe. Taking the signal's discrete-time derivative ensures that the actual symbol strobe falls onto one of the samples.

In the same way as above, this detector still works with a phase or small frequency error when errors from I and Q paths are added up [37]. Its precision decreases with the signal's excess bandwidth (i.e. roll-off).

Phase and frequency detector for PSK and QAM

A detector for phase and frequency, combining solid jitter performance with a large acquisition bandwidth, is described in [68]. It is based on a classic phase de-



a) Detector output in track and hold modes in case of a frequency error.

b) Tracking and holding using threshold values for a QPSK signal.

Figure 2.3.: Left: Detector output in the presence of a positive frequency offset. Right: PFD construction for carrier recovery in QPSK. Figures drawn after [68]

tector, for instance by multiplication of the signal with a reference oscillator output, yielding a sine-shaped detection characteristic (*track mode*). As soon as the so-detected phase error trespasses a predefined threshold θ on the rising carrier flank, the detector outputs a constant value (*hold mode*) until the condition is relieved in the following cycle. In this way, a frequency error is caught up quickly and the device acts as a *phase and frequency detector* (PFD). An example detector output is shown on the left side of Figure 2.3.

For simple modulation formats, like BPSK and QPSK, a fast modulation wipe-off can be performed by accepting phase errors in any of the modulation quadrants. This is illustrated on the right side of Figure 2.3. This gets more involved for higher-level QAM constellations where this gradually approaches a decision-directed detector.

2.5.2. Phase-locked loops (PLLs)

The detectors described in the previous section are primarily used in PLLs to track essential oscillatory clock signals in the received signal. A phase-locked loop consists of three elements: a phase detector, a loop filter and a voltage-controlled oscillator (VCO) [39, p. 1]. The VCO tries to *follow* the input signal, while the phase detector measures the time or phase difference between the VCO output and the input signal. The filter smooths the error signal and feeds it back as a control sig-

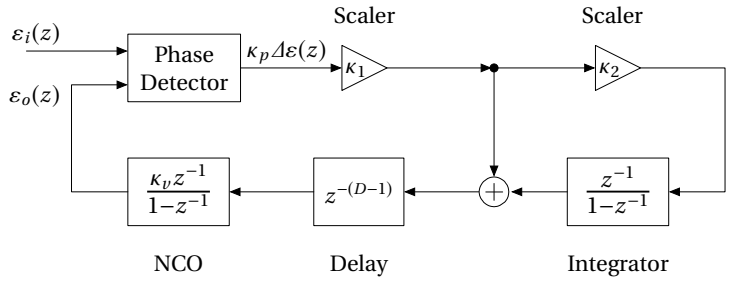


Figure 2.4.: Block diagram of the very common digital type 2 PLL, drawn after [39]

nal to the VCO. The filter is responsible for the dynamic behaviour of the loop.

Digital PLLs operate in a similar way to their analogue counterparts. They work on sampled sequences of numbers rather than real physical quantities. The output can be simple binary clock signal, or the output of a numerically controlled oscillator (NCO).

Figure 2.4 shows a simplified block diagram of a digital type 2 PLL[9] from [39, p. 69]. The phase-detector compares the incoming phase ε_i with the phase output ε_o of an NCO and delivers an error signal. Although the phase detector output is usually not proportional to the phase error $\Delta\varepsilon$, it is often so for small errors and can be modelled as such. κ_p is the phase detector gain. The loop filter of a type 2 digital PLL consists of a *proportional* path and an *integral* path. The two paths are added and fed to the NCO. The delay block models unavoidable computational delay in the loop elements. The NCO accumulates the filtered samples, adjusted by the *NCO scaling coefficient* κ_v, and creates an oscillation based on the resulting phase. The feedback loop is closed by feeding the NCO output back to the phase detector for comparison with the incoming oscillation [39].

Essential design parameters of a type 2 PLL are its loop bandwidth and damping coefficient. Thermal noise in the signal will lead to fluctuating errors in the synchronization estimates, commonly called jitter. Jitter may be reduced by decreasing the loop bandwidth, i.e. performing the estimate over a longer measurement period. This however in turn reduces the synchronizer's capability to follow quick variations in the quantities to be estimated like timing variations or phase noise. Therefore, the designer of a synchronization system needs to adjust the loop bandwidth to find a good compromise between the robustness against thermal noise and the tracking speed.

The term *loop bandwidth* is ambiguous; when AWG noise is considered the

[9]The type indicates the number of integrators in the loop. It is related to its capacity to track variations of different order. A type 1 PLL tracks phase without error only in the absence of a frequency error. A type 2 PLL tracks frequency without error in the absence of frequency drift.

predominant disturbance, Gardner recommends to use the noise bandwidth [39, p. 17f.]. The expression for the noise bandwidth B_N for the PLL shown in Figure 2.4 is [39, p. 130]:

$$B_N = \frac{\kappa f_s}{4} \frac{1 + \frac{\kappa_2}{\kappa} - \frac{\kappa_2}{2}(3 - \kappa_2)}{1 - \kappa_2 - \frac{\kappa}{4}\left(2 - \kappa_2 + \kappa_2^2\right)} \tag{2.33}$$

$\kappa = \kappa_p \kappa_v \kappa_1$ and f_s is the sampling frequency of the system.

The damping coefficient ν determines the dynamic behaviour of the loop by controlling the impact of the integral path on the total loop behaviour. Low damping allows faster roll-off in the loop frequency response but a strong gain peaking effect, which is undesired. The damping coefficient is often chosen to be around $\nu \simeq 1/\sqrt{2}$. It can be estimated using the following formula [39, p. 85]:

$$\nu = \frac{1}{2}\sqrt{\frac{\kappa}{\kappa_2}} \tag{2.34}$$

2.5.3. Detector performance estimation

In many cases it is appropriate to study PLL and detection behaviour by simulations. Analytical bounds are very useful to compare the achieved performance with best- or worst-case expectations. The *modified Cramer-Rao bound* [22] is such a lower bound for the variance in the unbiased estimation of frequency, phase and timing error in additive white Gaussian noise. It is derived from the Cramer-Rao bound found in estimation theory, modified to ease calculation for the present problem. Its derivation is given in [22] and [57, pp. 54f.,64f.] and is briefly reiterated here.

Among the random variables $\boldsymbol{\gamma} = \{\Delta\omega, \phi_0, \tau, \mathbf{c}^T\}$ in a signal $s(t)$, a single parameter λ shall be estimated. All other parameters in $\boldsymbol{\gamma}$ are regrouped in the random vector $\mathbf{u}$ [10]. A realization of λ should now be estimated from the observed noisy signal $r(t) = s(t, \lambda, \mathbf{u}) + n(t)$ [11].

According to the *maximum-likelihood* principle, given a series of observations, the estimated parameter $\hat{\lambda}$ should be selected such as to render these observations the most probable [81]. Formally, this is expressed by maximizing a *likelihood* function $\Lambda(r(t)) = p_\lambda(r(t))$, where $p_\lambda(x)$ is the probability of observing x for a given λ. Then

$$\hat{\lambda} = \arg\max_{\lambda} \Lambda(r(t)) \tag{2.35}$$

[10] $\mathbf{c}$ contains the data symbols c_i which are not directly part of the synchronization problem. Nonetheless, for the receiver it is a random variable and belongs in $\mathbf{u}$.

[11] For brevity and simplicity, there is no notational difference made between the statistical model and its realization.

The noise $n(t)$ is AWGN, the likelihood function can therefore be expressed using the Gaussian probability density function. After taking the logarithm and removing some constant terms which are irrelevant for the optimization, this *log-likelihood* function takes the following form [22]:

$$\ln \Lambda(\lambda, \mathbf{u}) = \frac{-1}{2N_0} \int_{T_0} \left| r(t) - s(t, \lambda, \mathbf{u}) \right|^2 dt \tag{2.36}$$

T_0 is the duration of the observation (i.e. measurement). It corresponds more or less to half the inverse of a tracking PLL's loop bandwidth [57, pp. 130,216,383]. The MCRB is defined as follows [22]:

$$\mathrm{MCRB}(\lambda) = \frac{1}{\mathrm{E}_{n,\mathbf{u}}\left[\left(\frac{\partial \ln \Lambda(\lambda, \mathbf{u})}{\partial \lambda} \right)^2 \right]} \tag{2.37}$$

$$= \frac{N_0}{\mathrm{E}_{\mathbf{u}}\left[\int_{T_0} \left| \frac{\partial s(t, \lambda, \mathbf{u})}{\partial \lambda} \right|^2 dt \right]} \tag{2.38}$$

Note that in (2.37) the expected value is over the noise n and the vector $\mathbf{u}$, while in (2.38) only $\mathbf{u}$ remains. The statistics of $\mathbf{u}$ are assumed to be known. As an illustration, the specific MCRB for $\lambda = \tau$ is computed in the following, departing from the expression of $s(t)$ (cf. equation (2.30)):

$$s(t) = e^{\mathrm{j}[(\omega_\mathrm{c} + \Delta\omega)t + \phi_0]} \sum_i c_\mathrm{i} \cdot g\left(t - iT - \tau\right) \tag{2.39}$$

The denominator of equation (2.38) is:

$$\mathrm{E}_{\mathbf{u}}\left[\int_{T_0} \left| \frac{\partial\, s\left(t, \tau, \mathbf{u} = (\Delta\omega, \phi_0, \mathbf{c}^\mathrm{T})\right)}{\partial \tau} \right|^2 dt \right] \tag{2.40}$$

$$= \mathrm{E}_{\mathbf{u}}\left[\int_{T_0} \left| e^{\mathrm{j}[(\omega_\mathrm{c} + \Delta\omega)t + \phi_0]} \sum_i c_\mathrm{i} \cdot \frac{\partial}{\partial \tau} g\left(t - iT - \tau\right) \right|^2 dt \right] \tag{2.41}$$

Due to the absolute value, the complex exponential can be removed after taking the derivative and $\Delta\omega$ as well as ϕ_0 disappear from the expression:

$$\mathrm{E}_{c_\mathrm{i}}\left[\int_{T_0} \left| -\sum_i c_\mathrm{i} g'(t - iT - \tau) \right|^2 dt \right] \tag{2.42}$$

The expected value can be evaluated before the integral:

$$\int_{T_0} \mathrm{E}_{c_\mathrm{i}} \left[\left(\sum_i c_\mathrm{i}^* g'(t - iT - \tau) \right) \left(\sum_k c_\mathrm{k} g'(t - kT - \tau) \right) \right] dt \tag{2.43}$$

After distributing the product, the property of independent symbols is used to discard all mixed terms (cf. equation (2.10)). Remembering that

$$\mathrm{E}\left[c_\mathrm{i}^* c_\mathrm{k}\right] = C \text{ for } i = k \tag{2.44}$$

this leads to

$$\int_{T_0} C \sum_i \left(g'(t - iT - \tau)\right)^2 dt = \int_{T_0} C \sum_i p(t - iT - \tau)\, dt \tag{2.45}$$

with the helper function $p(t)$, defined as follows:

$$p(t) = \left(g'(t)\right)^2 \tag{2.46}$$

Again, the Poisson-Sum formula from [40, p. 344] transforms the series into the frequency domain. Analogous to the description in Section 2.2.1, the measurement duration is chosen as an integer multiple of T, such as to discard all terms with $n \neq 0$.

$$\frac{C}{T} \int_{T_0} \sum_n P\left(\frac{2\pi n}{T}\right) e^{\mathrm{j}\frac{2\pi n}{T}(t-\tau)}\, dt = C\frac{T_0}{T} P(0) \tag{2.47}$$

Using the definition of $p(t)$ (equation (2.46)) and elementary Fourier-transform calculus [64, p. 313], $P(\omega)$ is further dissected:

$$P(\omega) = \frac{-1}{2\pi} \int_{-\infty}^{\infty} (\omega - \xi)\, G\left(\omega - \xi\right) \xi G\left(\xi\right)\, d\xi \tag{2.48}$$

and $P(0)$:

$$P(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \xi^2 G\left(-\xi\right) G\left(\xi\right)\, d\xi \tag{2.49}$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \xi^2 \left|G\left(\xi\right)\right|^2 d\xi \tag{2.50}$$

The last equality uses the fact that the pulse-shape $g(t)$ is real [64, p. 313]. Defining the factor ζ, depending on the pulse-shape as

$$\zeta = \frac{T^2}{4\pi^2} \frac{\int_{-\infty}^{\infty} \omega^2 \left|G\left(\omega\right)\right|^2\, d\omega}{\int_{-\infty}^{\infty} \left|G\left(\omega\right)\right|^2\, d\omega} \tag{2.51}$$

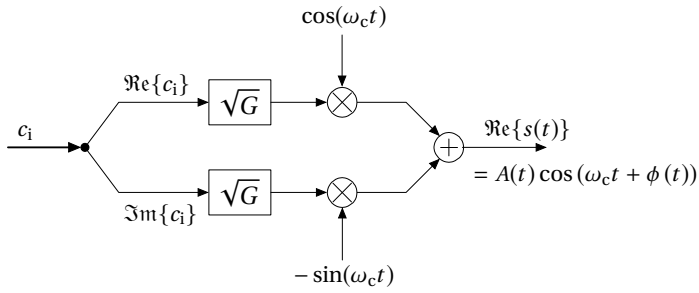


Figure 2.5.: Basic block diagram of a digital modulator (from [63, p. 24])

the MCRB for symbol clock estimation, normalized to the symbol period, is:

$$\frac{1}{T^2} \cdot \mathrm{MCRB}(\tau) = \frac{T}{8\pi^2 \zeta T_0} \frac{N_0}{E_\mathrm{s}} \tag{2.52}$$

The MCRB for timing estimation depends on the pulse-shape; wideband pulses are easier to synchronize to. The measurement duration with respect to the symbol duration and the signal-to-noise ratio are evidently important factors for the quality of the estimate as well.

2.6. Basic modulator and demodulator

To prepare the implementation of a transmitter and receiver pair for the experimental demonstration, the core functionality of a modulator and demodulator are briefly reviewed in this last section.

2.6.1. Modulator

A transmitter is a device that realizes the expression $s(t)$ from equation (2.7), for instance as a voltage to be fed to an antenna port. A quadrature-amplitude modulator, as shown in Figure 2.5, is a standard technique to achieve this. It is fed with the sequence of complex symbol values corresponding to the data to be transmitted. The real and imaginary parts (or in-phase and quadrature components) are filtered individually by identical pulse-shaping filters. Frequency upconversion from baseband to the target band is achieved by multiplication of the filter outputs with cosine and sine carriers from a local oscillator. Both components are summed to yield the final amplitude and phase-modulated signal.

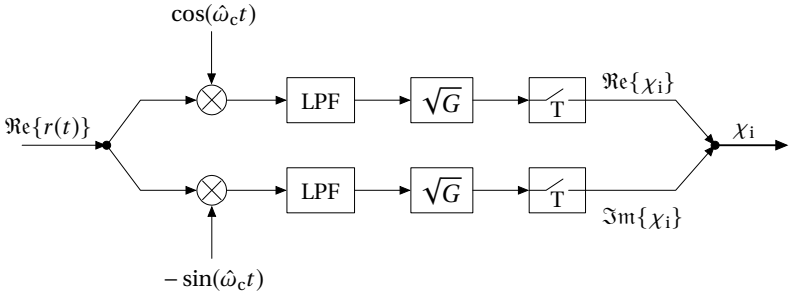


Figure 2.6.: Basic block diagram of a digital receiver

2.6.2. Demodulator

The purpose of the receiver is to estimate, from the noisy received signal $r(t)$, the original symbol sequence c_i sent by the transmitter. It reverses the processes from the modulator, by first downconverting $r(t)$ back into baseband. As the LO frequency used in the transmitter is not known, it uses an estimate $\hat{\omega}_\mathrm{c}$. This conversion creates an image frequency which is removed by a low-pass filter (LPF). The matched filter is applied; it has the same frequency response as the pulse-shaping filter in the transmitter. The matched filter removes a maximum of noise to yield a noisy symbol χ_i.

In a digital receiver, the symbols χ_i are sampled after receive-filtering. The sampling occurs at regular intervals according to the symbol clock estimate. This is the so-called instant of *maximum eye opening*. It ensures optimal sampling with the best de-mapping performance. Finally a decision-device determines the transmitted symbols c_i from χ_i according to a maximum-likelihood scheme.

3. Polarization Shift Keying

The principal objective of the present work is the analysis, simulation and implementation of polarization shift keying of the satellite channel. In this chapter, the performance of this modulation technique under ideal and realistic conditions is investigated, considering the prevalent conditions in the satellite scenario defined above.

To transmit a PolSK signal over satellite transponders, it has to be divided into its constituent horizontal and vertical carrier signals, which are processed individually for the major part of the signal path. This is a key difference to PolSK over optical fibres and leads to a certain number of challenges.

Both constituent carriers have to be *balanced* in their power and envelope variations to ensure that processing and external impediments (e.g. amplification or thermal noise) do not disproportionately impact only one polarization. By proposing a stochastic analysis of the PolSK symbol sequence in in-phase/quadrature form, this can be assured and some guidelines for good constellation design be derived.

Ground stations and satellite transponders are rarely designed to have equal path lengths on each polarization. Depending on the magnitude of this difference, two implications have to be distinguished. For small path length differences only the phase relationship is modified, leading to a rotation of the constellation in Stokes-space. This effect can be neutralized by a polarization-tracking device. Should the path length difference correspond to the same order of magnitude as the symbol duration, a receiver may encounter difficulties acquiring a proper lock on the symbol clock.

As mentioned in the previous chapter, the amplifier non-linearity poses a challenge as well. A PolSK signal of constant power, having all symbols placed on a Poincaré-sphere of fixed radius, will nevertheless experience envelope variations on each of the carriers. As these carriers are amplified separately on the transponder, their peak-to-average power ratio is relevant and needs to be controlled to limit distortion.

The PolSK demonstrator is used in measurements on a satellite transponder simulator and in a real satellite link. The equipment available for these experiments poses an additional difficulty. For redundancy reasons the local oscillators in the frequency converters are independent on each polarization. This leads to a frequency incoherency between the polarizations and an important advantage of

PolSK is lost: the ability to demodulate *directly* (i.e. incoherently) while entirely cancelling phase noise. The demonstrator is required to operate in this environment, without the general principle of PolSK modulation being put into question. It has to be ensured that measurement results obtained in these settings are applicable to the actual application target setting, where both polarizations are coherent.

The following sections describe the signal model for PolSK over satellite. The modulation technique is modified such as to lead to a balanced signal and to provide powerful hooks for synchronization. The frequency spectrum of this new PolSK signal is derived. The simulation software and the implemented demonstrator are presented. Simulated performance predictions in various scenarios are shown and discussed. Finally they are compared with the obtained measurement results. The chapter closes with a discussion on the applicability and reasons to use PolSK in satellite channels.

3.1. Theoretical analysis

The PolSK modulator design proposed in Section 2.3 splits the power of a light source and adjusts the phase relationship between both resulting orthogonal polarization components. Such a direct modulation procedure is hardly possible in microwave communications, where the modulated signal is traditionally created on a lower frequency and then upconverted into the target band. To transpose the optical PolSK modulation to microwave frequencies, it is necessary to consider both polarization components as independent carrier signals that, only upon combination on the antenna interface, will yield the requested state of polarization.

This microwave perspective is adopted for PolSK and the signal is conceived as a pulse-shaped, dual-carrier signal. The carriers represent the horizontal (x or H) and vertical (y or V) polarizations available in a satellite communication link. For the notation, a dual version of equation (2.6) is adequate:

$$s_{\mathrm{x}}(t) = A_{\mathrm{x}}(t)e^{\mathrm{j}\left(\omega_{\mathrm{c}}t+\phi_{\mathrm{x}}(t)\right)} \tag{3.1}$$

$$s_{\mathrm{y}}(t) = A_{\mathrm{y}}(t)e^{\mathrm{j}\left(\omega_{\mathrm{c}}t+\phi_{\mathrm{y}}(t)\right)} \tag{3.2}$$

Assuming clock-synchronous signalling on both polarizations and defining the scalar symbol values as $c_{\mathrm{i}} = A_{\mathrm{i}}e^{\mathrm{j}\phi_{\mathrm{i}}}$, the expression is rewritten in vector form:

$$\mathbf{s}(t) = e^{\mathrm{j}\omega_{\mathrm{c}}} \sum_i \begin{pmatrix} c_{\mathrm{i,x}} \\ c_{\mathrm{i,y}} \end{pmatrix} g(t - iT) \tag{3.3}$$

The symbol vector $\mathbf{c}_\mathrm{i} = \left(c_{\mathrm{i,x}}\, c_{\mathrm{i,y}}\right)^\mathrm{T}$ contains two amplitude and two phase values. The Stokes-vectors, as dictated by the PolSK symbol constellation, need to be translated into this representation prior to modulation.

3.1.1. Signal-to-noise ratio

In analogy to equation (2.8), an average signal power P_s and the corresponding average symbol energy E_s can be defined:

$$P_\mathrm{s} = \frac{1}{2T_\mathrm{m}} \int_{T_\mathrm{m}} \mathbf{s}(t)^\mathrm{H} \mathbf{s}(t)\, dt = \frac{E_\mathrm{s}}{T} \tag{3.4}$$

The superscript $^\mathrm{H}$ denotes a vector's hermitian, i.e. transposed and complex conjugate. In this way, the signal power is defined as the sum power of both individual bandpass signals. Adhering to the definition in equation (2.8), the symbol energy becomes:

$$E_\mathrm{s} = \frac{1}{2}\mathrm{E}\left[\mathbf{c}_\mathrm{i}^\mathrm{H}\mathbf{c}_\mathrm{i}\right] = \frac{1}{2}\mathrm{E}\left[S_0\right] \tag{3.5}$$

The second equality is verified by equation (2.4). As above, noise power in the receiver is limited by the reception filter. As two carrier signals are received, there are two filters. In consequence the received noise power is proportional to twice the signalling bandwidth:

$$P_\mathrm{n} = \frac{1}{2}\,(2B)\,(2N_0) = \frac{2}{T}N_0 \tag{3.6}$$

This leads to the following expression for the signal-to-noise ratio:

$$\mathrm{SNR} = \frac{P_\mathrm{s}}{P_\mathrm{n}} = \frac{E_\mathrm{s}}{T}\frac{T}{2N_0} = \frac{E_\mathrm{s}}{2N_0} \tag{3.7}$$

A factor ½ appears, as opposed to the expression E_s/N_0 derived in the previous chapter. Although this might suggest that the SNR for a dual-polarization signal is only half as large as for a classical bandpass signal, this coefficient stems from the definition of the signal energy as the sum energy on both polarizations. In consequence, the noise power spectral density also has to be considered twice.

In order to achieve the same SNR as for a traditional signal on a single polarization at a fixed symbol rate, the dual transmission requires twice the power inside twice the bandwidth. This is easily guaranteed in satellite communications, where two transponders will automatically dispense twice the power of a single one. This difference has to be considered when comparing dual-polarization with single-polarization signalling techniques; to account for the doubled power and bandwidth requirements, the bit-rates should be doubled for the latter.

In practice it can be the case that the SNR differs on each individual polarization. For all simulation and measurement results shown below, both signal and noise power are equally split between polarizations. Then, the SNR on each polarization individually equals the total, dual-polarization SNR:

$$\mathrm{SNR}_{\mathrm{dual}} = \frac{E_\mathrm{s}}{2N_0} = \frac{E'_\mathrm{s}}{N_0} = \mathrm{SNR}_{\mathrm{single}} \tag{3.8}$$

The apostrophed symbol energy E'_s represents the symbol energy per polarization, analogous to the situation in equation (2.10).

In Section 3.3 below, different constellations are compared by their error robustness. The signal-to-noise ratio is expressed as E_b/N_0 such that the results can easily be compared with those from the literature:

$$\frac{E_\mathrm{b}}{N_0} = \frac{2}{b} \cdot \mathrm{SNR} \tag{3.9}$$

The variable b stands for the number of bits transmitted per dual-polarization symbol.

3.1.2. Symbol representation

Equation (3.3) describes a dual carrier signal with amplitude and phase modulation, whereas a PolSK symbol constellation is defined in Stokes-space. To derive the amplitude and phase values from the Stokes-vector associated to each symbol, the formulas from equation (2.4) need to be inverted.

When looking at the phases, only their difference δ has an influence on the Stokes-parameters. Therefore, when transforming Stokes-parameters into amplitude and phase-representation, the carrier phase on one polarization can be freely chosen. This freedom is not fully recognized in the PolSK literature, where, according to equation (2.25), the phase difference δ is created by simply modulating the phase on each polarization by $\pm\delta/2$.

Here, I propose an alternative method of modulation that is more suitable for application in microwave settings. The freedom to choose one phase parameter is used to introduce pilot symbols into the signal that facilitate symbol clock recovery and polarization tracking. Furthermore, the pilot symbols allow the recovery of lost frequency coherency between the polarizations, which should permit the measurements on the available equipment. This modification entirely preserves the original PolSK signal and a receiver may or may not make use of the proposed improvement. To present this improvement, the concept of *fixed* and *moving* polarizations is introduced, constellation design is discussed and, finally, the insertion of the pilots is pointed out.

When synthesizing a requested state of polarization, absolute phase values are irrelevant and one may choose an arbitrarily fixed value $\phi_x = \xi$ for the horizontal polarization and derive ϕ_y from δ such as to produce the required result:

$$\phi_y = \xi - \delta \tag{3.10}$$

In this case, the horizontal polarization is called the *fixed* polarization, while the vertical is the *moving* polarization. The moving polarization changes its phase for each symbol. The expressions received for c_x and c_y by inverting (2.4) depend on whether the horizontal polarization is *fixed*, by $\phi_x = \xi$, or the vertical, by $\phi_y = \upsilon$:

fixed $\phi_x = \xi$	fixed $\phi_y = \upsilon$
$c_x = \sqrt{\frac{S_0+S_1}{2}}e^{j\xi}$	$c_x = \frac{S_2+jS_3}{\sqrt{2S_0-2S_1}}e^{j\upsilon}$
$c_y = \frac{S_2-jS_3}{\sqrt{2S_0+2S_1}}e^{j\xi}$	$c_y = \sqrt{\frac{S_0-S_1}{2}}e^{j\upsilon}$

(3.11)

3.1.3. Constellation design

The choice of a constellation in Stokes-space, where symbols should have a large minimum distance between themselves, has consequences on the average and peak power distribution between polarizations as well as on the presence of residual carrier signals in the final waveforms. To balance these parameters between horizontal and vertical carriers, a set of criteria has been determined.

Noise power can be expected approximately equal on both polarizations. It seems wise to equally distribute the signal power as well in this case. In addition, an equal peak-to-average power ratio ensures equal amplifier operating points as well as comparable signal distortion on each polarization.

A PolSK signal contains residual carriers [6]. It is judged beneficial that the power of discrete spectral lines is equally distributed between both polarizations.

Finally, the present work considers only constant-power constellations. This is a prerequisite for applying the maximum-likelihood de-mapping technique proposed in [5].

In the following, the expressions from equation (3.11) are used to transform these specifications into tangible conditions on the PolSK constellation. The dual symbol sequence is split into its two constituent I/Q symbol sequences. As it is important to emphasize the temporal nature of the subscripted i here, it is replaced by a notation in square brackets:

$$c_x[i] = A_x[i]e^{j\phi_x[i]} \tag{3.12}$$

$$c_y[i] = A_y[i]e^{j\phi_y[i]} \tag{3.13}$$

Consequently, i is the discrete-time index. These sequences are modelled as random processes with all symbols chosen independently and with equal probability among the constellation points. This permits the expected statistical moments for the sequence to be computed by averaging over all symbols in the constellation:

$$\mathrm{E}\big[c[i]\big] = \mathrm{E}\big[c_\mathrm{m}\big] = \frac{1}{M}\sum_{m=0}^{M-1} c_\mathrm{m} \tag{3.14}$$

$$\mathrm{Var}\big[c[i]\big] = \mathrm{Var}\big[c_\mathrm{m}\big] = \frac{1}{M}\sum_{m=0}^{M-1} (c_\mathrm{m})^2 - \left(\frac{1}{M}\sum_{m=0}^{M-1} c_\mathrm{m}\right)^2 \tag{3.15}$$

Here m is an index running over all M symbols in the constellation. Due to this equivalence, the indices i and m are omitted in the following. By formulating requirements on the symbol sequences and transforming those into Stokes-space, the conditions for constellation design are derived.

Energy repartition

It is a straightforward choice to equally divide the average symbol energy between the polarizations:

$$E_\mathrm{x} = E_\mathrm{y} \tag{3.16}$$

$$\Longleftrightarrow \mathrm{E}\big[A_\mathrm{x}^2\big] = \mathrm{E}\big[A_\mathrm{y}^2\big] \tag{3.17}$$

$$\Longleftrightarrow \mathrm{E}\big[S_0 + S_1\big] = \mathrm{E}\big[S_0 - S_1\big] \tag{3.18}$$

$$\Longleftrightarrow \mathrm{E}\big[S_1\big] = -\mathrm{E}\big[S_1\big] = 0 \tag{3.19}$$

The power repartition between polarizations is thus dictated by the S_1 parameter. In a balanced constellation, it is equal to 0 on average.

To balance signal distortion from non-linear amplification, a similar requirement is set on the peak power:

$$\max\{S_0 + S_1\} = \max\{S_0 - S_1\} \tag{3.20}$$

$$\Longleftrightarrow \max\{S_1\} = \max\{-S_1\} \tag{3.21}$$

The second line results from the constant-power requirement, which fixes S_0 for all symbols.

Residual carriers

The carrier on the *fixed* polarization is only amplitude-modulated and shows a residual in its frequency spectrum. Its magnitude should be equal on both polar-

izations. This is ensured by respecting the following relationship for *fixed* polarizations:

$$\mathrm{E}\big[|c_\mathrm{x}|\big] = \mathrm{E}\big[|c_\mathrm{y}|\big] \tag{3.22}$$

$$\Longleftrightarrow\ \mathrm{E}\Big[\sqrt{\frac{S_0+S_1}{2}}\Big] = \mathrm{E}\Big[\sqrt{\frac{S_0-S_1}{2}}\Big] = m_0 \tag{3.23}$$

m_0 gives an indication of the residual's amplitude. On the *moving* polarization, carrier residuals should be entirely avoided. For this to be the case, the sequence average has to vanish on the *moving* polarization:

$$\mathrm{E}\big[c_\mathrm{x}\big] = \mathrm{E}\big[c_\mathrm{y}\big] = 0 \tag{3.24}$$

$$\Longleftrightarrow \mathrm{E}\Big[\frac{S_2+\mathrm{j}S_3}{\sqrt{2S_0-2S_1}}e^{\mathrm{j}\nu}\Big] = \mathrm{E}\Big[\frac{S_2-\mathrm{j}S_3}{\sqrt{2S_0+2S_1}}e^{\mathrm{j}\xi}\Big] = 0 \tag{3.25}$$

$$\Longleftrightarrow \mathrm{E}\Big[\frac{S_2+\mathrm{j}S_3}{\sqrt{2S_0-2S_1}}\Big] = \mathrm{E}\Big[\frac{S_2-\mathrm{j}S_3}{\sqrt{2S_0+2S_1}}\Big] = 0 \tag{3.26}$$

This last condition is fulfilled in all modern coherent modulation formats (e.g. in 2^n-PSK or 2^m-QAM) by their point-symmetric design. Note that equations (3.22) and (3.24) compare two *fixed* and two *moving* polarizations, respectively. On a real PolSK signal one polarization is fixed, while the other is moving.

Although constellation design does not imperatively need to respect all these constraints, they lead to a signal with similar behaviour on both polarizations, independently of which of them is *fixed* or *moving*.

Looking at the equations, it seems quite involved to deduce a constructive design prescription. This is especially true when, in addition, a high number of symbols has to be placed onto the Poincaré-sphere at a large mutual distance. In Section 3.2.2 below, a 16-point constellation design showing all these preferred properties is presented.

3.1.4. Synchronization

Polarization tracking and carrier recovery

In PolSK over optical fibre, polarization tracking is performed to account for slow fluctuations in the state of polarization [7]. In the present work, such fluctuations due to propagation effects are considered negligible (cf. Section 2.2.3). Still, received Stokes-vectors might be wrongly oriented on the Poincaré-sphere because of the practically unavoidable difference in path length between both polarizations.

This path length difference leads to a delay difference, which in turn leads to a phase difference. Prior to symbol de-mapping, this difference needs to be

estimated to recover the original phase relationship δ and to bring the Poincaré-sphere back into its original orientation. This is a synchronization problem that can be solved by transmitting phase-reference signals on both polarizations.

This is achieved by using the *fixed* polarization to transmit the local oscillator's phase over both polarizations. Consequently, the *fixed* carrier will have the phase set to zero: $\xi = 0$ and $\upsilon = 0$. The receiver, by comparing the phase values, can measure the difference in phase between both polarizations and derotate the received PolSK symbols on the Poincaré-sphere. To transmit this phase information on *both* polarizations, the *fixed* and *moving* carriers need to be exchanged from time to time (i.e. after a certain number of symbol-clock cycles). As the path length is a static quantity, this estimation can be performed over a very long time and is practically noiseless. For this reason, the PolSK phase-noise immunity is not affected.

If necessary, and such is the case for the planned measurements, these phase pilots can be used to recover a frequency difference between both polarizations as well. After an incoherent frequency conversion, phase noise is no longer correlated. The bandwidth for estimating phase and frequency can then be chosen larger, in accordance with the system's specific requirements on tracking and acquisition speed or allowed noise in the loop.

As the phase on the *fixed* polarization is zero at the symbol strobe, the proposed mechanism does not introduce any self-noise, as long as the symbol clock is precise.

Symbol-clock recovery

There are no symbol recovery schemes for polarization shift keying published in the literature. In optical communications, pulse-shaping is rare and rectangular pulses are usually employed, which are easy to recover. In the experimental demonstrations, it is either implicitly assumed, as in [15], or offline processing is used, as is done in [78]. Rectangular pulses are not acceptable in satellite communications, where stringent bandwidth limitations demand low excess bandwidth.

Many timing-error detectors rely, directly or indirectly, on a square-based non-linearity to extract timing information from a band-limited information signal [36, 37, 61]. This method suffers from high self-noise with the increasingly lower excess bandwidths used in contemporary satellite communications (For instance in DVB-S: 35 % [27], in DVB-S2: down to 20 % [28]). This self-noise can be suppressed using very narrow loop bandwidths, at the price of increased acquisition time [35]. For PolSK, long acquisition times are not desirable, because the direct-detection ability allows in principle the immediate start of the demodulation process.

In the present case, the dilemma is solved by exchanging *fixed* and *moving*

carriers after each single transmitted symbol. As is shown in the subsequent section, this introduces discrete spectral components into the PolSK signal at half the symbol rate above and below the carrier. Upon squaring, these components lead to a high-intensity spectral line at the symbol-frequency which should substantially reduce the self-noise influence. For PolSK, a squared baseband signal containing the timing information is readily available in the S_0-parameter, independent of any carrier frequency offsets or phase differences.

With this alternation between fixed and moving polarizations, square-based timing recovery systems can be employed with high fidelity and low roll-off values. In addition to the so-gathered symbol-timing information, phase pilots are transmitted at every second symbol strobe.

Note that this new method for PolSK modulation is totally transparent to the original PolSK. The Stokes-vectors themselves are unaffected such that PolSK receivers employing different recovery techniques will be able to decode the signal as well. This is true in particular for direct-detection receivers mixing both polarizations.

The concrete implementation of the presented mechanisms in the PolSK demonstrator is exposed in Section 3.2.2.

3.1.5. Frequency spectrum

The exchange of fixed and moving polarizations after every symbol, as proposed in the previous section, introduces a periodic component into the signal that is visible in the frequency spectrum.

The mean and autocorrelation function of the symbol sequence $c[i]$ on one polarization are computed in the following and used to derive the signal's spectral shape and the energy allocation. It is assumed that $c[i]$ fulfils all requirements from Section 3.1.3, whereby the carrier is *moving* for even and *fixed* for odd time indices:

$$c[i] = \begin{cases} A[i]e^{\mathrm{j}\phi[i]} & \text{if } i \text{ even} \\ A[i] \geq 0 & \text{if } i \text{ odd} \end{cases} \tag{3.27}$$

The sequence mean values can be evaluated using the requirements from Section 3.1.3:

$$\mathrm{E}\big[c[i]\big] = \begin{cases} \mathrm{E}\big[A[i]e^{\mathrm{j}\phi[i]}\big] = 0 & \text{if } i \text{ even} \\ \mathrm{E}\big[A[i]\big] = m_0 \geq 0 & \text{if } i \text{ odd} \end{cases} \tag{3.28}$$

To compute the sequence's autocorrelation function $R_{cc}(i;k)$, different cases have

to be distinguished:

$$R_{cc}[i;k] = \mathrm{E}\big[c[i]^* c[k]\big] = \begin{cases} \mathrm{E}\big[A[i]^2\big] = 2E'_\mathrm{s} & \text{for } i = k \\ \mathrm{E}\big[A[i]A[k]e^{\mathrm{j}(\phi[k]-\phi[i])}\big] = 0 & \text{for } i, k \text{ even} \\ \mathrm{E}\big[A[i]A[k]e^{-\mathrm{j}\phi[i]}\big] = 0 & \text{for } i \text{ even}, k \text{ odd} \\ \mathrm{E}\big[A[i]A[k]e^{\mathrm{j}\phi[k]}\big] = 0 & \text{for } i \text{ odd}, k \text{ even} \\ \mathrm{E}\big[A[i]A[k]\big] = E'_\mathrm{R} \geq 0 & \text{for } i, k \text{ odd} \end{cases} \tag{3.29}$$

E'_s is the symbol-energy on one polarization only. As symbols are chosen independently, the covariance between $A[i]$ and $A[k]$ will be zero. Therefore:

$$E'_\mathrm{R} = \mathrm{E}\big[A[i]A[k]\big] = \mathrm{Cov}\big[A[i];A[k]\big] + \mathrm{E}\big[A[i]\big]^2 = m_0^2 \tag{3.30}$$

The autocorrelation function is different for odd and even values of i. $c[i]$ is thus a cyclostationary process of period 2 [63, p. 70]. An average mean and autocorrelation over one period can be computed; they have the following expressions upon careful inspection:

$$\overline{m_c} = \overline{\mathrm{E}\big[c[i]\big]} = \frac{\sqrt{E'_\mathrm{R}}}{2} \tag{3.31}$$

$$\overline{R_{cc}[n]} = \overline{\mathrm{E}\big[c[i]^* c[i+n]\big]} = \left(2E'_\mathrm{s} - \frac{E'_\mathrm{R}}{2}\right)\delta[n] + \frac{E'_\mathrm{R}}{4}\left(1 + \cos\left(\frac{\pi}{n}\right)\right) \tag{3.32}$$

By applying the discrete Fourier-transform (DFT) to the average autocorrelation function, the average power spectral density is obtained [64, p. 320f.]:

$$\overline{S_{cc}[\mu]} = \mathcal{F}\left\{\overline{R_{cc}[n]}\right\} \tag{3.33}$$

$$= \underbrace{\frac{1}{N}\left(2E'_\mathrm{s} - \frac{E'_\mathrm{R}}{2}\right)}_{\text{information payload}} + \underbrace{\frac{E'_\mathrm{R}}{4}\delta[\mu]}_{\text{residual carrier}} + \underbrace{\frac{E'_\mathrm{R}}{8}\left(\delta[\mu + \frac{N}{2}] + \delta[\mu - \frac{N}{2}]\right)}_{\text{symbol clock}} \tag{3.34}$$

N is the number of samples in the DFT. The spectrum of the symbol sequence consists of three distinct parts. A certain fraction of the signal energy is equally distributed into all frequency bins. Of the remaining energy, half is allocated to the bin at the DFT centre, the other half is equally distributed into both extremal bins. Upon modulation, these different parts correspond to the digitally modulated carrier, the residual carrier and spectral components at ± half the symbol rate. Evidently, the shape of the transmit filters will also influence the continuous part of the spectrum.

This representation clearly shows how the total signalling power is distributed between the actual symbol information and the discrete spectral components. It is valid on both polarizations to the same extent. Simulations reveal that

this does not change the modulation's power efficiency with respect to PolSK as described in the literature. The quantity E'_R determines how much energy is allocated to the discrete spectrum. It can be computed from the PolSK constellation using equations (3.23) and (3.30).

The measured frequency spectrum of such a PolSK carrier is displayed in Figure 3.4 at the end of Section 3.2. The root-raised cosine spectral shape is clearly discernible, dominated by the three spectral lines as predicted by equation (3.34). The side residuals are each 3 dB below the central carrier residual.

3.2. Tools and methods

The analysis of polarization shift keying is focussed on power efficiency and on the impact of various impairments. It is conducted in two stages. First a simulation program is used to try different parameter sets, symbol constellations and impairment combinations.

Based on the insight from the simulations, a PolSK demonstration modulator and demodulator have been implemented for the second stage of the analysis. They are used for the experimental measurements and for the satellite demonstration of polarization shift keying.

In practical modulation settings, transmission data is encoded before modulation. *Source-coding* compresses the data by removing as much redundancy as possible. Subsequently, a *channel-code* provides the ability for the receiver to detect and correct wrong symbol decisions by inserting again some well-defined redundant information into the signal. For simplicity and scope, the focus is set on the bare physical modulation and coding is, in consequence, not considered.

3.2.1. Software simulator

The implemented simulation software, written in the `Matlab` programming language, simulates a basic dual-polarization transmitter and receiver as well as the associated satellite channels. Besides polarization shift keying, it supports various other linear modulation techniques on one or both polarizations. The channel simulator holds a range of expected impairments at its disposal. The principal output is a symbol-error rate estimate for the selected set of parameters.

The signal flow in the simulator is illustrated in Figure 3.1. A modulator is fed with a sequence of random input values and maps them to a sequence of symbols according to the selected constellation. This symbol sequence is pulse-shaped and transferred to the channel simulator.

The channel simulator models the different impairments to be expected in the satellite link. They are selected and configured among non-linear amplifica-

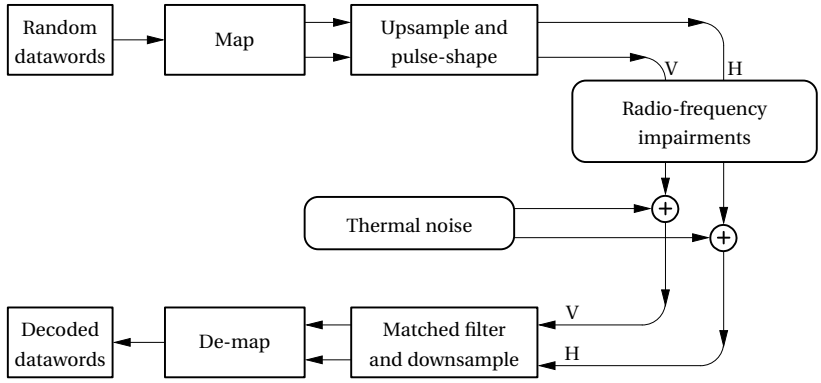


Figure 3.1.: Waveform-level simulator steps

tion, phase noise and frequency drift and are applied in the order determined by the user. Thermal noise, modelled as additive white Gaussian noise, is considered in the downlink only and, as such, is added in a final step.

The receiver model applies the matched filter on the impaired and noisy waveform. Symbols are de-mapped according to the maximum likelihood principle. A final step compares the de-mapped data stream with the initially transmitted random values and counts the number of wrongly detected symbols. This is how an SER estimate is produced. Symbol-clock jitter is not simulated; the optimal downsampling instant is hard-coded into the software.

Besides the SER as a main output, all intermediate waveforms generated during a simulation run are saved for later inspection. In this way, the effects of impairments or signal processing can be visualized and interpreted. It is possible to inspect the time signal, unprocessed or as an eye-diagram, and the frequency spectrum. Scatter diagrams can be displayed in in-phase and quadrature, or in Stokes-space representation. Finally, the average and peak power of waveforms are computed.

The simulation software has been implemented according to the guidelines and models proposed in [49]. Random phenomena (i.e. thermal noise and phase noise) are simulated by creating large, sampled representations of the modelled random process: this is called the Monte-Carlo approach. A modular implementation allows the software to be easily extended by new modulation types and channel impairments.

The simulator has been validated by comparing the created symbol-error rate curves with examples shown in the literature and to the output of `Matlab`

built-in functions[1]. A more detailed explanation of the simulator's inner workings and the employed models is to be found in Appendix A.

3.2.2. Demonstrator

For the measurements and the PolSK demonstration over satellite, a demonstration transmitter and a receiver have been developed. Both systems were implemented on a programmable logic platform with analogue-to-digital (ADC) and digital-to-analogue converters (DAC) as inputs and outputs respectively.

Unlike the simulations, where different constellations and parameter sets can be tried out with ease, the signalling properties of the demonstrator are fixed. An important design choice is the selected PolSK constellation.

Symbol constellation

The most popular modulation method in satellite communications is quaternary phase shift keying (QPSK). It is an inevitable reference due to its simplicity, robustness and power efficiency. To allow comparison with QPSK, a 16-point PolSK constellation is selected. In this way, the same spectral efficiency as for QPSK on both polarizations is achieved (dual QPSK: 2×2 bit$persymbol$, 16-PolSK: 1×4 bit per symbol). Note that no PolSK implementation of such a high spectral efficiency has been reported as of today.

In line with the limitation of the scope to constant-power PolSK, the constellation is created by distributing 16 points on the two-dimensional Poincaré-sphere. By carefully inspecting and adjusting the result of a minimum distance optimization on a spherical surface, conducted by Sloane in [76], a good 16-point constellation, compliant with all requirements from Section 3.1.3, is found. The table below contains the constellation symbols' final positions, in spherical coordinates.

#	0	1	2	3	4	5	6	7
ϑ	$\pi-\beta$	β	α	$\pi-\alpha$	$\pi-\beta$	β	α	$\pi-\alpha$
φ	$\frac{\pi}{2}$	$\frac{3\pi}{4}$	$\frac{\pi}{2}$	$\frac{3\pi}{4}$	π	$-\frac{3\pi}{4}$	π	$-\frac{3\pi}{4}$
#	8	9	10	11	12	13	14	15
ϑ	$\pi-\beta$	β	α	$\pi-\alpha$	$\pi-\beta$	β	α	$\pi-\alpha$
φ	$-\frac{\pi}{2}$	$-\frac{\pi}{4}$	$-\frac{\pi}{2}$	$-\frac{\pi}{4}$	0	$\frac{\pi}{4}$	0	$\frac{\pi}{4}$

[1] Specifically, the `berawgn` function, which provides analytically evaluated symbol-error rate estimates.

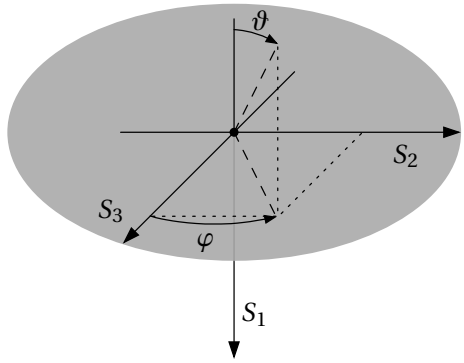


Figure 3.2.: Angle definitions for Stokes-coordinates

All constellation points have equal power and lie therefore on a spherical surface with radius $r = S_0 = 2E_s$. The angle ϑ is measured from the $-S_1$ axis, while φ increases from the S_3 towards the S_2 axis. This representation, illustrated in Figure 3.2, allows the elegant notation of the coordinates shown above. Two angular parameters remain to be defined:

$$\alpha = 38{,}5° = 0{,}672\,\text{rad}$$
$$\beta = 76{,}4° = 1{,}333\,\text{rad}$$

A three-dimensional representation of the constellation in Stokes-space is shown in Figure 3.3. The next paragraph describes step by step how this constellation fulfils all requirements from Section 3.1.3.

The value of S_1 is dictated by the angle ϑ and determines the power allocation between the H and V polarizations. S_1 takes four different values, chosen symmetrically around the origin; this satisfies the requirement from equation (3.21). In Stokes-space this creates a set of four planes, arranged symmetrically around the $S_1 = 0$ plane, which by intersection with the spherical surface defined by S_0 lead to four circles onto which symbols can be placed by picking the remaining parameters S_2 and S_3. By placing onto each of these circles the same number of symbols as on its symmetrically opposed counterpart, equations (3.19) and (3.23) are fulfilled. Making sure that on each of the circles the symbols are placed in regular arcs (like in a PSK constellation), fulfils equation (3.26) for each circle (i.e. each value of S_1) individually, and thus also for the entire constellation.

For the proposed constellation, $E'_R = 1{,}8E'_s$. In consequence, 55 % of the total power is allocated to information transport, while the discrete spectral components carrying frequency and timing information share the remaining power.

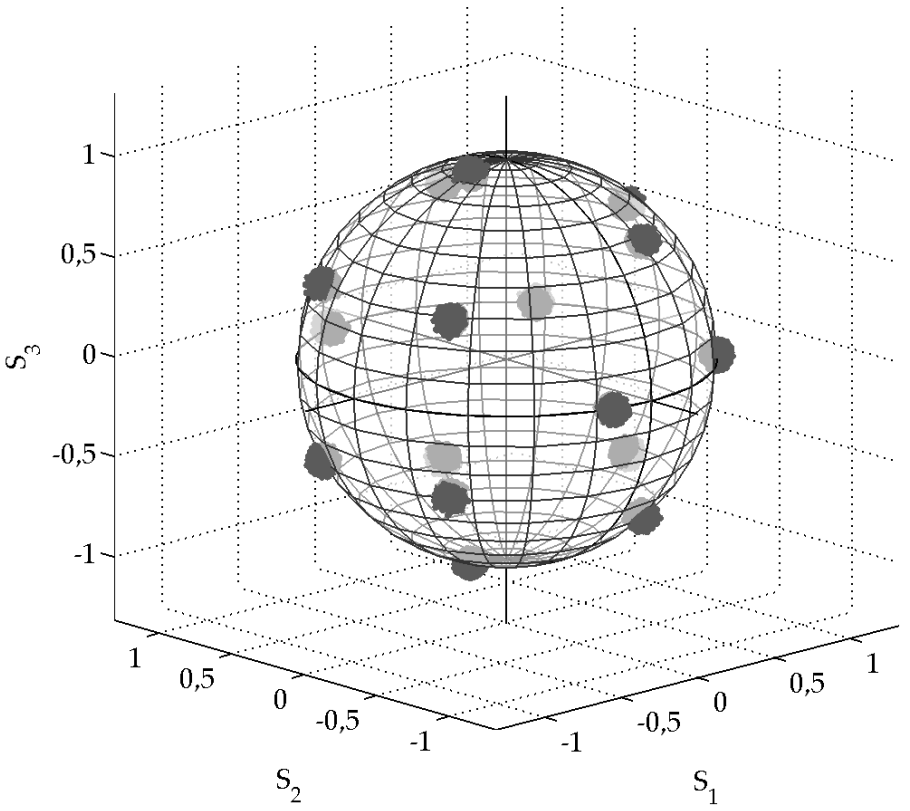


Figure 3.3.: 16-PolSK symbol constellation

General structure of the demonstrator

The remainder of the present section describes the implemented demonstrator system. The level of detail is limited to what is required to understand the meaning of the measurement results presented further below. For a more thorough presentation of the signal processing and to inspect the relevant block diagrams, the interested reader is referred to Appendix B.

The demonstrator system is specified to process signals on an intermediate frequency (IF) of 70 MHz. This is the prevalent IF in satellite communications, with many generic parts and measurement equipment available. Two D/A converters and two A/D converters realize and sample the horizontal and vertical carrier signals, respectively. The superimposition of both carriers in orthogonal polarization modes forms the actual PolSK signal. The DACs and ADCs operate on a sampling frequency of 245,76 MHz, fast enough to comfortably handle the 70 MHz analogue waveform.

Even though a satellite transponder can usually accommodate 36 MHz-wide signals, the selected symbol clock frequency is 1,024 MHz. Thereby the computational complexity in the receiver is limited to a reasonable level and the requirement for identical path lengths is less stringent. In addition, such a narrowband signal can be used on transponders with non-aligned channel definitions (cf. Section 2.2.2).

PolSK transmitter

The transmitter is implemented as a dual quadrature amplitude modulator. A functional block diagram is shown in Figure B.1 in Appendix B.

A linear feedback shift register (LFSR, [41]) generates pseudo-random numbers from zero to 15 at the symbol clock rate. This data addresses a lookup table containing the symbol constellation in I/Q notation. After each symbol strobe, the table contents are flipped to change the fixed polarization from the horizontal to the vertical polarization and vice-versa, in accordance with the method outlined in Section 3.1.4.

A root-raised cosine filter upsamples and shapes the resulting baseband symbol sequence. The roll-off is $\alpha = 0{,}35$. A second upsampling stage, based on a cascade of half-band filters, increases the internal sampling frequency to the DAC sampling rate.

Subsequently, the baseband signal is upconverted to the target frequency by multiplication with a fixed 70 MHz signal from a numerically controlled oscillator. A digital automatic gain-control unit (AGC) adjusts the signal amplitude for an optimal use of the DACs' dynamic range. Finally, the digital signal is physically realized through two synchronized DACs and reconstructed by a pair of analogue

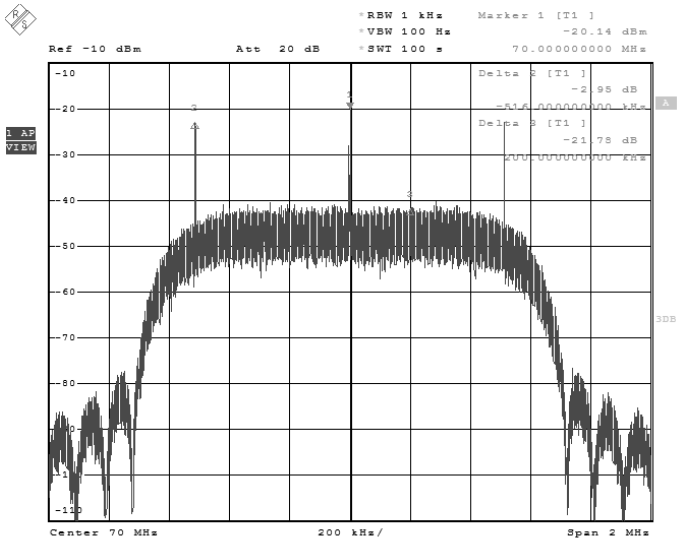


Figure 3.4.: Spectrum analyser screenshot of the PolSK transmitter's H-pol output. The roll-off is $\alpha = 0{,}35$. Three discrete spectral lines are discernible, at the centre frequency and at $\pm 1/2T$

low-pass filters.

The frequency spectrum of the so-created waveform is displayed in Figure 3.4. It is a spectrum analyser screenshot taken on the running transmitter's horizontal output port.

PolSK receiver

The PolSK receiver inverts the transmitter's processing steps to determine the initially transmitted data sequence. The receiver is slightly more complex, as additional circuitry is required to estimate unknown timing variables before effective demodulation can occur. Figure B.4 in Appendix B displays the receiver's block diagram.

Two synchronous ADCs sample the IF PolSK signal. A digitally controlled preamplifier adjusts the signal level for optimal dynamic range. Two wide bandpass filters remove excess noise in out-of-band regions. This relaxes the requirements on filtering after downconversion and thereby reduces delay in the tracking loops.

The signal is downconverted into baseband by multiplying it with a pair of local oscillators. In a plain PolSK receiver, the phase difference between these os-

cillators is adjusted to compensate the difference in path delay between both polarizations. In the present case, all experiments involving path delay differences will additionally suffer from non-coherent frequency conversion. For this reason a frequency tracking system is implemented using the pilot symbols described above. For all other experiments, this subsystem can be disabled, leading to incoherent downconversion with a single, fixed local oscillator.

Multiplying the signal with the LOs leads to a spectral copy at 105,76 MHz which is removed by a subsequent filter stage. This band-limited signal is downsampled and fed to the matched filters.

The amount of noise removed by this operation depends on the signal's SNR, such that the resulting signal power is not known a priori. Therefore a second AGC provides a constant signal level to the final reception stages. This also constrains the loop gains for the symbol- and carrier-tracking PLLs within defined limits.

At this point the baseband components are transformed into Stokes-vectors, using the formula in equation (2.4). The symbol-recovery loop determines the symbol strobe using the oversampled S_0-parameter. The optimal instant for detection is selected and the corresponding Stokes-vector is de-mapped by a minimum distance decision algorithm.

Using a shift register identical with that in the transmitter, the receiver is able, upon register synchronization, to predict future symbols [14]. This permits symbol errors to be detected and counted in order to produce an estimate of the error rate.

Symbol clock recovery

The square timing-error detector, described in Section 2.5.1, is combined with a PLL to estimate the symbol clock from the S_0 Stokes-parameter. This is an analogue technique, which can only be used because the available baseband signal is highly oversampled (30-fold).

The phase-locked loop, tracking the spectral line at $f = 1{,}024\,\text{MHz}$ in S_0, has a bandwidth of 100 Hz, with an additional low-pass filter ($f_{3\text{dB}} = 10\,\text{kHz}$) to remove the strong disturbance from the residual carrier. In this way, a local copy of the symbol clock is created.

Using sample counters, the index n at which the symbol strobe is located, is determined. This index is forwarded to the final downsampling stage and selects the optimal decision sample.

Flip-flops are used to divide the frequency of the symbol clock by two. The divided clocks control the sampling of the quadrature values on both polarizations. These quadrature values are used as an error measure by the carrier-recovery system.

Carrier frequency and phase recovery

The carrier-recovery system in the receiver is optional and can be activated if frequency coherency is lost between the horizontal and vertical polarizations. It uses the knowledge that, by construction, every second PolSK symbol has a phase of zero. It uses the hints from the symbol-recovery to sample the quadrature values of both signals at each second symbol strobe. Knowing that

$$\phi_x = 0 \iff X_Q = 0 \tag{3.35}$$

$$\phi_y = 0 \iff Y_Q = 0 \tag{3.36}$$

these samples are directly used as a phase error, leading to a sine-shaped detection characteristic. The phase-errors on each polarization are tracked independently by two PLLs. The carrier-recovery PLLs are of type 2, so that phase and frequency errors can be tracked without residual error. The loop bandwidth is 10 kHz.

When a carrier recovery system is used, polarization de-rotation is no longer required and is therefore not implemented. For polarization de-rotation, the difference between both quadrature values can be taken as an error signal, and the loop bandwidth can be reduced to the order of Hz.

The calculation of S_0 does not depend on a previous carrier-lock and can therefore be accomplished with fixed oscillators set at the approximate centre frequency. On the other hand, the carrier-recovery system requires previous symbol-lock. For this reason, reception is performed with fixed local oscillators, until the symbol-clock PLL is successfully tracking.

It is not known in advance which of the polarizations is *fixed*. The carrier-recovery system will therefore try both alternatives until a lock can be achieved.

3.3. Simulation results

The simulations are carried out to provide first impressions on the behaviour of a PolSK signal. They allow different constellations to be quickly compared and various parameter and impairment combinations to be tried out.

The output of a simulation run is presented as a symbol-error rate curve, containing the signal-to-noise ratio in the abscissa and the estimated SER in the ordinate. *Symbol* errors, rather than *bit* errors are indicated. This avoids choosing a symbol-to-bits correspondence for every one of the constellations. Many different choices are possible on the spherical surface, all leading to a slightly varying error performance and it is not trivial to unambiguously determine an optimal solution.

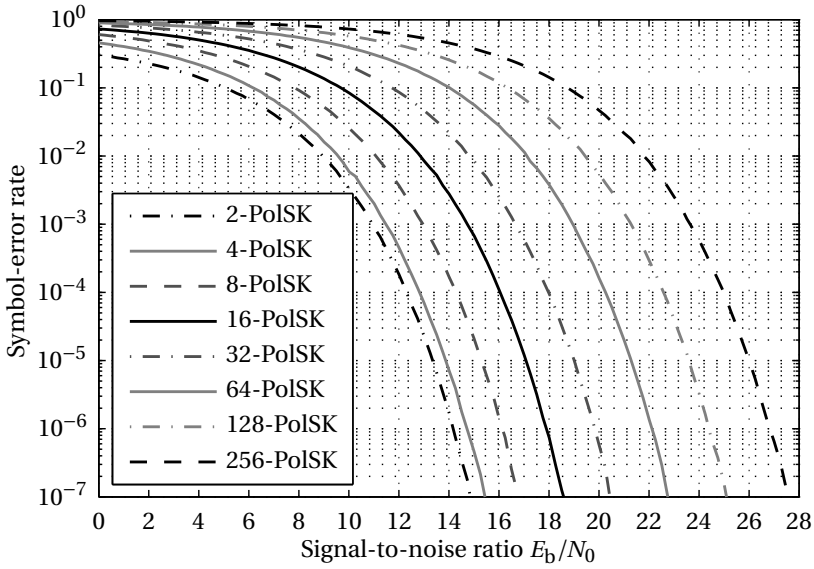


Figure 3.5.: Estimate of the power efficiency for different PolSK constellations by simulating the SER as a function of E_b/N_0.

In most of the following cases, the SNR required for an error rate of $1 \cdot 10^{-4}$ is chosen as representative for comparisons between the results of different simulations. As only uncoded modulation is simulated, this is estimated to be the point at which the error-correcting code would fill in and suppress further errors below a measurable threshold.

3.3.1. Different PolSK constellations

A first simulation estimates the symbol-error rate of polarization shift keying constellations of varying spectral efficiency in plain thermal noise.

Constellations with 2 up to 256 symbols have been simulated. Except for the dual, 16- and 256-point PolSK, all constellations have been taken as is from the sphere-packing database in [76]. The dual constellation simply consists of two diametrically opposed Stokes-vectors, while the 16-point PolSK is the constellation presented above. The 256-point constellation is generated by the packing algorithm described in [55].

The results are plotted in Figure 3.5. The simulated symbol-error probability corresponds very well to the results published in [5] for the constellations up to

32 points[2]. This effectively validates the PolSK implementation in the simulation program and confirms that the newly introduced modulation technique does not alter the power efficiency.

The difference in power efficiency between dual and quadruple PolSK is relatively small. Only in exceptional cases should one be forced to use the dual in favour of the quadruple variant. For every additional bit, the cost in required asymptotic SNR increases with the constellation size and approaches 2,4 dB, for the steps from 32 to 64 and above.

The 64-, 128- and 256-point constellations have been introduced for comparison with the modulation techniques of similar spectral efficiency readily available in satellite communications, i.e. 8-PSK and 16-APSK. However, the power cost for PolSK is prohibitively high for these spectral efficiencies. This insight led to the decision to focus the further analysis on the 16-PolSK constellation, corresponding to the smallest spectral efficiency in use in satellite communications: QPSK. A direct comparison of 16-PolSK with existing modulation techniques is provided further below.

3.3.2. Realistic operating conditions and phase-noise immunity

The simulations presented next have been undertaken to estimate the symbol-error performance of the proposed 16-PolSK constellation in a realistically dimensioned satellite transmission, including phase noise and non-linear amplification, in addition to plain thermal noise.

Phase noise

First, the impact of phase noise is considered. Simulations with identical phase noise on both polarizations should confirm the respective immunity described in the literature. Opposed to identical phase noise, fully independent phase noise is simulated, to illustrate the performance drop expected when this immunity cannot be exploited.

As a baseline, the applied phase-noise power and spectral shape is loosely based on an SES technical recommendation for LNBs [69]. Phase-noise levels ten and 100 times higher are simulated as well. The applied phase-noise characteristic is displayed in Figure A.4 in Appendix A.

The simulation results are shown in Figure 3.6, together with the previous result for 16-PolSK in plain AWGN. For identical phase noise on both polarizations, as expected, no deterioration is encountered. The respective curves are exactly

[2]There are no published results for 64 points and above.

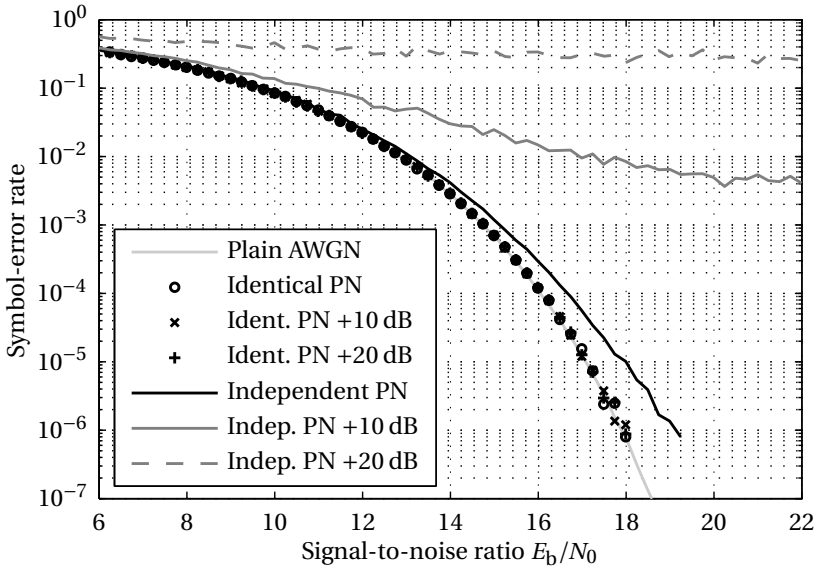


Figure 3.6.: Simulated 16-PolSK SER with phase noise

equal to the included reference case without phase noise[3].

With independent phase noise, at a level expected from widely-used consumer equipment, good reception is possible as well, at a slight power cost. With 10 dB more, reception is seriously hindered, while with 20 dB above the phase-noise mask, received data is more or less random.

The simulations reveal that the PolSK modulation can bear a certain level of independent phase noise with a tolerable loss in power efficiency. Above this level, successful reception is only possible when exploiting the immunity described in the literature. I would expect partially correlated phase noise can be separated into an identical and an independent contribution; only the independent contribution would then have an impact on the detection performance.

Non-linear amplification

Two different models are used to simulate the impact of non-linear amplification. The Saleh-model simulates a regular travelling wave tube amplifier (TWTA) [67, 49] while for the more modern, linearized amplifiers the characteristic is taken from a reference curve shown in the DVB-S2 standard description [28]. Details

[3]The variance of simulation results is higher at low SER, which is the reason for the visible differences.

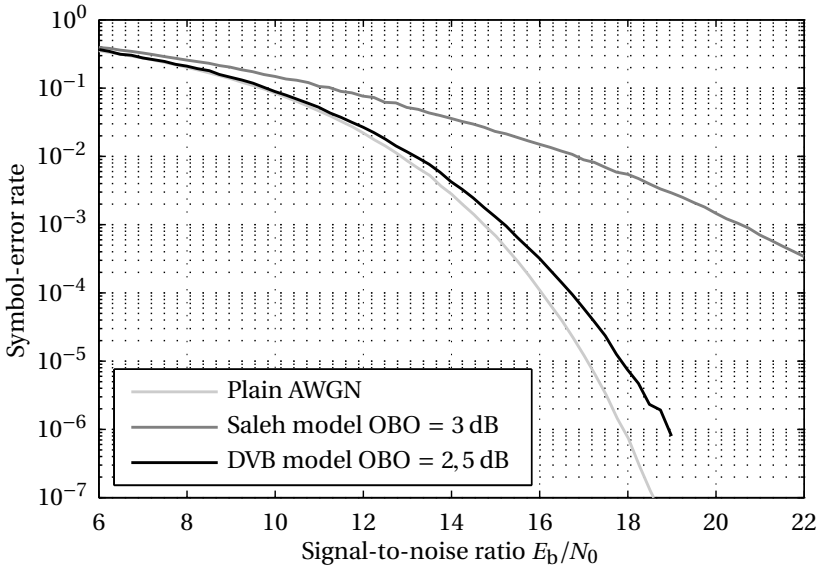


Figure 3.7.: Simulated 16-PolSK SER upon non-linear amplification

on the models are shown in Appendix A. For the simulations, both amplifiers are operated backed off, to allow reasonable reception quality.

The results are shown in Figure 3.7. A PolSK signal amplified with a regular TWTA suffers from unacceptably high distortion, even with an output back-off (OBO) of 3 dB. The simulations using the linearized tube model show that good results can be achieved even with less back-off, given more sophisticated amplifiers.

A constant-power constellation, like the proposed 16-PolSK, will, for symbols with high absolute values for the S_1 parameter, have both amplifiers in opposing states; while one is in saturation, the other will be in linear operation and vice versa. In that sense, phase distortion affects the signal doubly. This effect is reinforced in the simulation by the relatively steep phase characteristic of the Saleh-model. It can be judged that for this reason the problem is significantly less acute for the linearized tube, which shows plateaus in the AM-PM characteristic for the operating points in question (cf. Figure A.3 in Appendix A).

Realistic operating conditions

In a third simulation effort, both effects are analysed together. The plot in Figure 3.8 shows the simulation results using both amplifier types and various phase

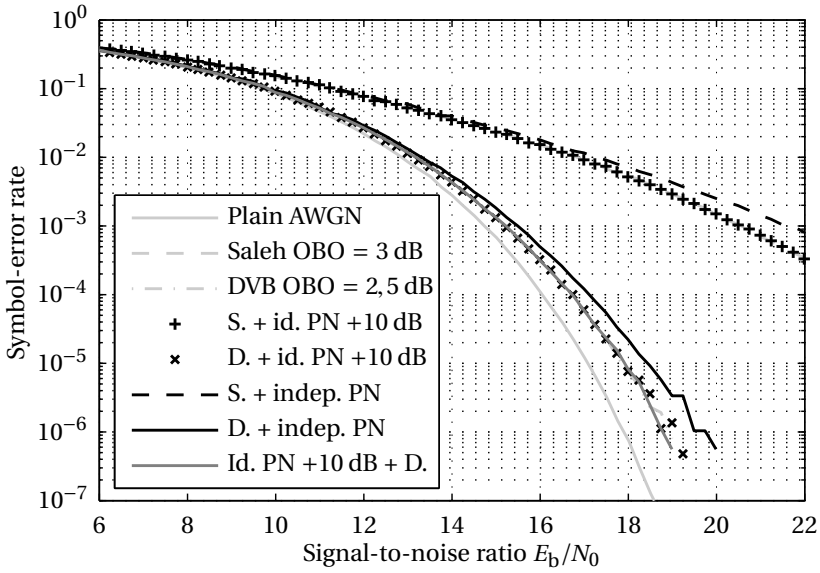


Figure 3.8.: Simulated 16-PolSK SER with non-linearity and phase noise

noise effects. Except for the curve in the lowermost legend position, phase noise is always applied after the non-linearity. The three curves from Figure 3.7, containing the results in the linear AWGN channel with non-linear amplification, are included in light grey for reference.

A couple of implications can be drawn from these results. When adding independent phase noise to an already distorted signal, the additional degradation is less important than from phase noise with respect to the plain AWGN channel.

The immunity towards phase noise remains intact upon non-linear amplification, irrespective of whether it is applied before or after the amplifier. As a consequence, it should be possible to compensate not only phase noise contributions in the receiver, but also the phase noise originating from the frequency conversion on the satellite transponder.

3.3.3. PolSK compared with classic modulation techniques

Polarization shift keying is envisaged as a possible alternative to other modulation techniques currently employed in satellite communications. Therefore a direct comparison with these alternatives imposes itself.

Considering the bandwidth and power use of dual-polarization signalling, PolSK modulation schemes have to be compared with classical I/Q modulation

techniques in dual operation (e.g. QPSK simultaneously on both polarizations). This ensures a fair comparison between single- and dual-polarization modulation techniques, as spectrum usage and transmission rates are equal.

The SER of the contemporary reference techniques is derived from known error probability formulas [80][4]. For a dually operated technique to lead to a correct decision, *both* underlying symbols need to be detected correctly. In consequence, the error probability $p(E)$ in dual operation has to be adjusted accordingly:

$$\begin{aligned} & p_{\text{dual}}(C) = p_{\text{single}}(C) \cdot p_{\text{single}}(C) \\ \Longleftrightarrow\ & 1 - p_{\text{dual}}(E) = \big(1 - p_{\text{single}}(E)\big) \cdot \big(1 - p_{\text{single}}(E)\big) \\ \Longleftrightarrow\ & p_{\text{dual}}(E) = 2p_{\text{single}}(E) - p^2_{\text{single}}(E) \end{aligned} \tag{3.37}$$

The event C stands for a correct decision, while E is an error.

Figure 3.9 shows the resulting plot comparing 16-PolSK with various other modulation techniques of equal spectral efficiency in the thermal noise channel. The reference techniques operate on a single or on both polarizations — some are based on absolute, others on differential phase detection.

It can be deduced that 16-PolSK is indeed more power-efficient than 16-PSK and 16-DPSK on a single polarization. On the other hand, PolSK cannot compete with the spectral efficiency of coherent PSK schemes in polarization division multiplex, as the curves for dual QPSK or even dual 8-PSK imposingly demonstrate.

Pfeiffer argues in [24] that a direct comparison with PSK can hardly be justified as PolSK is, in the sense that the phase *difference* between two carriers is modulated, a differential technique and should be compared with techniques of the same family. Comparing thus PolSK with dual differential QPSK, a little more than 4 dB of additional SNR is required by a PolSK signal to achieve the same symbol-error probability.

Knowing that almost half (i.e. 2,6 dB) of the transmission power is consumed in the residual carriers and that symbol decisions are taken in a particular *power-space*, where noise statistics differ from the Gaussian distribution, this efficiency gap seems legitimate, although slightly disappointing.

3.3.4. Preliminary conclusions

The simulations conducted, validated by findings in the literature, reveal that polarization shift keying suffers from a considerable power penalty with respect to modulation techniques currently used in satellite communications. This penalty can only be justified under circumstances where the main advantages of PolSK, the phase-noise immunity and reduced reception complexity, can be fully exploited.

[4] See the documentation of the `berawgn` function

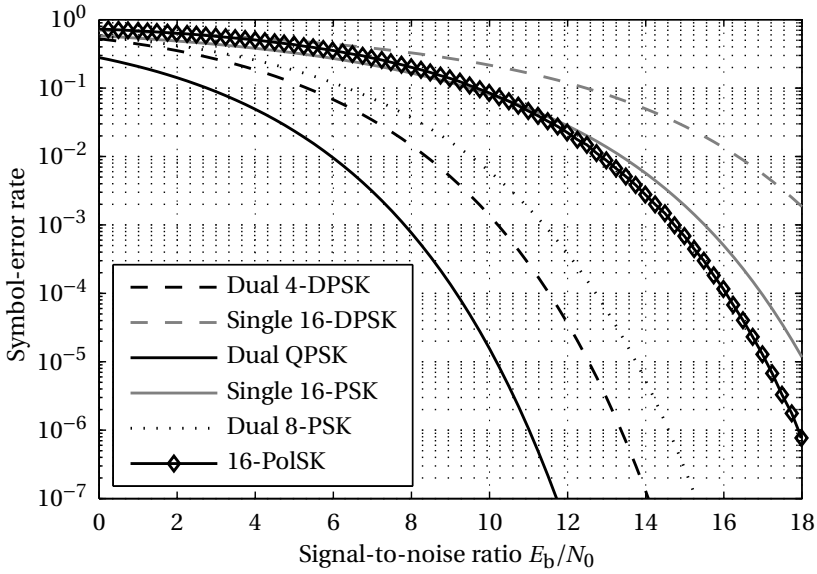


Figure 3.9.: Simulated 16-PolSK SER, compared with other popular modulation methods

One example for circumstances favourable to PolSK application is narrowband communications. Here, the impact of phase noise becomes increasingly dominant as the signalling bandwidth and phase-noise bandwidth are within the same order of magnitude. Also in systems relying on bursty transmissions, PolSK can be used to its advantage; the capability of direct-detection and the explicit symbol clock allow reception without prior lock[5]. Polarization de-rotation is only required once, as it is does not vary from one transmission to the next.

A second issue is that a PolSK signal is amplified by two separate amplifiers. A constant-power constellation like the proposed 16-PolSK still has considerable amplitude variations when looking at each polarization separately. These can be minimized by choosing appropriate constellations, but for their complete avoidance a further considerable loss in signalling efficiency would have to be accepted. Good reception quality can be achieved with linearized amplifiers in back-off. A more adventurous solution would be to look for an amplifier that could operate on both polarizations simultaneously (cf. Section 4.3.2 in the following chapter).

[5] In this specific case, the symbol clock has to be extracted quickly as well, possibly using a feedforward recovery technique like that presented in [61].

3.4. Measurements

The simulations have considered several constellations with different spectral efficiency. 16-point PolSK was selected for a more thorough investigation, examining it under realistic operating conditions and comparing it with existing bandpass modulation techniques. This section continues these investigations and describes the measurement results obtained using the implemented demonstrator.

The PolSK demonstrator modem is tested in various experimental assemblies. First measurements are conducted in thermal noise only. The carrier recovery system is tested and the sensitivity of the modulation towards path-length differences is measured. Then, the PolSK link is put to the test on a transponder simulator and over actual satellite transponders.

The measurements yield curves similar to those obtained by the simulation software: symbol-error rate as a function of the SNR. Decision errors are counted by the receiver itself, as described in Section 3.2.2. Signal and noise power are measured using a power sensor, except during the satellite demonstration.

Note that for this section the signal-to-noise ratio is indicated as average symbol energy *per polarization* $E_s/2$ over noise-power spectral density N_0. This corresponds to the physical SNR (cf. equation (3.7) above) and is more representative of how these quantities were measured. To maintain the link to the previous simulations, the SER for 16-PolSK as simulated in plain thermal noise is included in the displayed plots.

3.4.1. Intermediate frequency channel

A first measurement series is taken in an assembly on the 70 MHz intermediate frequency. The transmitter outputs are fed to the receiver without frequency conversion and over equal-length cables. In this case, frequency and phase coherency is maintained between both polarizations and the carrier recovery subsystem is disabled.

Thermal noise is coupled into the signal path to emulate noisy reception amplifiers. The HP noise generators (HP 3708A) used for this task allow the noise power to be adjusted in discrete steps to synthesize different SNR levels. A block diagram of the assembly is shown in Figure 3.10.

Thermal noise

A first measurement compares the error performance of the demonstrator with the corresponding simulation results. The only impairment is thermal noise. The results are plotted in Figure 3.11.

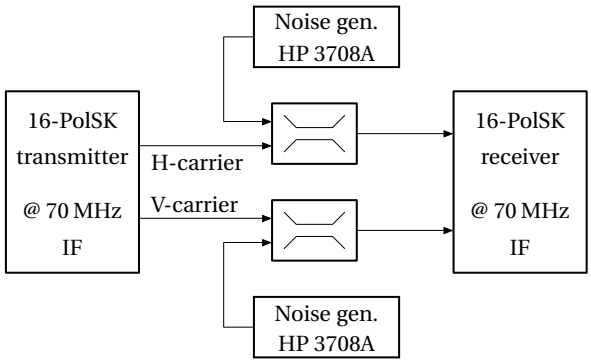


Figure 3.10.: Experimental assembly for the measurements in the intermediate frequency (IF) channel

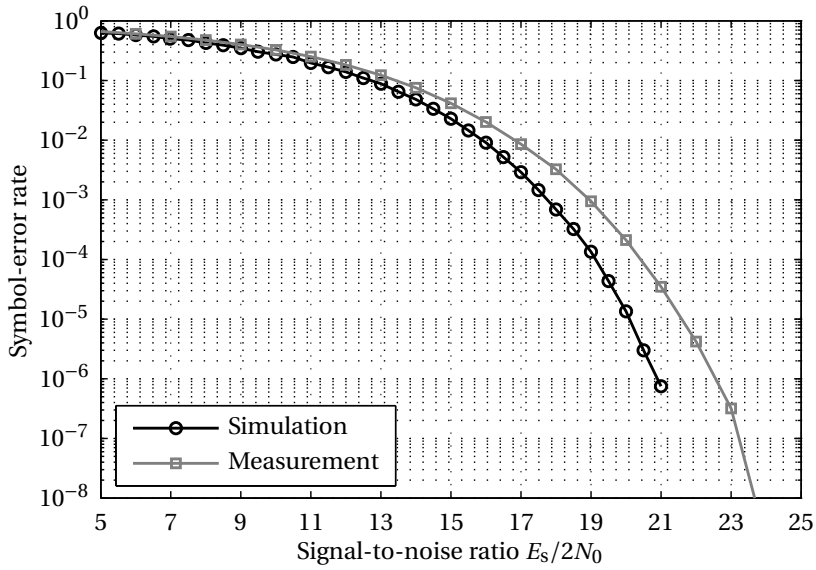


Figure 3.11.: Measured 16-PolSK SER in thermal noise

In high noise, the error rates of simulation and measurement are identical. As the noise is reduced, a discrepancy progressively shows between the curves. At the highest simulated error rate of $1 \cdot 10^{-6}$, the measurement lacks 1,6 dB behind the prediction. As the measurement is significantly faster than the simulation, it can be pursued towards lower error rates, even though the precision eventually suffers as well.

It can be concluded that the demonstrator correctly implements the PolSK modulation technique. The real system does however not quite achieve the performance predicted by the simulations.

I believe that this implementation loss is the combined consequence of two effects: First, the frequency response of the ADCs' analogue reconstruction filters did not correspond to the documentation. A digital filter, specifically implemented for compensating both the reconstruction filters as well as the DAC frequency response, had to be removed from the transmitter[6]. Some linear signal distortion is the consequence. Second, the simulations were carried out with perfectly synchronous timing (cf. Section 3.2.1). The receiver however selects one of 30 samples for downsampling to the symbol rate, leading to a timing imprecision distributed in the interval $\pm T/60$. Indeed, the aspect of the measured curve hints towards a noise-floor, similar to those shown in [57, pp. 14f.] for non-ideal timing.

Carrier recovery test

The first measurement was conducted with the carrier recovery disabled. To assess its functionality and accuracy, it is enabled for a second experiment. The estimated error rates with and without carrier recovery are displayed in Figure 3.12.

The plot shows that detection performance is almost equal for both cases. The fixed-carrier receiver has an almost indiscernible advantage in the high-noise region, but performs slightly worse when noise is low. Two conclusions can be drawn. First, the proportion of phase noise outside the bandwidth of the carrier recovery ($B_N = 10\,\text{kHz}$) is insignificant. Second, using PolSK with a carrier recovery does not significantly change the symbol-error rate, as long as this condition on phase noise is fulfilled. As a consequence, measurement results obtained with activated carrier recovery should at least qualitatively be applicable to *plain* PolSK reception, without carrier recovery.

A possible explanation for the subtle difference between both curves is the following. High noise power in the carrier-tracking PLL leads to phase jitter and slightly reduced tracking capability. This explains why using fixed-carriers is beneficial in this case. When noise levels are low, this advantage is lost and the fixed-

[6] This filter is not included in any of the descriptions, as it was never used in the end.

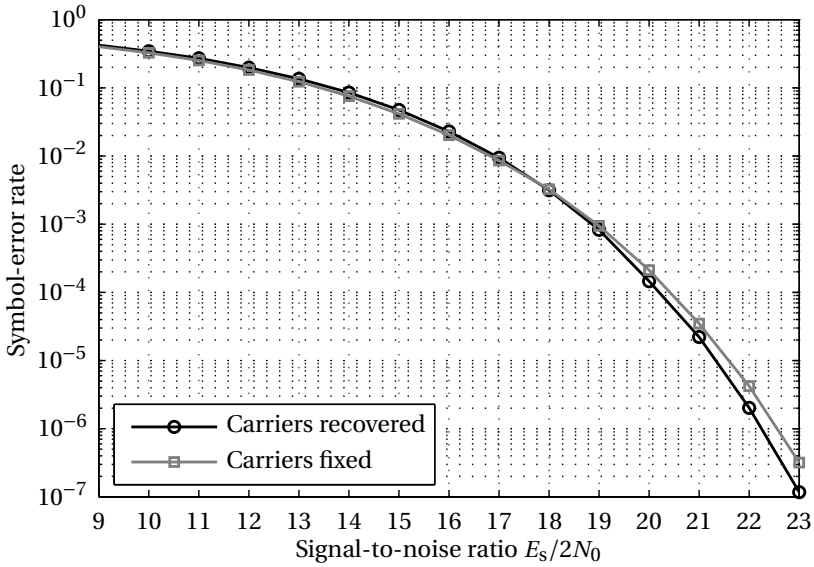


Figure 3.12.: Measured 16-PolSK SER in thermal noise, with active and disabled carrier recovery

carrier variant is a little less reliable because the incoming signal is, frequency-wise, not perfectly aligned to the matched filters in the receiver. In this case a small fraction of the signal energy is simply lost.

Unequal path length

The position of the symbol strobe becomes ambiguous when the path length between the horizontal and vertical polarizations differs by a substantial fraction of the symbol-*length*. The symbol duration of $T = 1/1{,}024\,\text{MHz} = 976{,}56\,\text{ns}$ corresponds to a length of 292 m in free space. Measurements are conducted with the path on the vertical polarization being 4,7 and 9,9 m longer than on the horizontal[7]. Assuming a phase velocity of $c/\sqrt{2{,}3} = 197{,}9 \cdot 10^6\,\frac{\text{m}}{\text{s}}$ for the RG-59 cable, the delay differences are:

$$\Delta t_1 = 23{,}75\,\text{ns} = 2{,}43\% \cdot T$$
$$\Delta t_2 = 50{,}03\,\text{ns} = 5{,}12\% \cdot T$$

[7] There is no specific reason for this choice of lengths besides that those were readily available cables.

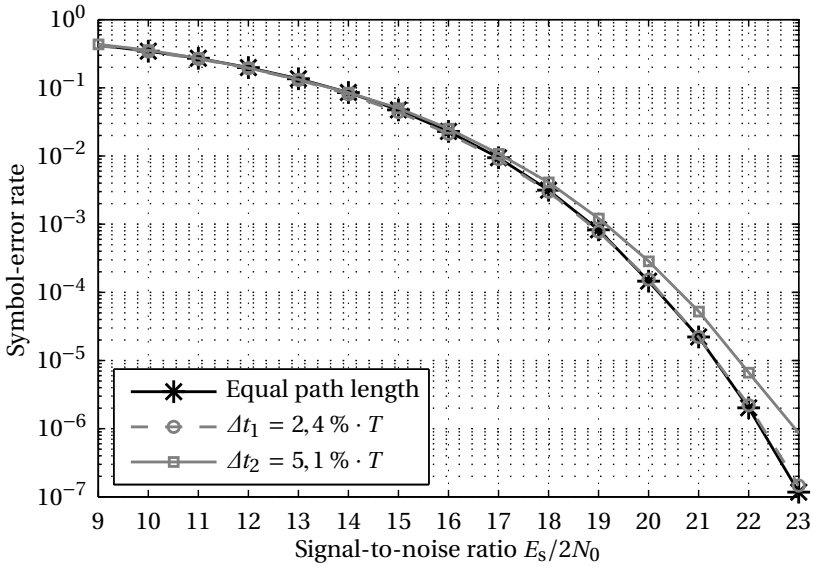


Figure 3.13.: Measured 16-PolSK SER with different path lengths between the polarizations

The receiver has the carrier recovery enabled, to compensate the phase shift resulting from the delay difference. Measurement results are displayed in Figure 3.13.

For the smaller delay, no perceptible deterioration in reception quality is observed with respect to the first measurement above. As the delay is increased however, the symbol-tracking loop cannot maintain the ideal sampling instant fixed on both signals at the same time: one of them is off by one sample. The consequence is the reduced noise robustness, as shown by the plot.

This finding can be extrapolated to a faster signal, occupying a full satellite transponder at 27 MBd. A tolerance in 5% translates into a maximum allowed free-space path difference of 55 cm, a difference which could quickly accumulate in ground stations or satellite payloads. Still, as long as the symbol frequency is equal on both signals, which can safely be assumed when the uplinks are made from the same ground station, an adequately designed receiver can measure and account for this delay difference (cf. Chapter 5).

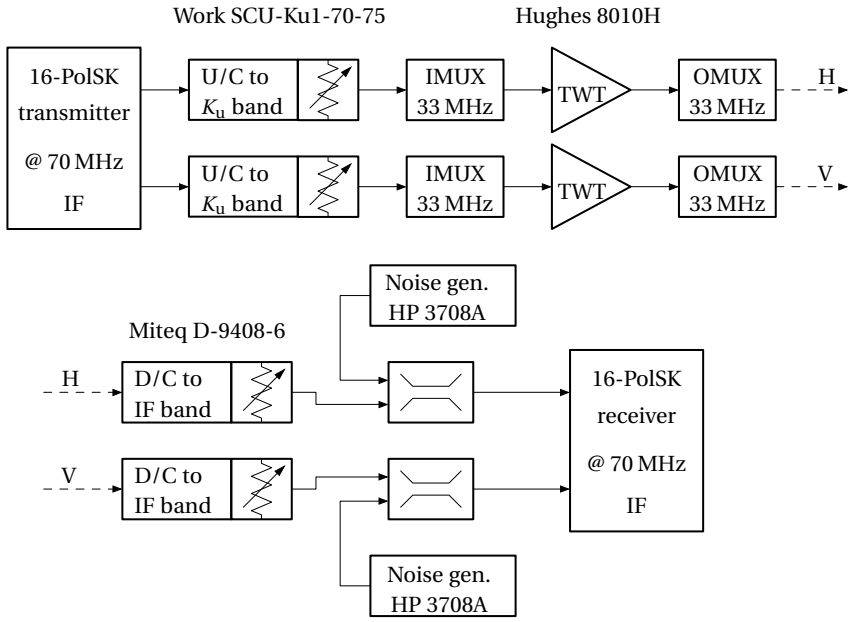


Figure 3.14.: Experimental assembly for SER measurements on the transponder simulator

3.4.2. Satellite transponder simulator

In a second measurement series a basic satellite transmission is simulated using a *transponder simulator*. The transponder simulator comprises input multiplex filters (IMUX), travelling wave tube amplifiers (TWT) and output multiplex filters (OMUX). These devices all operate in K_u-band so prior frequency conversion is required. A block diagram of the setup is displayed in Figure 3.14.

The transponder simulator allows a better appreciation of the PolSK modulation in realistic conditions: conversion to microwave frequencies, two filter stages and non-linear amplification by real tubes. The up- and downconverters are not locked to a common frequency reference, so that the final carrier frequencies will be slightly different. The converters are professional-grade units, therefore the introduced phase noise, even though not correlated between polarizations, can be considered negligible.

Consequently, the carrier recovery is enabled for all measurements with the transponder simulator. Different back-off settings are selected by modifying the amplifiers' operating point using variable attenuators in the upconverters.

In preparation for the experiment, the amplitude-to-amplitude modulation

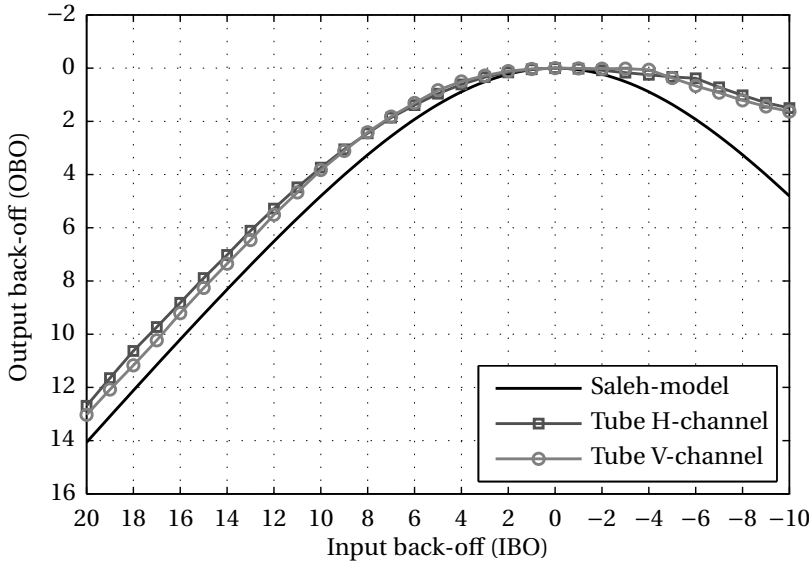


Figure 3.15.: Transponder simulator's TWTA characteristics, measured AM-AM curves and Saleh-model

characteristic of the amplifiers has been measured using clean-carrier signals. They are depicted in Figure 3.15, together with the *Saleh*-model [67] used by the simulator[8]. The small difference visible between the amplification characteristics of the H and V tubes are unlikely to have any significant effect on the results. Alongside the experimental SER measurements, simulations are run to provide a prediction of the transponder simulator's expected behaviour. As the AM-AM curves differ quantitatively from the Saleh-model, slight mismatches should be expected between measurement and simulation.

A first series of measurements is conducted with the amplifiers in a linear operating point at 11 dB output back-off. In a second series, the back-off is tentatively reduced to 3 dB. The results of both measurements are displayed in Figure 3.16. As a comparison, the previous measurement result in thermal noise as well as the pertinent simulation outcomes are included in the plot.

In the case of linear amplification, the modem performance is very close to that from the previous measurements on the intermediate frequency; there is no

[8]The phase-distortion of the tubes was not measured, therefore only AM-AM curves are shown. In addition, although the Saleh-model is generally considered accurate for amplifier AM-AM characteristics, this is not necessarily so for the description of amplitude-to-phase modulation, which can vary considerably among real tubes.

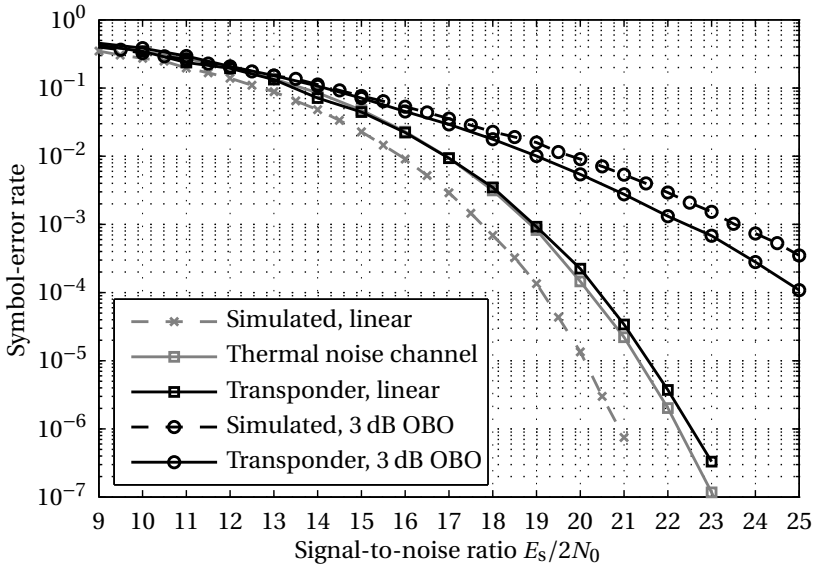


Figure 3.16.: Measured 16-PolSK on the transponder simulator in linear conditions and with OBO = 3 dB

significant difference up to an SNR of about 20 dB. From there on, a small degradation can be observed for the transponder measurement. It can be attributed to a combination of the non-identical amplification characteristics, converter phase noise and the analogue filters' group delay variations.

With a distorted signal, the receiver performance drops considerably. More than 4 dB of power efficiency is lost at an SER of $1 \cdot 10^{-4}$. This is a consequence of the symbol constellation being deformed by amplitude compression and phase shifts. In addition, matched filtering of the distorted waveform introduces strong inter-symbol interference (ISI). Interestingly, the real error rate is lower than predicted by the simulation, even though the simulation significantly outperforms the modem in the linear channel. This confirms the suspicion that the Saleh-model leads to somewhat pessimistic predictions.

Backing off by 3 dB is clearly not enough for faithful PolSK reception. A popular strategy is to amplify with *peaks-at-saturation*, where the output back-off corresponds to the baseband peak-to-average power ratio. The PAPR of pulse-shaped 16-PolSK, as determined by the simulation, is 5,5 dB. Accordingly, a third measurement on the transponder simulator has been conducted. The results are displayed in Figure 3.17.

The penalty from distortion is significantly reduced in this setting. At the

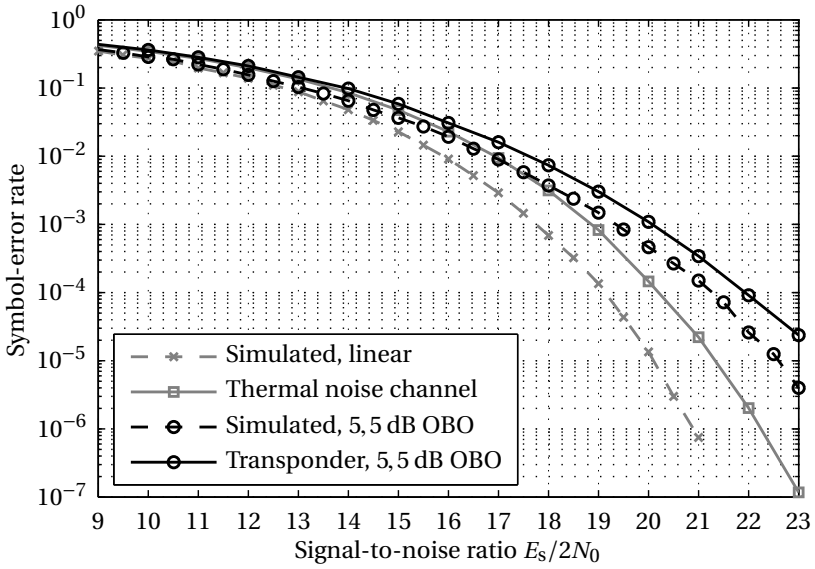


Figure 3.17.: Measured 16-PolSK on the transponder simulator with OBO = 5,5 dB

SER = $1 \cdot 10^{-4}$, 1,3 dB is lost with respect to the measurement on the intermediate frequency. Again, the corresponding simulation is qualitatively close but does not exactly predict the observation. 5,5 dB of back-off is considered *a lot* in satellite engineering. An additional correction of 5,5 + 1,3 = 6,8 dB has to be accounted for in a link budget, with respect to PolSK in the linear channel.

3.4.3. Satellite demonstration

The final feasibility proof for polarization shift keying was brought by the actual demonstration over satellite. This experiment was conducted on January 29th, 2014 in Betzdorf, Luxembourg, using transponders on the SES *Astra 3B* satellite. The weather was favourable with a clear, blue sky for the test.

Uplink and downlink were accomplished by the same 11 m-parabolic-dish antenna. The horizontal carrier was put on the antenna's horizontal feed, the vertical correspondingly. The uplink frequency was 14 269,25 MHz, with an estimated power of 40 dBW EIRP (equivalent isotropically radiated power) per polarization. The satellite transponders downconverted the signals to 11 469,25 MHz and retransmitted them after amplification, with tubes operating in the linear region. Regulatory restrictions did not allow the amplifiers to be operated anywhere

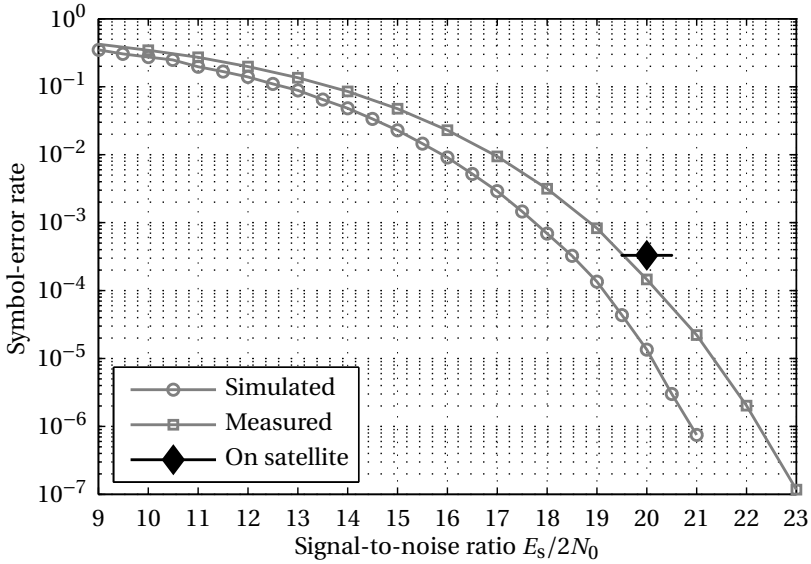


Figure 3.18.: Satellite demonstration of polarization shift keying

close to saturation, neither the transmission power to be modified. The polarizations are flipped on the satellite, such that the H-signal was coming back on vertical and V on horizontal. In consequence, the receiver ports had to be plugged inversely. Downlink EIRP from the satellite was a measured 33 dBW on average on both polarizations.

All on-ground frequency converters were synchronized, such that the uplink signal was a *plain* PolSK signal. After frequency conversion on the satellite, it was retransmitted with an unknown frequency difference between both carriers.

With enabled carrier recovery, the receiver locked onto the signal successfully and an error-rate measurement was initiated. Over the period of approximately 28 minutes, 1 763 556 488 symbols were transmitted, containing 578 271 errors. Accordingly the SER was $3{,}3 \cdot 10^{-4}$. The SNR, roughly estimated using a spectrum analyser, was between 19,5 and 20,5 dB. Due to the aforementioned restrictions, only a single measurement could be taken. The result is plotted as a diamond in Figure 3.18, including the SNR uncertainty.

Performance of the PolSK system over satellite is, at best, the same as in the IF channel and at worst 1 dB below that, at the measured error rate. More important than the exact error probability is the successful receiver lock and the continued stable connection. It shows that the devised recovery systems are able to successfully track a PolSK signal in its target setting. With this demonstration,

the technical feasibility of PolSK over satellite has been shown, confirming a major hypothesis of the project. This concludes the measurements on polarization shift keying.

3.5. Discussion

In this chapter polarization shift keying has been reviewed in the context of satellite communications. The PolSK modulation technique and constellation design have been adapted to cope with path-length differences and, to the extent that was possible, non-linear amplification. The proposed modifications also allow the reception of PolSK upon coherency loss between both polarized carrier signals without appreciable performance drop.

The techniques developed were successfully proven by simulations and by a technical demonstration. Initial performance predictions were obtained from a simulation program, which subsequently guided the further development of the demonstrator equipment. The demonstrator allows PolSK transmission and online reception including symbol-error counting. The investigations revealed that PolSK pays for its phase-noise immunity and direct-detection capability with a certain penalty in power efficiency. In this respect, PolSK cannot compete with traditional modulation techniques in the analysed broadcast channel conditions.

It was shown that the proposed PolSK modulation variant leads to three discrete spectral lines in the frequency spectrum, which reunite almost half of the signal's power. This led to the claim that recovery systems utilizing these discrete carriers to extract required clock signals would be particularly robust. The devised recovery systems worked perfectly well and fulfilled their duty flawlessly even in strong thermal noise and with different frequencies on each polarization. Except from such isolated hints however, this claim has not been verified and remains a hypothesis until it can be confirmed by a rigorous scientific examination.

The unconditional phase-noise immunity of PolSK is lost upon conversion using non-coherent local oscillators. Even though the proposed techniques still allow successful demodulation, an experiment in phase-noise-dominated channel conditions could bring further insight. One might even investigate whether an identical phase-noise process, applied to carriers on different frequencies, could upon conversion into baseband still show some degree of correlation. Due to the involvement of non-linear processes, the outcome is difficult to predict.

As a conclusion it can be said that PolSK can be used in satellite communications, when two polarizations are available, as for example in traditional broadcasting. The use of PolSK should be limited to cases where the phase-noise immunity and the reduced reception complexity justify the disadvantage in power efficiency. As examples, narrow-band communications and bursty transmission

schemes have been mentioned. Further research might reveal other scenarios.

In the PolSK variant presented here, the *fixed* carrier was used to transmit phase pilots, thereby allowing the recovery of the transmitted state of polarization. Another approach would be to use this *idling* carrier for additional signalling. Instead of transmitting a zero phase, an additional phase modulation could be impressed onto this carrier, like a BPSK or QPSK signal. This would increase the spectral efficiency without additional power cost as the power in the discrete spectral components would automatically be redirected to this additional modulation. Direct-detection capability and phase-noise immunity should be conserved, but the compensation of the path length difference would have to be accomplished in a different way.

Puncturing the PolSK signal with an additional phase modulation is certainly worth investigating. Some tentative research revealed that this leads to a combination of two modulation techniques with entirely different properties, requirements on SNR and probably different target applications. Both modulation techniques would have to be carefully matched to each other and analysed in a joint manner.

The approach chosen for the further work is a different one. Having a three-dimensional PolSK signal plus an additional one-dimensional phase modulation leads to a combined four-dimensional signalling space. This four-dimensional signal space can be considered as one, having two amplitudes and two phases for modulation. This approach, called 4-Quadrature or four-dimensional signalling is analysed in the following Chapter 4.

4. Four-dimensional signalling

The feasibility of using polarization shift keying in a satellite link was successfully demonstrated. The investigations on PolSK led to two important observations. Firstly, polarization shift keying suffers from a considerable power penalty with respect to modulation techniques currently employed in satellite communications. The second is that when non-coherent frequency conversion is used on the satellite transponders, the advantages of PolSK practically vanish.

For this and other cases where PolSK cannot be justified, another candidate for the conjoint exploitation of both polarizations is proposed. Four-dimensional, or *4-quadrature signalling* [8]. As for PolSK, the dimensionality increase allows more freedom in constellation design than classic in-phase/quadrature modulation. But here, all four available dimensions in the dual-polarization channel are exploited, such that no loss in power efficiency should be expected.

The constellation design principles introduced in Section 2.4 are employed to generate constellations for the four-dimensional (4-D) signal. Their power efficiency is simulated and compared with that of two-dimensional constellations of comparable spectral efficiency. As these four-dimensional constellations are more aggressively oriented towards power efficiency, the question of their peak-to-average power ratio is addressed in a more rigorous fashion. Potential gross signalling gains might be put into perspective by their limited applicability in non-linear channels.

In a similar procedure to that in Chapter 3, one constellation candidate is selected for implementation. The demonstrator is created by modifying the previously built PolSK modem and is used to prove the feasibility on the ground and in space of the four-dimensional modulation technique.

A related signalling method was recently published in [52, 85], showing a different way of using both polarizations to increase bandwidth efficiency. The technique is called poly-polarization signalling and consists of superimposing a third carrier signal diagonally on top of existing horizontal and vertical carriers. Due to the unmistakable similarities between poly-polarization signalling and the four-dimensional technique presented here, it is analysed as well in Section 4.3.3.

4.1. Constellation design

In Section 2.4.2 the concepts of coding and shaping gain were introduced to guide the constellation design of four-dimensional symbol constellations. According to these ideas, different constellations are proposed hereunder.

4.1.1. Lattice-based constellations

Targeting maximally power-efficient signalling, the first set of constellations is based on the D_4-lattice. To this end, a large sample of D_4-lattice is generated and positioned in space such as to free the origin from a node. The constellation is then carved out by delimiting the lattice by a spherical boundary centred around the origin [54]. In this way, a maximally dense and shaped constellation is created with minimum average power. The lattice itself provides for the *coding gain*, the spherical cut-out for the *shaping gain*.

The so-created constellations will be referred to as *lattice amplitude modulation* (LAM) constellations in this thesis work. LAM constellations can be viewed as four-dimensional correspondents to classic, rectangular QAM, or to the cross-shaped constellations from [32].

An 88-LAM and a 256-LAM constellation have been generated[1]. The latter corresponds to a dual 16-APSK or 16-QAM constellation in spectral efficiency. The former is comparable to 8-PSK or to 8-QAM.

These LAM constellations are similar to regular QAM in the sense that they are lattice-based instead of being arranged along one or several discrete power levels like for (A)PSK. By construction they will show a relatively high peak-to-average power ratio and are therefore of limited applicability in satellite communications. Even though they could be used, preferably in multi-carrier operation where amplifier back-off is customary, the main reason for their inclusion into the study is that they can serve as a good reference for what can effectively be gained by the dimensional increase under idealized channel conditions. By subsequently restricting the freedom in constellation design according to the physical limitations, more appropriate designs are found below.

The same argument justifies that the LAM constellations should be compared to QAM, rather than PSK constellations of equivalent spectral efficiency. For the dual 8-point constellation, a *hexagonal* QAM constellation is preferred over the rectangular arrangement. The hexagonal grid is the most dense in two dimensions [19] and an 8-point constellation is readily created. Such a comparison

[1] The generation procedure does not necessarily create constellations with a number of points being a power of 2. It was not possible to elegantly arrive at 64-points, so the choice falls on the next larger constellation with 88 lattice nodes.

can reveal the actual net gain from passing into four-dimensions. The hexagonal 8-QAM constellation is shown in Appendix C.

To illustrate the application of the coding and shaping techniques, the orthogonal projections onto each polarization[2] of the 256-LAM constellation are displayed in Figure 4.1. The points represent *symbol-elements*, i.e. the projections of 4-D symbols onto either polarization. The numbers indicate how many symbols will resolve to a given symbol-element in either projection. They represent a frequency that a certain symbol-element will be selected when all symbols are equiprobable. Note that different orthogonal projections are possible. The present one is chosen such that the number of symbol-elements is minimized.

The displayed constellation looks similar to two QAM constellations. The choice of a dense lattice leads to a higher number of apparent symbols than in plain rectangular QAM. Where 16-QAM has four rows of four symbols, up to 6 symbol-elements are arranged in a line or column for 256-LAM. This effect is called *constellation expansion* [32].

Looking at the relative frequency of the symbol-elements in Figure 4.1, it is observed that low energy points are selected more often than those on the projections' outer boundaries. This is the consequence of shaping; the displayed constellation approaches a two-dimensional, Gaussian probability density function.

It has to be mentioned that both generated LAM constellations correspond exactly to the *four-dimensional codes* shown in [86] for the given number of symbols. These constellations are thus known, but have so far not been considered as a replacement for the polarization division multiplex in satellite communications.

4.1.2. Sphere-based constellations

Phase shift keying (PSK) or amplitude and phase shift keying (APSK) are used for transmissions with amplifiers at or close to saturation. Constellations like QPSK and 8-PSK have all constellation points on an equal-power circle in the I/Q-plane. In this way, envelope variations are kept to a minimum and distortion, small.

To construct similar constellations in four-dimensions, one can place a fixed number of points equidistantly onto the three-dimensional surface of a sphere in 4-D space; then all the symbols have equal energy. These are the constellations initially proposed in [8].

As a specific example, a 64-point constellation was generated, denoted 64-4D-PSK, to be compared to dual 8-PSK. The exact arrangement of points is taken from Sloane's website on spherical packings [76]. A constellation diagram is not

[2] These are called *constituent 2-D constellations* in the coding literature.

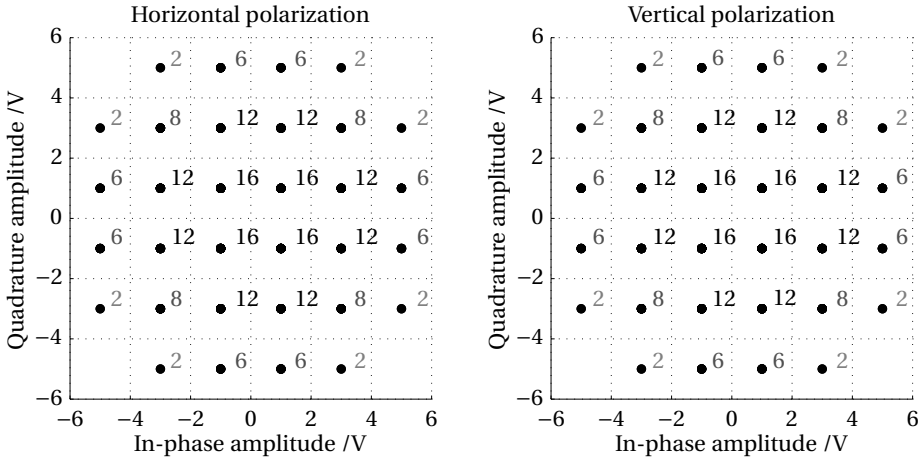


Figure 4.1.: 256-LAM constellation projected onto both I/Q planes. Numbers indicate the symbol-elements' frequency if every symbol in the constellation is selected once.

shown; being the result of a numerical optimization procedure, it has a chaotic aspect that is ultimately not very enlightening.

The constant-power limitation on 64-4D-PSK is only meaningful when both polarizations are considered together. As satellite amplifiers have to amplify both polarizations separately, the constellation suffers from a similar penalty to the PolSK constellation from the previous chapter. Thus, even for 64-4D-PSK, amplifiers would require some back-off.

4.1.3. Cylinder-based constellations

If signal envelope variations are to be completely avoided also on individual polarizations, the constraints on constellation design have to be further restricted. By fixing the power on both carriers, only two degrees of freedom remain: the phases. The same principle of dense lattices, used above in the *quadrature amplitude space*, can be applied inside the two-dimensional *phase space*. As in two dimensions the hexagonal grid is the most dense, a constellation can be generated by laying out points accordingly. The constellation layout can be imagined as intersecting cylinders in 4-D space, hence the classification as *cylinder-based*.

The phase space is of quadratic size, $[0;2\pi[\times[0;2\pi[$ and a closed surface[3]; whereas the hexagonal grid has a non-rational side-ratio: A quadratic chunk can

[3]The boundaries are *glued* together.

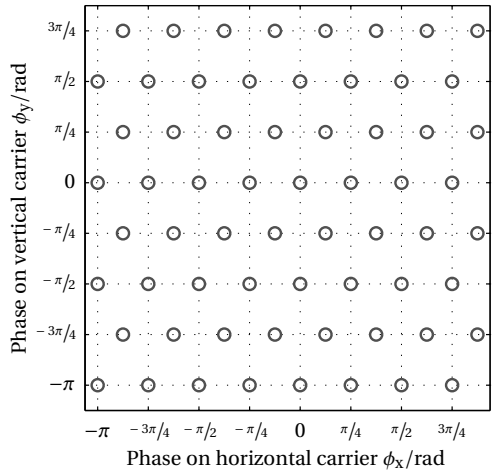


Figure 4.2.: Phase diagram for 64-point cylindrical PSK

only be created approximately, for a very high number of points. If 64 points are used, the lattice needs to be slightly squeezed in order to *fit*; this deformation costs power efficiency. A constellation plot in the two-dimensional *phase-space* is shown in Figure 4.2.

Taking APSK as a model, one can envisage similar constellations with multiple cylindrical shells. The construction of such constellations seems to be more involved and probably best undertaken with the help of numerical optimization procedures. An algorithm like that proposed in [16] should be extendible to four-dimensional signals. An extensive search for APSK-like four-dimensional constellations would have gone beyond the scope of this work.

4.1.4. Bi-orthogonal constellation

A particular cylinder-based constellation is the four-dimensional, bi-orthogonal signal. It was identified in [51, 2] as the most power-efficient modulation in four dimensions.

As stated in Section 2.4.4, a carefully chosen projection yields constellation diagrams similar to QPSK. A representation in phase space becomes possible as well and is shown in the scatter plot in Figure 4.3. A comparison with dual QPSK shows that the bi-orthogonal constellation contains half of its symbols, thereby separating the QPSK constellation into two distinct sets. In this way, 1 bps of spectral efficiency is traded in to increase the mutual distance by $\sqrt{2} = 1{,}51$ dB.

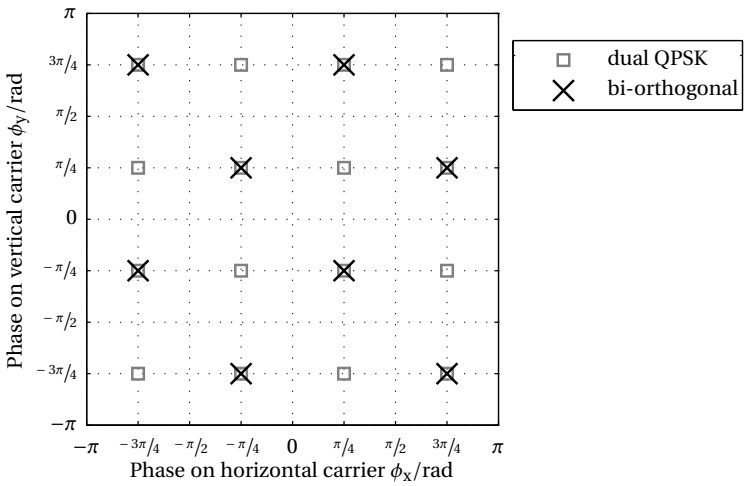


Figure 4.3.: Phase diagram comparing dual QPSK and bi-orthogonal constellations

Due to their structural similarity the bi-orthogonal constellation can be compared to QPSK in dual operation. This is, however, only partly justified as QPSK's spectral efficiency is considerably higher. As an alternative, one can consider a comparison with a 3-point PSK constellation. Such a 3-PSK constellation, laid out as vertices of an equilateral triangle, permits a better illustration of the net gain from the dimensional increase. The spectral efficiency of 3,17 bps for dual 3-PSK comes quite close to the 3 bps of a bi-orthogonal signal. The 3-PSK constellation is shown in Appendix C.

4.2. Tools and methods

For each of the symbol constellations introduced in the previous section, a two-dimensional counterpart was identified. For both the proposed constellations and their respective 2-D references, the power efficiency is simulated and compared using the same `Matlab` simulation tool as used in the analysis of polarization shift keying. For this, the new four-dimensional constellations have been introduced as an extension module into the software. As an additional validation, union bounds are used to confirm the simulation results for the lattice-based constellations.

As for PolSK, a demonstrator was implemented. It is based on the same plat-

form and uses the bi-orthogonal constellation for modulation. Most of the previously accomplished developments could be reused for its implementation. Only PolSK-specific functions had to be replaced; the general structure of the modems and the hardware controlling features are identical.

Union bounds

Due to their regular structure, signal constellations based on the D_4-lattice are particularly suited to analysis using *Union Bounds* (UB). Union bounds are used to approximate the probability of a union of events by the sum of the events' probabilities [49, p. 318]:

$$p\left(\bigcup_i A_i\right) \leq \sum_i p\left(A_i\right) \tag{4.1}$$

The union bound can be used as an upper bound for the symbol-error probability in a digital communication system. It approximates the probability of a symbol error by summing over the probabilities that a received symbol is closer to every erroneous candidate in the constellation individually. This is easy to calculate as it involves only the one-dimensional integration of the noise probability-density function but yields a pessimistic estimate, as the numerous intersections between the *error regions* are not considered. The final union bound is obtained by averaging over the estimates for every symbol in the constellation.

The error in the union bound can be reduced by considering only the nearest neighbours of each constellation point. Conway [19, p. 71] suggests the following enhanced formula for computing the union bound for lattice-based constellations:

$$\mathrm{UB} = \frac{\hat{\tau}}{2}\,\mathrm{erfc}\left(\frac{d_{\min}}{2\sigma}\right) \tag{4.2}$$

$d_{\min}$ is the minimum distance between two symbols, while $\hat{\tau}$ is the average number of nearest neighbours for the constellation symbols. Both are functions of the given constellation. $\mathrm{erfc}\,(\cdot)$ is the well known complementary error function:

$$\mathrm{erfc}\,(x) = \frac{2}{\sqrt{\pi}} \int_x^{+\infty} e^{-t^2}\,dt \tag{4.3}$$

The average energy per bit is proportional to the squared minimum distance. Using this, combined with the relationship between noise standard deviation and spectral density,

$$2E_\mathrm{b} = \alpha\frac{d_{\min}^2}{b} \quad \text{and} \tag{4.4}$$

$$2N_0 = \sigma^2 \tag{4.5}$$

equation (4.2) can be brought into a form comprising the signal-to-noise ratio:

$$\mathrm{UB} = \frac{\hat{\tau}}{2}\,\mathrm{erfc}\left(\sqrt{\frac{b}{4\alpha}\frac{E_\mathrm{b}}{N_0}}\right) \tag{4.6}$$

b is the number of bits per symbol, like in Equation (3.9). $\hat{\tau}$ and the proportionality coefficient α have to be computed separately for each constellation. The union bound also points to a factor that has been neglected during constellation design. In addition to the distance between symbols, the number of nearest neighbours also plays a role [63]. This is what limits the effectiveness of coding algorithms when increasingly dense lattices are used (cf. [30] and Section 2.4.2).

4.2.1. Demonstrator

The bi-orthogonal signal constellation is used for the demonstrator implementation. Compared to the other proposed constellations, the complexity is easier to manage and it should be suitable for operation on saturated amplifiers. In addition, the QPSK-like constellation diagram permits the use of QPSK carrier-recovery systems.

Both the transmitter and the receiver build upon the PolSK modem. The same development platform is used and as a result the technical details are very similar. The signal is modulated onto a 70 MHz IF carrier at a symbol rate of 1,024 MBd. The pulse roll-off is reduced to the more progressive value of $\alpha = 0{,}20$.

In the transmitter, the symbol table and transmission filters are exchanged accordingly and the AGC thresholds are adapted.

More thorough modifications were necessary to develop the corresponding receiver. The symbol clock is recovered in the same way as for the PolSK modem by simply squaring the baseband signal [36]. Also here, the loop bandwidth is 100 Hz combined with an additional single-pol filter with $f_{3\mathrm{dB}} = 10\,\mathrm{kHz}$.

For the carrier recovery the method by Sari and Moridi [68] for QPSK, described in Section 2.5.1, is employed on each polarization individually. The so-detected error signals are fed to two type 2 PLLs with a loop bandwidth of 1 kHz. Each PLL controls an NCO and provides a local oscillator for downconversion.

The system works without pilot-symbols. As such, an initial phase ambiguity cannot be avoided in the carrier recovery. This is solved by iterating through all sixteen possible phase states until the error-counting LFSR is able to properly synchronize. For measurements in low SNR, the ambiguity needs to be resolved beforehand, while the SER is roughly below $1 \cdot 10^{-2}$.

4.3. Simulation results

Simulations are conducted to compare the four-dimensional constellations with the proposed references. For the results presented hereunder, the linear AWGN channel model is supposed. Further investigations on phase noise or non-linear distortion have been left aside.

The power efficiency of the different constellations is estimated by creating curves of the estimated error probabilities for different signal-to-noise ratios. In addition to evaluating a curve's general aspect, direct comparisons between two modulation techniques are made by looking at the signal-to-noise ratio required to achieve an error rate of $1 \cdot 10^{-4}$.

As before, symbol-error rates are compared rather than bit-error rates so that the tedious search for a good symbol-to-bits allocation is avoided. This time however, more precaution is advised when interpreting the results: Many of the four-dimensional constellations have more neighbours than their two-dimensional counterparts. It is often not possible to define the allocations such that neighbouring symbols will differ only in 1 bit. In consequence, a symbol error in four dimensions will more likely lead to multiple wrong bits than in two dimensions.

4.3.1. Power efficiency of new constellations

A first simulation run compares constellations with (approximately) 6 bit per symbol: Figure 4.4 shows the estimated power efficiency of 88-LAM and 64-4D-PSK. Dual 8-PSK and dual hexagonal 8-QAM are included as the corresponding reference modulations.

A first observation is that both four-dimensional constellations have almost the same error probability. The SER of 88-LAM is confirmed by the union bound. For a SER of $1 \cdot 10^{-4}$, their advantage over the classic constellations hexagonal 8-QAM and 8-PSK is approximately 0,7 dB and 2,2 dB, respectively. The advantage with respect to 8-PSK is significant, but the comparison with a constant-power modulation is rather unfair. Comparison with the result for the hexagonal constellation shows that indeed a moderate net power benefit is realizable when going into four dimensions.

It is interesting that the design limitation in the step from the lattice-based to the sphere-based constellation has only a small impact on the power efficiency. Even considering the fact that the 88-point constellation carries almost half a bit more, it leads to the conclusion that a combined power limitation for both polarizations is notably less severe than the traditional limitation to constant power on each polarization individually.

A similar simulation is conducted with the 256-point LAM constellation. The results are shown in Figure 4.5 and validated by a union bound approximation.

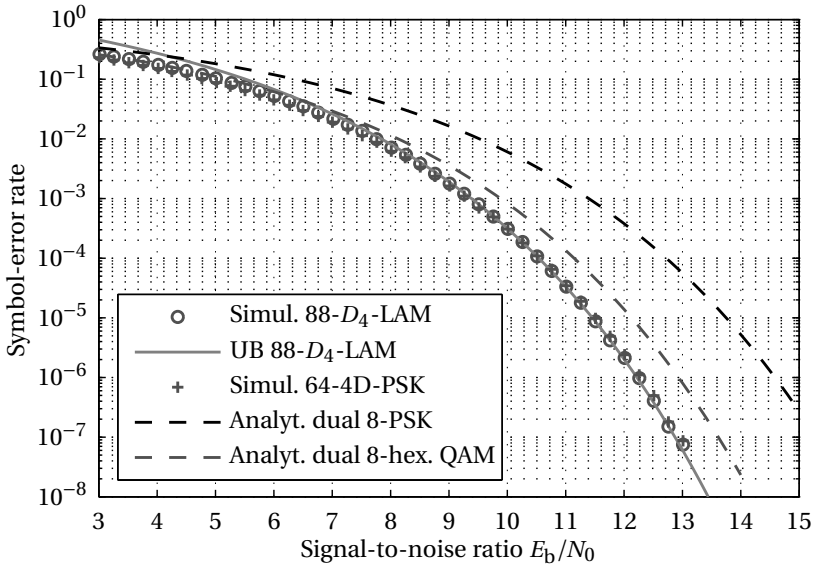


Figure 4.4.: Simulated SER of modulation techniques carrying approximately six bits per symbol

They are compared with the performance of dual 16-QAM. The power advantage for the LAM constellation is even higher in this case: approximately 1,3 dB can be gained.

In a third simulation, the PSK-like cylindrical constellation with hexagonally arranged phase values is compared to plain 8-PSK. This time the power constraints on constellation design are exactly the same as for the PSK signal. The results, plotted in Figure 4.6, show an improvement for the 4-D candidate, although of only 0,2 dB. It is interesting that this advantage can be obtained by simply rearranging the phase values of a dual 8-PSK signal. This is a case where the neighbourhood of a symbol point could become important. In the phase arrangement shown in Figure 4.2, a symbol has two neighbours at $d_{\min}$ and four at $d_{\min}\sqrt{5}/2$. In a comparable rectangular arrangement, there would be four neighbours at $d_{\min}$ and four at $\sqrt{2}d_{\min}$. Whether this subtle difference will neutralize the gain again when bit-error rates are considered has not been investigated further.

The last simulated constellation is the bi-orthogonal. The symbol-error rate predictions are plotted in Figure 4.7, along with those of dual QPSK and dual 3-PSK. The bi-orthogonal constellation shows very good performance with respect to both traditional methods. For the somewhat unjust comparison with QPSK, a gain of 1,6 dB can be achieved at an SER $= 1 \cdot 10^{-4}$. With 3-PSK, still 0,9 dB is pos-

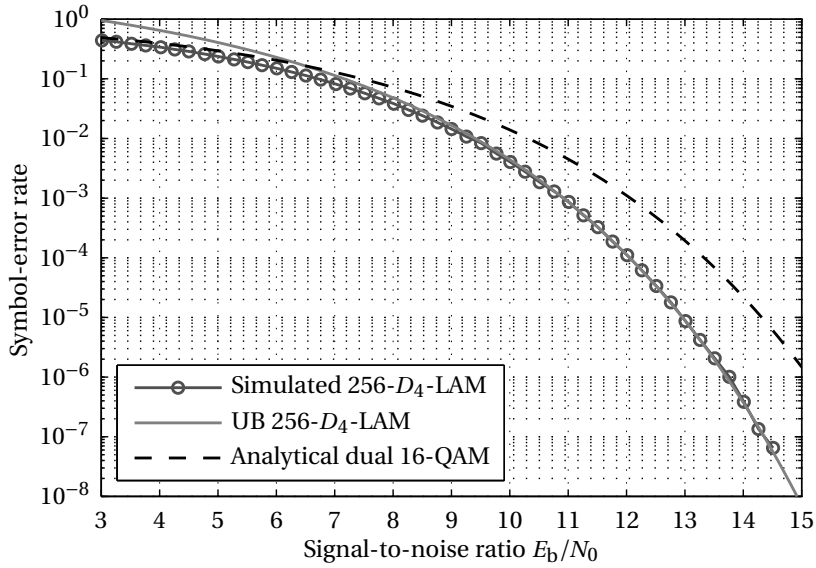


Figure 4.5.: Simulated SER of the 256-point LAM constellation and dual 16-QAM

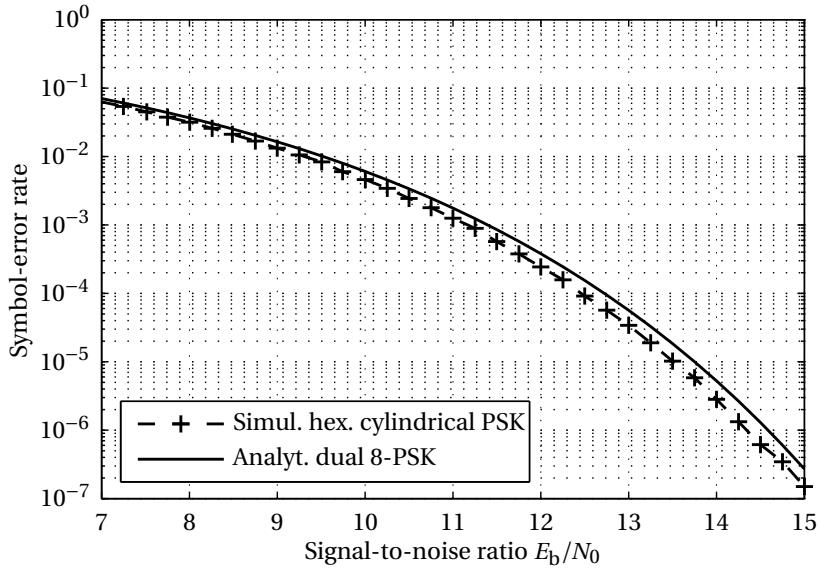


Figure 4.6.: Simulated SER of cylindrical PSK with hexagonal phase arrangement and dual 8-PSK

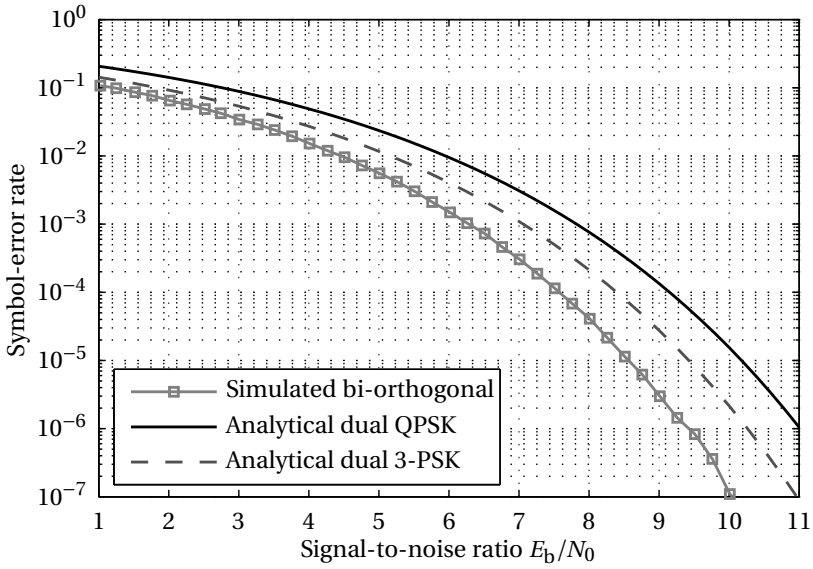


Figure 4.7.: Simulated SER of the bi-orthogonal constellation, dual 3-PSK and dual QPSK

sible. Note that the number of nearest neighbours is equal for all three candidates. A further deterioration when going to the bit-level is not expected in this case.

The simulations in the linear AWGN channel confirm the potential of four-dimensional constellations to yield a power advantage with respect to more conservative methods. Even when pushing the two-dimensional channel to its limit by applying more *exotic* constellations like hexagonal 8-QAM or 3-PSK, this advantage remains, even though it is less significant. For the very large and very small constellations (256 and 8 points), the achieved gain is considerable, while in the intermediate range of 3 bit per symbol and polarization, it is smaller.

A question only briefly addressed so far is the peak-to-average power ratio of these dual-polarization constellations. The non-linear amplifier behaviour requires backed-off operation when signals with high envelope variations are to be transmitted. The back-off, due to a reduced output power, as well as the remaining distortion, both reduce the potential net signalling gain.

4.3.2. Peak-to-average power ratio in 4-D signals

Constellation shaping increases the frequency of low-energy symbol-elements. Also the choice of the lattice structure, by expanding the constellation, increases

Table 4.1.: Table of peak-to-average power ratios for all proposed modulation techniques

Modulation		**Symbols**		**RRCS** $\alpha = 0{,}20$	
		comb.	single	comb.	single
	256-D_4-LAM	1,3	2,5	3,6	6,6
dual	16-QAM	1,8	1,8	4,3	5,6
	88-D_4-LAM	1,3	2,3	3,7	6,4
	64-4D-PSK	1	1,9	3,0	5,5
dual	8-hex QAM	1,6	1,6	4,2	5,3
	64-PolyPM	1,8	1,8	5,0	5,6
	hex. cyl. 64-PSK	1	1	3,0	3,4
dual	8-PSK	1	1	3,0	3,5
	bi-orthogonal	1	1	3,2	3,5
	bi-orthogonal alt.	1	1	2,5	3,5
dual	QPSK	1	1	3,0	3,5
dual	3-PSK	1	1	2,8	2,9

the per-polarization peak-to-average power ratio (PAPR) of the carriers.

Especially when each transponder is loaded with a single carrier, the amplifiers operate in or close to saturation for maximum output power[4]. This distorts signals with high envelope variations; distortion which can be reduced by moving the amplifiers' operating point away from saturation, i.e. backing off, by using various pre-distortion techniques or linearized amplifiers. The price is reduced output power and the link might not close. The SER to SNR curves shown in the previous section are oblivious of this, as E_b designates only the *average* power. The issue is less critical when multiple carriers are transmitted over the same transponder. Back-off is then mandatory by default and modulation types with higher PAPR may be used.

A good heuristic to determine the amount of back-off required to maintain signal integrity is to look at the PAPR of the baseband signal. The PAPR of all proposed constellations was estimated from the waveforms generated during the simulations; they are shown in Table 4.1. The table is arranged in four columns: the two leftmost show the PAPR values computed from the symbols in the constellation. The two rightmost columns show the PAPR for pulse-shaped signals with roll-off $\alpha = 0{,}20$, as yielded by the simulations. Each of those two groups has a column for the *combined* i.e. dual-polarization and the *single*, per-carrier

[4] This is where the amplifiers show the highest power efficiency, which is important as the limited power available on a satellite should be optimally used.

PAPR. The lines are ordered with decreasing spectral efficiency and also grouped accordingly.

Focussing on the pulse-shaped single-polarization column, the table allows the expected amount of back-off, required for the proposed constellations, to be compared. It shows that the PAPR of the large 4-D constellations (88-D_4-LAM, 256-D_4-LAM and 64-4D-PSK) is larger than that of the traditional constellations with which they are compared. By taking the PAPR *ratio*, the required additional back-off can be roughly estimated. As an example, using 256-LAM with an output back-off increased by 6,6/5,6 = 0,7 dB with respect to 16-QAM results in a reduction of the power gain from 1,3 dB down to 0,6 dB.

Constant-amplitude constellations like QPSK or 8-PSK usually require no or only very little back-off in practical operation, despite the PAPR of around 3,5. The same should be expected for the bi-orthogonal and the cylindrical 64-PSK constellation. Note that the PAPR of the 3-PSK is especially low, probably because symbol transitions do not pass through the origin.

The bi-orthogonal constellation was created by dividing the QPSK constellation into two complementary sets, of which one is removed (cf. Section 4.1.4 illustrated by Figure 4.3). A variant of the bi-orthogonal signalling method can be imagined where the signal is created by flipping this set after every symbol. A first symbol would be mapped according to the crosses in Figure 4.3, the following by the rectangles, the third by crosses again and so forth. This variant was simulated and is tabulated under *bi-orthogonal alt.*. The combined PAPR of this waveform is the lowest of all simulated constellations. It seems that the modification reduces the number of symbol transitions through the origin in 4-D space.

The observation of low combined PAPR is valid for all of the proposed 4-D constellations: it is in all cases significantly lower than for the individual carriers. When looking at the pulse-shaped signals on the right side, this is even true for some of the classic constellations. If there was an amplifier able to operate on both polarizations, at the same carrier frequency simultaneously, the power constraint could be alleviated from

$$E_{\mathrm{s,x}} < E_{\max} \tag{4.7}$$

$$E_{\mathrm{s,y}} < E_{\max} \tag{4.8}$$

to

$$E_{\mathrm{s,x}} + E_{\mathrm{s,y}} < 2E_{\max} \tag{4.9}$$

This would allow the use of 4-D constellations without the strong penalty due to back-off. The 64-4D-PSK constellation (and also the PolSK constellation from Chapter 3) was designed with such a constraint in the design prescriptions and has a combined symbol PAPR of one. If such an amplifier could be manufactured

and mounted on satellites, it would allow more efficient signalling in microwave broadband and satellite communications.

Neither my colleagues nor myself are aware of this kind of amplifier nor do we know whether a practical operation principle exists. Possible tracks could be the multi-port amplifiers that have become increasingly popular in satellite communications [96]. They allow for multiple frequency-multiplexed signals to be amplified using a combined *pool* of amplifiers, all in the same operation state. Multi-port amplifiers are, by default, also operated in back-off and in the present case the amplification of two carriers at exactly the same frequency is required. During discussions, gyrotrons were mentioned [53], but without practical hints on how to apply the principle to the present purpose. The use of semiconductor-based amplifier systems might open up new possibilities.

4.3.3. Poly-polarization modulation

In a recent publication another way of exploiting both polarizations is presented [52]. Starting from a classic polarization division multiplex signal, a third signal is added by *diagonally* superimposing it onto the original two. The interference from this non-orthogonal superimposition can be removed by an adequately designed receiver and all streams can be properly decoded. Indeed the paper shows how the power efficiency is increased by this method. The technique is called *Poly-Polarization Modulation* or PolyPM.

Constellation diagrams of the symbol-elements appearing as a result of this superimposition are shown in [99]. A superimposition of three QPSK signals leads to constellation diagrams similar to 16-QAM on both polarizations. It seems that the superimposition creates virtual constellations with similar properties to those obtained when directly designing constellations in four-dimensional space. This is acknowledged in [52], where a direct maximum-likelihood method for symbol de-mapping is indicated, which operates on these *virtual 4-D constellations*. An enhanced poly-polarization modulation (E-PPM) is proposed in [99], where constellation shaping is used to further increase the power efficiency.

The 64-point poly-polarization constellation disclosed in [52] has been integrated into the `Matlab` simulator to reproduce these recent findings. The simulation results are plotted in Figure 4.8. On one hand, PolyPM has a significant power advantage of more than 1 dB for all practical cases with respect to 8-PSK[5]. On the other hand, comparison with 8-PSK is inappropriate because it is a constant-power modulation. Comparing the 64-point PolyPM with the 64-4D-PSK constellation proposed above reveals a gap of approximately 1 dB. It seems that PolyPM cannot fully exploit the possibilities offered by the four-dimensional channel to

[5] 8-PSK was chosen by the authors of [52], their results are only restated here.

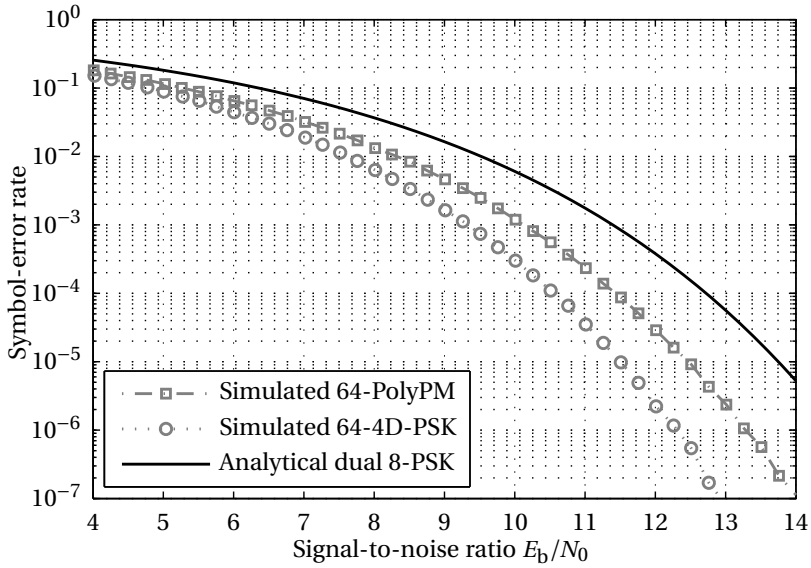


Figure 4.8.: SER for Poly-Polarization Modulation and 4-D modulation.

optimize power efficiency.

The PolyPM signal is constructed from three constant-power signals. By their superimposition, more than one power level is created leading to a signal with a higher PAPR than the underlying QPSK signals. The PAPR of the analysed 64-point constellation is displayed in Table 4.1; it is comparable to that of 16-QAM. As a first conclusion, considering the lower power efficiency and the higher PAPR with respect to the 64-4D-PSK constellation, PolyPM does not show any obvious advantage.

An alternative interpretation of the PolyPM signal can be attempted. The signal consists of three independent, but synchronous, QPSK signals. It is possible to uplink these QPSK signals individually, amplify them using saturated tubes and combine them only *after* amplification. Only some downlink spectrum might be recovered, but the disadvantage from high envelope variations would vanish. In this case, one will have to find a way to combine these interfering high-power signals in a lossless manner. The inventors of PolyPM do not seem to have considered such a solution, as they have carefully avoided the non-linearity question so far.

In conclusion, it can be said that while PolyPM seems to achieve comparable results to the constellations presented above, it does not permit the power

efficiency of the four-dimensional approach to be attained. In a slightly different scenario, when only downlink is considered, poly-polarization modulation could be a powerful way to increase the signalling efficiency if lossless combination is possible.

4.3.4. Preliminary conclusions

A selection of different four-dimensional constellations has been proposed in an effort to increase the signalling efficiency for satellite communications. Their respective power efficiencies were simulated in the AWGN channel. All the proposed constellations showed increased performance upon comparison with the traditional reference methods. The simulated gains are reasonably high for the lattice-based constellations, but diminish once more design constraints are introduced to adapt the constellations to the non-linear satellite channel.

The envelope variations of these constellations plays an important role in their evaluation as well. The four-dimensional bi-orthogonal constellation is an exception to this and a substantial gain can be achieved without the price of increased PAPR.

The reader familiar with the DVB-S2 [28] standard might wonder why the 16-QAM constellation was selected for the comparison with 256-D_4-LAM, instead of 16-APSK, which is more popular in satellite communications for its increased robustness towards non-linear amplification. Like 16-QAM, the 256-point LAM constellation is lattice-based and their PAPR values are similar. A comparison with 16-APSK would have resulted in a stronger advantage for the four-dimensional constellation, much like in the comparison of 88-D_4-LAM with 8-PSK. Therefore, such a comparison was judged inappropriate because unfair.

I consider it more helpful to directly search for APSK-like constellations that are optimized for 4-D signalling. Unfortunately, this would have gone beyond the scope of the present survey, as a specialized optimization algorithm would have to be implemented for this purpose.

4.4. Measurements

The previous experiments on polarization shift keying have demonstrated the possibility to conjointly modulate two carriers on orthogonal polarizations. Contrarily to PolSK, the frequency conversion on the transponder does not need to be coherent to properly receive a four-dimensional signal. This kind of transmission can be more appropriate if the coherency between polarizations cannot be ensured or if the advantages of PolSK are not specifically required in the given channel.

As this is the case for the satellite used in the experiments, it was deemed to be of value to demonstrate a four-dimensional transmission link over this satellite as well. In addition to the proof of feasibility there is a psychological aspect. The impact of showing an uncoded transmission with higher power efficiency than QPSK has the potential to shift the dual-polarization signal model into focus and to possibly change convictions in the satellite community about the utility of conjoint modulation.

Two measurement series were undertaken with the implemented modem for bi-orthogonal signalling. The first, in the thermal noise channel, compares the actual modem performance with the theoretical predictions and prepares the satellite demonstration. The second measurement series was gathered during the live satellite transmission. The initially planned measurements on the transponder simulator had to be abandoned due to sporadic equipment failures which made these experiments extremely time-consuming and imprecise.

4.4.1. Thermal noise channel

The test assembly is the same as for the PolSK measurements in the intermediate frequency channel, displayed in Figure 3.10. The receiver locked to the signal successfully. Again, the symbol-error rate is measured in the linear channel for different levels of thermal noise. The results are plotted in Figure 4.9, alongside the corresponding prediction from the simulator. As the symbol-error rate for QPSK, having higher spectral efficiency, is included as well, the SNR values are reported as E_b/N_0.

The measured differences between the simulated and the implemented system are small; they overlap for low SNR until about 5 dB, then a slight loss becomes visible for the demonstrator. This implementation loss looks qualitatively like that found for the PolSK demonstrator in Figure 3.11. The difference is smaller, because the bi-orthogonal signal is operated in a noisier environment and the impact of self-noise reduces with increasing external noise. It is highly probable that the reasons for the deviation are the same: non-compensated imperfections in the frequency response and the absence of symbol interpolation.

Despite this implementation loss, the power efficiency of the link is clearly higher than for QPSK.

4.4.2. Realistic channel tests

In the satellite experiment, the system's end-to-end functionality is demonstrated in its target setting. The same satellite, transponder channels and ground equipment were used as for the PolSK demonstration. The uplink was transmitted from the 11 m ground station antenna in Betzdorf to the *Astra 3B* satellite. Signal and

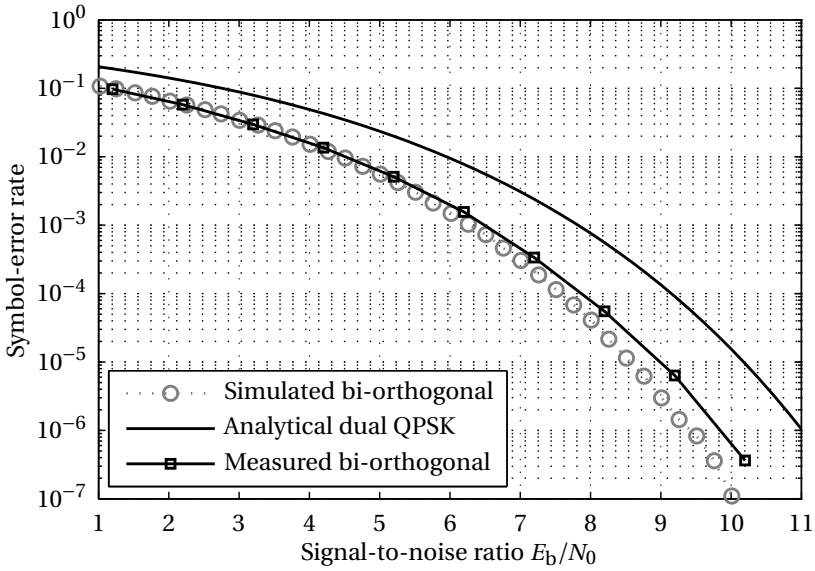


Figure 4.9.: Measured power efficiency of bi-orthogonal signal in thermal noise.

noise power were measured using noise markers on a spectrum analyser. During the installation, the sky was covered with thick clouds. A slight precipitation turned gradually into dense rain during the measurements. The experiment was conducted on Wednesday, 13th August 2014.

The receiver locked flawlessly to the returned bi-orthogonal signal. During the transmission the uplink power is varied, thereby modifying the relationship to the constant noise power level. In this way, the error rate was measured for different SNR levels; the results are displayed as diamonds in Figure 4.10.

The satellite demonstration confirms the previous on-ground findings, even though the measured curve looks slightly *steeper* than in the IF experiment. The noise power spectral density was measured at a frequency of 4 MHz above the actual carrier, to allow simultaneous measurement of signal and noise power. This led to a small bias in the SNR which might explain this phenomenon.

4.5. Discussion

The preceding sections have shown how a signalling advantage can be achieved by combining signals from both polarizations. Four-dimensional constellations, derived according to various design constraints, have been shown to provide an

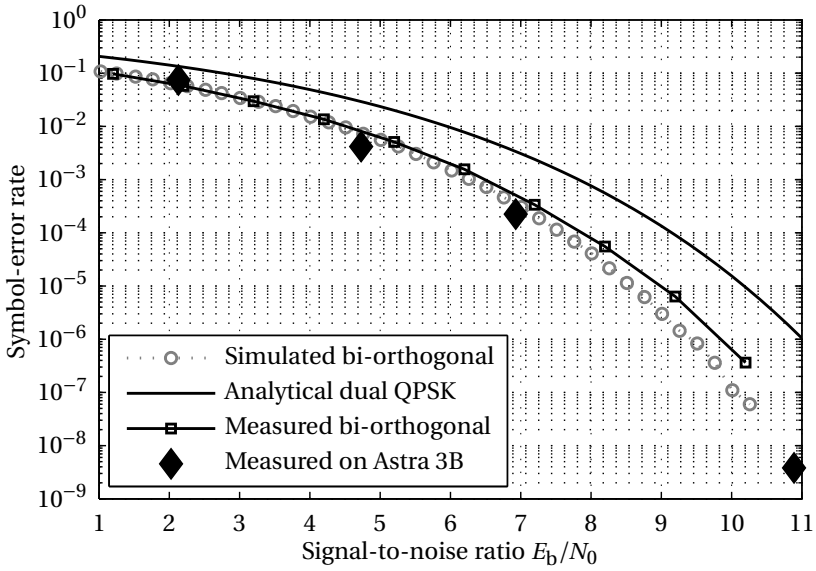


Figure 4.10.: Results from the SER measurement during the satellite demonstration of a 4-D, bi-orthogonal signal.

advantage in power efficiency with respect to traditionally used 2-D constellations. These *gross* advantages might be put into perspective when the associated increase in PAPR is considered as well. For this reason the field of application for these designs needs to be carefully chosen.

It has to be remembered here that the power gains from these advanced constellation designs have been achieved by applying techniques from coding theory. Also, the proposed constellations suffer from an increased peak-to-average power ratio and constellation expansion. Both phenomena are well known in the coding literature [31]. This leads to the question about the actual difference between the presented modulation techniques and a coded signal. Can the same benefits not be achieved by simple coding techniques?

4.5.1. Similarity to coding

On one hand, all the analysed modulation techniques have been considered with the premiss of uncoded signals. The simulated and measured symbol-error rates are those achieved from raw modulation without any error correction. On the other hand, techniques from coding theory were used to improve the signalling efficiency of these uncoded modulation techniques. By doing this, the bound-

ary between coding and modulation blurs, which may lead to wrong conclusions. The following discussion attempts to clarify this dilemma.

The signalling efficiency is improved by increasing the number of dimensions from two to four. The so-created freedom is used to optimally place constellation points; a placement that is *better* than would be possible in two dimensions. In coding, additional dimensions are created by merging consecutive symbols. For instance, a block of ten consecutive symbols can be redefined as a ten-dimensional complex, or 20-dimensional real codeword[6]. This is essentially what happens in a signal encoded with a block encoder Signals of an arbitrarily high number of dimensions can in principle be created, the limit is given by the decoding complexity and the acceptable time delay. In such a high-dimensional codeword, techniques to increase coding and shaping gain can be used and lead to improved error robustness, i.e. power efficiency, of the code.

In order to illustrate the difference between both approaches, a first step is to clarify two ways of decoding in the receiver: hard- and soft-decoding. A receiver, having sampled the matched-filtered waveform, ideally possesses a noisy symbol value. Based on the symbol's position in the constellation diagram, it can take a decision in favour of the most probable constellation symbol and extract the codeword associated to it. This codeword is then passed on to the error decoder. This is the hard-decoding technique; de-mapping and decoding are neatly separated.

A more efficient approach used by contemporary receivers is soft-decoding. Here the sampled constellation symbol is passed directly to the decoder. The decoder can use the additional information on the position of the symbol as a measure for the *certainty* with which a certain constellation symbol was transmitted. In a soft-decision receiver, de-mapping and decoding are thus performed together.

When hard-decisions are used in a dual-polarization receiver, the advantage of the dimensional increase is gathered upon de-mapping, before the decoder. In this case, the respective signalling advantage is independent of the actual coding. It is justified to show uncoded SER curves as a good performance indicator, as is done above.

The subtlety manifests itself when looking at soft-decoding. In this case, the received samples are passed directly to the decoder and there remains strictly no difference between symbol-elements on two polarizations or symbol-elements from consecutive symbols. The effort from 4-D constellation design plus the actual coding merge into a superordinate code. The impact on SER curves cannot be generalized in this case, as it depends on the coding technique used.

In 4-D signalling, the de-mapping can be regarded as a kind of inner soft-

[6] Also different frequency bands might be used to transmit more dimensions.

decode, while the actual decoding would be the outer decoding. The encoder stage is realized by increasing the dimension using two polarizations. Had it been implemented by merging every two symbols into a block code, the result would be exactly the same. Rather than using $B = 1/T$ on each polarization, the signal would require $2B$ on one polarization. In this sense, the four-dimensional bi-orthogonal constellation is equivalent to a block-coded QPSK on a single polarization with soft-decoding.

Such a block code, by increasing shaping or coding gain using the known techniques, poses the same challenges as the four-dimensional constellations. The *constituent two-dimensional constellations*, like those illustrated in Figure 4.1, expand and increase in PAPR [32]. A factor limiting the achievable coding gain from increasingly dense lattices in higher dimensions is that the number of direct neighbours grows for each symbol [31] as well. This has a detrimental effect on the symbol-error rate, as illustrated by the union bound expression in equation (4.6). In addition, a symbol error will more likely lead to multiple bit errors.

In conclusion, the proposed design techniques provide the same advantages and suffer from the same challenges irrespective of whether they are used as error-correcting codes or as symbol constellations for four-dimensional modulation. In the physical world of satellite communications, with amplifiers and bandwidth restrictions, the actual waveforms matter, rather than their abstract mathematical representation. These are the *constituent two-dimensional constellations* like those shown in Figure 4.1, together with the used pulse-shape.

To exploit the raw signalling advantages shown above, there has to be a physical reason or effect that can be exploited and which supports the choice for selecting a dual-polarization signal, rather than simply coding. Different physical constraints or conditions might act in favour or not of a polarization-, time- or frequency-based multidimensionality. Some possible effects are mentioned below in the conclusions.

4.6. Conclusions

This chapter analysed the coherent exploitation of the dual-polarization satellite channel. A range of different signalling constellations was proposed. It was shown that appropriate constellation design improves the modulations' power efficiency. A particularly interesting four-dimensional constellation, the bi-orthogonal signal, was implemented and demonstrated to be usable on contemporary satellite links without significant detriment.

Discussing the peak-to-average power ratio and the subtle link to coding, it was concluded that for their beneficial application, dual-polarization techniques require a physically advantageous setting. It can be difficult to justify receiving

two polarizations when the power efficiency gain, achievable equally well by coding, is the only reason. A selection of such physical advantages is proposed below.

The original project definition deliberately excluded error-correction codes to limit the scope of the work. Very sensible for the analysis of polarization shift keying, this decision may be questionable when modulation techniques with a strong link to coding techniques are to be analysed.

A strong argument in favour of dual-polarization signalling methods is the amplifier identified above in Section 4.3.2. An amplifier, able to amplify two orthogonally polarized waveforms at the same time, using a common power constraint, removes the detrimental effects of constellation expansion and increased PAPR. Very efficient constellation designs could be used without requiring back-off. The question of whether such a device can be constructed remains open.

There are other benefits, maybe less significant, but not relying on any hypothetical devices. For example, dual-polarization signalling allows bandwidth reduction. Signal processing devices could replace components with a bandwidth B by two identical parts of bandwidth $B/2$. This possibility of reduction can be joined into a general tendency for flexible resource usage in the channel [25].

In channels where crosstalk between polarizations is an issue, demodulating both polarizations would permit digital compensation and interference cancellation. This is discussed in [97], where this is considered even for channels with non-equal local oscillators on each polarization. For example, this could be used to redress polarization skew from wide-angle multi-head LNB's (low-noise block converter) in consumer reception equipment. A dual-polarization receiver could also electronically transform a signal received using a circular-polarized antenna into horizontal and vertical components and vice-versa.

The proposed advantages do not necessarily require the four-dimensional constellations to be used. It can be sufficient just to receive signals on both polarizations, for instance a traditional modulation technique in dual operation. In general, any quantities that are equal, or strongly correlated, on both polarizations can be exploited somehow. In polarization shift keying, the correlation of phase noise is used.

Also the other timing values like symbol clock or carrier frequency can be correlated on both polarizations. A final physical argument for reception is that the energy from both received signals can be used for the purpose of synchronization. This is analysed in the following chapter.

5. Synchronization of dual-polarization receivers

In addition to the benefits in power efficiency expected from a four-dimensional modulation technique, a couple of physically more direct advantages are possible. One of these is the possibility of using the signal from both polarizations to synchronize receivers.

Equal timing quantities on both polarizations, like phase or frequency offset or the symbol clock, allow a dual-polarization receiver to enhance its tracking capabilities. This is interesting for satellite systems where the received power can be doubled in this way, while the transmitted power is generally limited by physical or regulatory constraints.

This chapter elaborates on this conjecture. The signal model for synchronization is extended to both polarizations and the modified Cramer-Rao bounds are calculated for the case of dual reception. The findings are subsequently confirmed by simulating a dual timing-synchronizer for 16-QAM.

For the presented technique to work, the receiver is required to lock onto the signal on both polarizations. Demodulating both streams is necessary only insofar as it is a prerequisite to locking. It is therefore not important whether a four-dimensional signal or a polarization division multiplex is received — as long as the timing quantities are equal.

5.1. Signal model

After extending equation (2.30) to two modulated carriers, the prerequisites of equal timing quantities are subsequently justified and formalized one by one:

$$\begin{aligned} s_{\mathrm{x}}(t) &= e^{\mathrm{j}\left[\left(\omega_{\mathrm{c}}+\Delta\omega_{\mathrm{x}}\right)t+\phi_{0,\mathrm{x}}\right]} \sum_i c_{\mathrm{i,x}} \cdot g\left(t - iT - \tau_{\mathrm{x}}\right) \\ s_{\mathrm{y}}(t) &= e^{\mathrm{j}\left[\left(\omega_{\mathrm{c}}+\Delta\omega_{\mathrm{y}}\right)t+\phi_{0,\mathrm{y}}\right]} \sum_i c_{\mathrm{i,y}} \cdot g\left(t - iT - \tau_{\mathrm{y}}\right) \end{aligned} \tag{5.1}$$

Usually these prerequisites can only be met when both polarizations are uplinked from the same location. Identical carrier frequency offsets $\Delta\omega$ are technically the least difficult requirement. Frequencies are equal when the transponder channels

overlap and the same reference is used for the local oscillator on the satellite:

$$\Delta\omega_{\mathrm{x}} = \Delta\omega_{\mathrm{y}} = \Delta\omega \tag{5.2}$$

It is not realistic to assume exactly the same time delay on both polarizations. It is technically challenging to set up equal path lengths on separate satellite transponders and even on ground stations. Therefore a fixed deviation is expected between the polarizations for the phases ϕ_0 and the timing offsets τ. Such a *geometrical* offset does not change over time however and could be estimated by a long-term average. In that case any delay differences can be neglected and phase and symbol clock be assumed synchronous:

$$\phi_{0,\mathrm{x}} = \phi_{0,\mathrm{y}} = \phi_0 \tag{5.3}$$

$$\tau_{\mathrm{x}} = \tau_{\mathrm{y}} = \tau \tag{5.4}$$

The expression in equation (5.1) can now be simplified and put into vector form:

$$\mathbf{s} = e^{\mathrm{j}[(\omega+\Delta\omega)t+\phi_0]} \sum_i \begin{pmatrix} c_{\mathrm{x,i}} \\ c_{\mathrm{y,i}} \end{pmatrix} \cdot g\left(t - iT - \tau\right) \tag{5.5}$$

The symmetry in timing quantities can be exploited to perform timing estimations over both polarizations. In this way, thermal noise-induced jitter can be reduced without extending the measurement period.

For each of $\Delta\omega$, ϕ_0 and τ, different detectors exist; some are presented in Section 2.5.1. They yield an estimate in form of an *error signal* that is subsequently averaged and tracked by a PLL (cf. Section 2.5.2). The enhanced tracking performance can be achieved by adding the error signals from both polarizations and feeding them to a single PLL.

5.2. Tools and methods

The analysis of the proposed enhancement is conducted in two parts. The first part uses the model in equation (5.5) to derive the MCRB for reception on both polarizations and compares it with the classic MCRB derived in Section 2.5.3. In the second part, a simulation is run to confirm the findings.

5.2.1. Dual-polarization MCRB

The modified Cramer-Rao bound was introduced in Chapter 2 and computed for the case of symbol timing estimation. It presents a lower bound for the variance achievable by an unbiased estimate in a Gaussian noise process [22].

Starting from the model in equation (5.5), the MCRB derivation from Section 2.5.3 is repeated here for the dual-polarization case. The probability for observing the dual-polarization signal $\mathbf{r}(t)$ for given λ and $\mathbf{u}$ is:

$$p_\lambda(\mathbf{r}|\mathbf{u}) = \frac{1}{(2\pi N_0)^2} \exp\left\{\frac{-1}{2N_0}\big(\mathbf{r}(t) - \mathbf{s}(t,\lambda,\mathbf{u})\big)^{\mathrm{H}}\big(\mathbf{r}(t) - \mathbf{s}(t,\lambda,\mathbf{u})\big)\right\} \tag{5.6}$$

By analogy to the derivation for scalar signals in [57], a log-likelihood function for the estimation of a constant parameter λ in a noisy, dual-polarization signal can be found:

$$\ln \Lambda(\lambda,\mathbf{u}) = \frac{-1}{2N_0}\int_{T_0}\big(\mathbf{r}(t) - \mathbf{s}(t,\lambda,\mathbf{u})\big)^{\mathrm{H}}\big(\mathbf{r}(t) - \mathbf{s}(t,\lambda,\mathbf{u})\big)\,dt \tag{5.7}$$

This allows in turn the expression for the MCRB to be generalized to a vector signal:

$$\mathrm{MCRB}(\lambda) = \frac{1}{\mathrm{E}_{\mathbf{n},\mathbf{u}}\left[\left(\frac{\partial \ln \Lambda(\lambda,\mathbf{u})}{\partial\lambda}\right)^2\right]} \tag{5.8}$$

$$= \frac{-1}{\mathrm{E}_{\mathbf{n},\mathbf{u}}\left[\frac{\partial^2 \ln \Lambda(\lambda,\mathbf{u})}{\partial\lambda^2}\right]} \tag{5.9}$$

$$= \frac{N_0}{\mathrm{E}_{\mathbf{u}}\left[\int_{T_0}\left(\frac{\partial\,\mathbf{s}(t,\lambda,\mathbf{u})}{\partial\lambda}\right)^{\mathrm{H}}\left(\frac{\partial\,\mathbf{s}(t,\lambda,\mathbf{u})}{\partial\lambda}\right)\,dt\right]} \tag{5.10}$$

$\mathbf{u}$ contains the parameters that are not to be estimated and thus modelled as random, as well as the data symbols $\mathbf{c}_\mathrm{i}$ [1]. The first equality is a known transformation for expected values exposed in [82, p. 67]. Equation (5.10) is obtained by performing the derivations and taking the partial expected value over the noise process $\mathbf{n}$.

The expression (5.10) is the vectorial equivalent of equation (2.38). To compute the MCRB from equation (5.10), the steps are analogous to the scalar derivations:

$$\mathrm{E}_{\mathbf{u}}\left[\int_{T_0}\left(\frac{\partial\,\mathbf{s}(t,\lambda,\mathbf{u})}{\partial\lambda}\right)^{\mathrm{H}}\left(\frac{\partial\,\mathbf{s}(t,\lambda,\mathbf{u})}{\partial\lambda}\right)\,dt\right] \tag{5.11}$$

$$= \int_{T_0}\mathrm{E}_{\mathbf{c}_\mathrm{i}}\left[\left(\sum_i \mathbf{c}_\mathrm{i}^{\mathrm{H}} g'(t-iT-\tau)\right)\left(\sum_k \mathbf{c}_\mathrm{k} g'(t-kT-\tau)\right)\right]\,dt \tag{5.12}$$

[1] Technically, $\mathbf{u}$ is no longer a vector, because $\mathbf{c}$ is a matrix. This has no influence on the argument however.

The only vectorial quantities are the symbol values. They are eliminated by resolving the symbol energy as $E_\mathrm{s} = \frac{1}{2}\mathrm{E}\left[\mathbf{c}_\mathrm{i}^\mathrm{H}\mathbf{c}_\mathrm{i}\right]$, according to equation (3.5). From this point onwards, the derivation is identical to the scalar case. The dual-polarization MCRB for timing estimates is:

$$\frac{1}{T^2} \times \mathrm{MCRB}(\tau) = \frac{B_\mathrm{N} T}{4\pi^2 \zeta} \frac{N_0}{E_\mathrm{s}} \tag{5.13}$$

with ([57, pp. 130,216,383]):

$$T_0 = \frac{1}{2B_\mathrm{N}} \tag{5.14}$$

In the same way, the calculation of the MCRBs for ϕ_0 and $\Delta\omega$ can be accomplished along the lines of the scalar derivations in [57, pp. 58-62]. The results are:

$$\left(\frac{T}{2\pi}\right)^2 \times \mathrm{MCRB}(\Delta\omega) = \frac{12\,(TB_\mathrm{N})^3}{\pi^2} \frac{N_0}{E_\mathrm{s}} \tag{5.15}$$

$$\mathrm{MCRB}(\phi_0) = TB_\mathrm{N} \frac{N_0}{E_\mathrm{s}} \tag{5.16}$$

Looking at the derivations, it seems that timing synchronization can profit from the dual-polarization estimate even in the event of non-equal carrier frequencies or phases. The same is not necessarily true for the estimation of ϕ_0 or $\Delta\omega$.

Preliminary conclusion

For each of the three parameters τ, $\Delta\omega$ and ϕ_0, the MCRB resolves to expressions equal to those shown in [57]. The difference lies in the interpretation of the symbol energy E_s, which here corresponds to the definition from equation (3.5).

A transmission over two polarizations will in general have twice the power of a single polarization transmission, because two amplifiers are used instead of one. This effectively doubles the value of E_s and the corresponding MCRB will be half as large for dual reception. As a consequence, synchronizing using a dual-polarization signal has the potential to reduce the jitter variance by a factor two.

A similar consideration can also be made for a pilot-aided tracking system. If the quantity to be estimated is equal on both polarizations, twice the number of pilots are available to estimate the same quantity. Alternatively, one could just transmit pilots on one carrier, leaving the second exclusively for payload data. In this case, the jitter variance would remain the same, but the efficiency of carrier usage would increase.

5.2.2. GED simulation

The MCRB only gives an indication of how well a best possible estimator would perform. To corroborate the previous findings by a specific example, the vari-

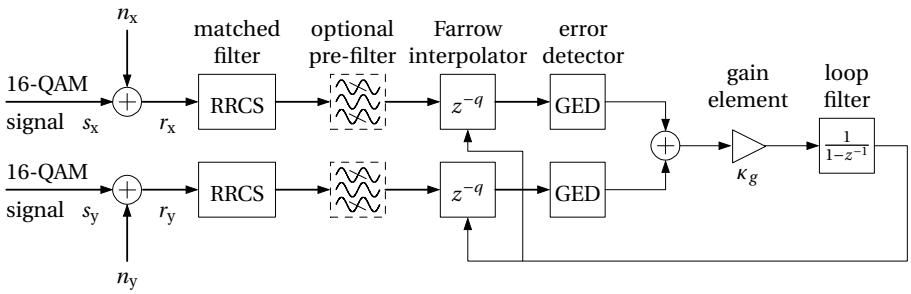


Figure 5.1.: Block diagram for Gardner-Error-Detector simulations on a dual-polarization 16-QAM signal.

ance of jitter in a timing recovery loop with the Gardner Error Detector (GED) [37] is simulated. For this, 16-QAM signals are simulated in a single- and dual-polarization setting.

A block diagram of the simulated system is shown in Figure 5.1. The simulation is based on the `Simulink` software which allows block-based processing in a signal flow-graph.

Two noisy 16-QAM signals, modelled as pulse-shaped, complex baseband signals, serve as inputs and are matched-filtered by a root-raised-cosine (RRCS). Optional pre-filters allow the GED's self-noise to be reduced, thus leading to better performance especially in the high-SNR case [21]. The pre-filters are 3-tap FIR filters; their impulse responses are drawn from the pertinent plot in [21].

The timing recovery loop has two GEDs, one for each polarization. They realize the formula from equation (2.32) in complex baseband. A block diagram of the implementation is displayed in Figure 5.2. Both error signals $\Delta\varepsilon[n]$ are summed and integrated to provide an input to two Farrow-interpolators [26]. This closes the recovery system's feedback loop.

The error variance of this tracking system is simulated and compared to the MCRB computed from equation (5.13). The symbol frequency in the simulation is 1 MBd. A simulation run lasts for 1,1 s in simulation time. The first 0,1 seconds permit the loop to settle into steady state. In the remaining time, the loop filter's register is read after every symbol, providing $1 \cdot 10^6$ filtered error samples. The final variance is estimated from these samples.

B_N is the noise bandwidth, chosen as $B_N = 500$ Hz. B_N is computed as follows from the loop parameters:

$$B_N = \frac{\kappa f_s}{4 - 2\kappa} \tag{5.17}$$

The loop gain κ is the combined gain from the detector and the gain element:

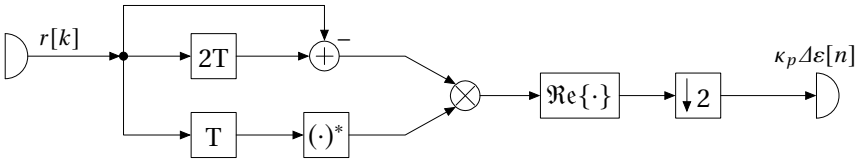


Figure 5.2.: Gardner Timing Error Detector implemented in complex baseband.

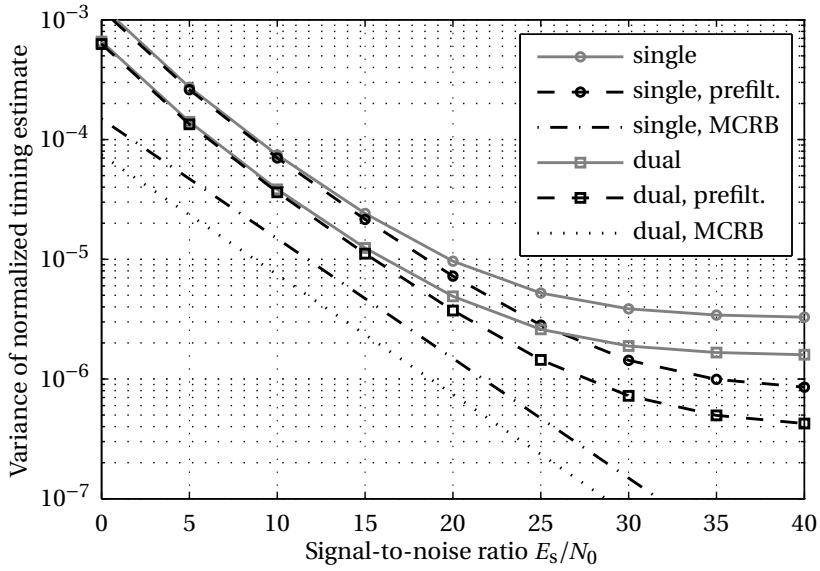


Figure 5.3.: Variance of normalized timing estimates

$\kappa = \kappa_p \kappa_g$. The sampling frequency f_s in the loop is equal to the symbol frequency. The parameter ζ, necessary for the computation of the MCRB, can be computed from equation (2.51). The simulation uses a roll-off value of $\alpha = 0{,}20$, leading to $\zeta = 0{,}852$.

5.3. Simulation results

Figure 5.3 shows the normalized timing variance as a function of E_s/N_0 for the simulated single- and dual-polarization signals with and without pre-filtering as well as the corresponding MCRBs.

In the low-SNR region, the variance of the estimates lies roughly five times

above the respective MCRBs. This is expected behaviour; according to [57, p. 58], the MCRB implicitly assumes known data symbols[2] and can be approached only by decision- or data-directed synchronization methods.

In the high-SNR region, the GED estimates saturate; self-noise in the detectors prevents the variance from decreasing further. As expected, the pre-filters are able to mitigate this phenomenon and the estimates are improved by a factor four at SNR = 40 dB.

The most important observation is that the jitter is reduced by half for all noise levels in the dual-polarization setting. This is true for the pre-filtered and non-pre-filtered sequences even up to the asymptotic performance limits. This means that the dual synchronization technique improves the estimates not only with respect to thermal but to the detector's self-noise as well. In the given scenario, dual-polarization tracking outperforms pre-filtered tracking on a single polarization up to a SNR of about 25 dB, which includes most practical cases.

5.4. Conclusion

It can be concluded that a dual-polarization receiver can increase its symbol clock tracking precision by performing the estimation over both polarizations. Judging from the respective MCRB expressions, the same should be true for carrier and phase recovery. In satellite communications the transmission power can usually not be increased, but a dual-polarization signal will have doubled symbol energy. In a receiver, where loop bandwidth is often the only tunable parameter, synchronization over both polarizations can be a satisfactory way of avoiding possibly conflicting requirements.

Both implemented demonstration receivers from Chapters 3 and 4 use such a dual-polarization error signal for symbol clock tracking. The idea has thus been implemented in practice, even though squaring devices were used instead of the Gardner error detector. No measurements of the improvement of tracking fidelity have however been undertaken.

One observation is that when a dual-polarization receiver implements symbol tracking using either the square or Gardner error detector, it should be immune to polarization crosstalk.

Section 3.4.1 discussed how significant differences in path delay can prevent a timing lock on both polarizations. In this case, a long-term average as mentioned in Section 5.1 might be an appropriate remedy to absorb the delay difference.

[2] and long data sequences

6. General conclusions

It was shown, by simulation and experimental demonstration, that polarization shift keying can be used in satellite links. With the help of a polarization tracking device, the path length difference between both polarizations is compensated and the phase noise immunity described in the context of optical communications can be maintained. Due to the relatively strong envelope variations on each carrier, even for constant-power constellations, it is recommended to operate PolSK on linearized amplifiers with a little back-off.

With respect to power efficiency, polarization shift keying cannot perform as well as the traditional modulation techniques in a polarization division multiplex. Its use should in consequence be limited to cases where phase noise dominates the channel impairments. For satellite communications, these could be narrow-band channels, where the signalling bandwidth and the bandwidth of the phase-noise spectrum are of similar magnitude. Another possible target scenario are bursty transmission schemes; with little time for synchronization, a direct-detection scheme is ideal. The implemented polarization-switching functionality should enhance the synchronization of the symbol clock, which for bursty transmissions could be recovered in a fast feed-forward manner.

When the transmission chain in a satellite link includes non-coherent frequency conversions, the use of polarization shift keying is not recommended. Such a frequency conversion neutralizes the phase-noise immunity – the principal advantage in favour of PolSK. For the demonstration shown in this dissertation, the PolSK receiver included a system to recover such incoherent carrier frequencies: this is certainly possible but can only be justified as a fallback option or in a scientific context but not when spectrum or power consumption are subject to economic considerations. In general, this recovery system should rather be dimensioned to compensate the difference in path length.

As an alternative to polarization shift keying, the four-dimensional modulation technique was proposed. It is a fully coherent technique, closer to the traditional modulation methods and to the polarization division multiplex. It can be used when the specific advantages of PolSK are not required or when the carriers on both polarizations are incoherent.

4-D modulations permit the exploitation of the channel at higher power efficiency than possible on a single polarization. The benefits achievable by constellation design in four dimensions should be encouraged by well-disposed physical

circumstances. Otherwise, the efficiency gain is indistinguishable from the one achievable by coding.

The most pre-eminent of these favourable circumstances could be the amplification by a dual-polarization amplifier, alleviating the power constraints for 4-D constellations and freeing up one signalling dimension. As such an amplifier seems out of reach at the present time, the advantages associated to dual-polarization signalling techniques are more modest.

First, in all shown examples, the power efficiency of the four-dimensional constellation was never worse than that of its traditional counterpart. There is always the possibility to resort to a 4-D constellation corresponding exactly to a traditional modulation technique in *dual* operation. The technology is in that sense backwards compatible. A range of additional benefits have been identified, associated to various aspects of dual-polarization reception.

Dual-polarization reception allows electronic compensation of polarization crosstalk. This is especially useful for small reception antennas or modern multi-head LNBs with wide view angles, where skew can become an issue. In the same category falls the capacity to transform two circular into two linear polarizations and vice versa, irrespective of the actual antenna equipment used.

A second category of advantages are the possible synchronization benefits. This was demonstrated in Chapter 5 using the example of timing clock recovery, but it should also work for frequency and phase recovery.

Finally, setting up a link on both polarizations allows the bandwidth to be split between two carriers. Two amplifiers are used in this case and the transmission power can be doubled with respect to using an equally wide bandwidth contingent on a single polarization. For future transmissions requiring more spectrum than usually fixed by the channel and filter bandwidths[1], this can be a valid alternative that avoids switching to another frequency.

Taken individually, these different advantages may seem modest, but in combination they show the potential to enhance the general flexibility for setting up a communication link over satellites at a low additional cost.

6.1. Perspectives and recommendations

Considering the findings presented here, I suggest that research on polarization shift keying should be continued only after more concrete application cases have been identified. Broadcasting and broadband data links are clearly not the target application for PolSK.

During the project, the actual tracking performance of the timing recovery with the modified PolSK signal has neither been simulated nor measured. As I

[1] 36 MHz for many broadcasting applications

consider the possible improvement to be a significant advantage, it would be interesting to investigate this further.

Due to limitations with the available equipment, demonstrating a functional polarization-tracking device, which would compensate an eventual path delay difference, was not possible. Using a different satellite, comprising a single local oscillator, such a transmission would demonstrate a plain PolSK signal both in uplink and downlink.

For private satellite operators providing large transponder bandwidths and high data throughput it is more interesting to consider fully coherent transmission on both polarizations. Any of the proposed advantages for dual-polarization systems, be it the four-dimensional constellations, synchronization, bandwidth reduction or cross-polarization compensation is useful by itself. In addition, a system specified to make use of one of these benefits gets all others practically for free; once the hardware capabilities are set up, all the features can be easily implemented using digital signal processing. From this perspective, using dual-polarization signalling can be very well justified.

The effort to equip satellite transponders and receivers with the required capability for dual-polarization transmission is relatively small.

The satellite transponder channels have to be frequency-aligned across polarizations. In addition, the local oscillators used for frequency conversion should be sourced from the same reference oscillator. In order to keep the same level of redundancy as in the case of individual LOs, each conversion module should have its own reference. Microchips can route the reference signal from one module to the other and switch it automatically in the event of a failure. The same changes need to be applied to large uplink or reception stations.

Concerning consumer equipment, today's LNBs already fulfil many of the requirements and are capable of providing two frequency bands on two polarizations to the reception equipment [70]. In a modern satellite receiver for television, the tuner plays only a minor role in the total cost. Dual and even triple tuner models exist, for the purpose of simultaneous recording and viewing. Doubling the number of tuners to allow combined reception on two polarizations is not a major cost factor.

Some investigations should be conducted to explore whether the proposed dual-polarization amplifier is physically possible and whether its development and construction can be justified economically.

There is a general trend in the satellite industry to consider optical links for feeder applications, combining light and microwave communications. I recommend to watch closely the developments in the optical signalling and modulation domain. Research in this field might yield some useful technologies that can potentially be *ported* to radio-frequency communication in the same way as was done here for polarization shift keying and four-dimensional signalling.

A. Simulation software

Initial performance estimates on the analysed modulation techniques were obtained by computer simulations. The simulation software, implemented in the `Matlab` computing environment, simulates a transmitter and a receiver providing the basic modulation and demodulation functionality described in Section 2.6. Three general classes of modulations are supported by the simulated modem.

Single-polarization: These are the classic modulation techniques, applied onto a single carrier wave. Examples are PSK, QAM or APSK.

Dual-polarization: The techniques based on conjoint modulation of both polarizations fall into this category: PolSK and 4-D signalling.

Dual operation: For a fair comparison with dual-polarization modulation techniques, single-polarization techniques are considered to be transmitted on both polarizations. The number of bits-per-symbol on each polarization is summed and the error probabilities are computed according to equation (3.37).

A large selection of single- and dual-polarization modulation techniques, variants and constellations can be selected for simulation.

Transmitter and receiver models are complemented by a channel simulator, which models the various impairments expected in the satellite link. According to the project's scope definition, only the impact of signal processing is simulated, whereas propagation effects have been left aside.

Waveforms are represented in complex baseband as sampled sequences of floating-point numbers. To handle random phenomena, like thermal or phase noise, the Monte-Carlo approach is followed; a random process is simulated by creating a sampled representation of a specific realization. When the number of samples is high enough, this realization can be considered to be representative for the entire process [49, pp. 371f.]. For the results shown in the present dissertation, around $1 \cdot 10^5$ to $1 \cdot 10^6$ samples have been generated per simulation run. For the accurate estimation of very low symbol-error rates, this number is increased up to thousand-fold by running consecutive simulations.

Two different simulation types have been developed and are described in the following. The first is a symbol-level simulation, where only the sequence of symbol values, subject to thermal noise, is simulated. The second simulation over-

samples the symbols and performs pulse-shaping. In this way, a complex baseband equivalent of the actual waveform is created and can be subjected to radio-frequency impairments, like non-linear amplification or phase noise.

A.1. Symbol-level simulation

The symbol-level simulation is the first implemented simulation type. It is simpler and faster than the full waveform simulation and is able to determine the detection performance in AWGN. A simulation run consists of the following steps.

1. Generation of a data stream as a random sequence of integer values.
2. Assignment of each data sample to a symbol according to the chosen constellation. A symbol contains the phase and amplitude values on either one or both polarizations, depending on the selected modulation class.
3. Estimation of the average bit-energy E_b and generation of white Gaussian noise of a variance corresponding to the targeted signal-to-noise ratio. The noise samples are added to the symbol values.
4. Symbol de-mapping according the optimal minimum-distance algorithm.
5. Comparison of the resulting data stream with the initial integer sequence to determine the number of symbol errors.

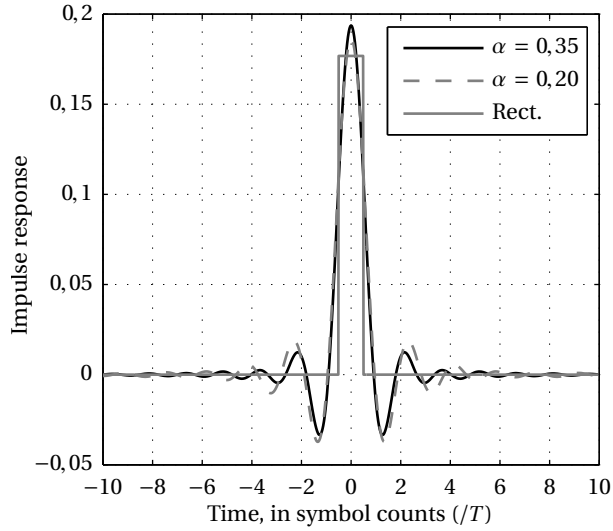
The simulation counts symbol errors, not bit errors. This keeps the resulting estimate independent of the chosen symbol-to-bit assignment. The choice of a particular assignment can change the bit-error-rate even though the detection capabilities of the demodulator are not altered. One purpose of the simulations is to try many different constellations. For many of those, finding a good assignment is an involved task which is therefore avoided. The bit-error rate is greater or equal to the symbol-error rate and generally approaches it in high SNR.

A.2. Waveform-level simulation

The simulation on the waveform-level is more flexible and allows the actual channel simulation of the satellite link. It simulates the physical waveform and, above thermal noise, various other impairments. The signal path of this simulator is displayed in Figure 3.1 in Chapter 3.

The crucial difference to the symbol-level simulation is the pulse-shaping. Symbol values are upsampled and interpolated using a root-raised cosine filter.

Figure A.1.: Root-raised cosine and rectangular pulse shapes



The excess bandwidth and upsampling factor can be tuned. For PolSK simulations, the default values are 35 % and 8 respectively. Four-dimensional signals have been simulated with an excess bandwidth of 20 %. The shape of the RRCS pulse is displayed in Figure A.1, together with a rectangular pulse for comparison.

Following the pulse-shaping, the signal is usually upconverted into the requested band and fed to an antenna. As the simulation operates in complex baseband, this upconversion is not explicitly performed. Consequently, all simulated impairments operate on the complex baseband signal. They are listed and described below.

Non-linearities Non-linear amplification causes signal distortion. The travelling wave tube amplifiers (TWTAs) used aboard of communication satellites are usually operated close to saturation and therefore show strongly non-linear behaviour. This behaviour expresses itself in an amplitude distortion and phase distortion, that are modelled by AM-AM and AM-PM curves. A widely used model for the amplification characteristic of such a tube is the Saleh-model [67], which is also used by the simulator. For the simulations, the Saleh-model was fitted to the characteristics a K_u-band amplifier tube on the *Astra 1K* spacecraft. The result is plotted in Figure A.2.

Using complicated pre-amplifiers, the non-linear behaviour of a TWT amp-

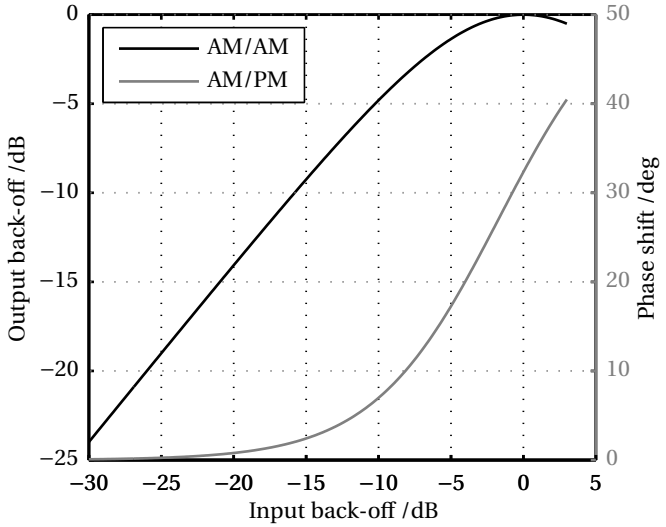


Figure A.2.: Saleh-model: AM-AM and AM-PM characteristics used by the simulator

lifier can be reduced. Such *linearized amplifier* systems may be used to support signals with higher envelope variations. The amplification characteristics of several linearized amplifiers are shown in Figure A.3.

The simulations shown in this dissertation used the characteristic from the DVB-S2 standard [28] because it represents a widely used and accessible model. The two other curves, showing a K_u and a K_a-band amplifier from the manufacturer *Tesat* have been measured during in-orbit tests of SES satellite transponders. The characteristics of the Tesat tubes differ qualitatively from the one in the standard. One might therefore question the legitimacy of this published data.

Phase noise The non-ideal behaviour of the oscillators in the modulation and signal processing chains leads to phase noise. Phase noise complicates the task of carrier tracking and may provoke detection errors. Phase noise is simulated by multiplying the signal with a complex exponential carrying the noise samples $\psi[n]$:

$$s_{\mathrm{PN}}[n] = s[n]e^{\mathrm{j}\psi[n]} \tag{A.1}$$

The noise samples themselves are white Gaussian noise, filtered by a single-order filter, as demonstrated in [29]. This leads to the phase noise power fall-

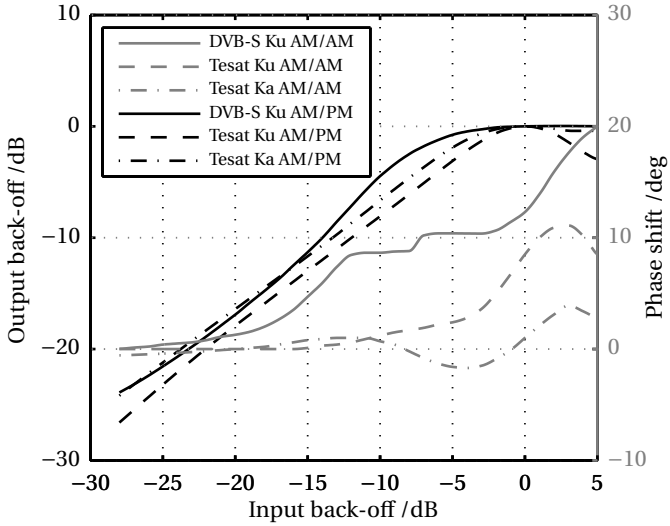


Figure A.3.: Linearized amplifier models used in the simulation software

ing off with $1/f^2$ distance to the carrier. The phase-noise level is chosen to match approximately a recommendation for LNBs issued by SES [69]. The mask and the realized power spectral density are displayed in Figure A.4.

As no actual tracking is performed in the simulator, only the visible effect of phase noise on the constellation diagram and related decision errors are accounted for. To simulate the effect of a *hypothetical* tracking loop with noise bandwidth $B_N = 1\,\text{kHz}$, an appropriate high-pass filter is applied to the noise process. This is visible in the plot in Figure A.4, where the phase-noise level drops considerably below 500 Hz. Thereby any strong low-frequency noise components that might lead to non-reproducible simulations are removed as well.

Frequency offset The exact carrier frequency is in general not known to the receiver in advance. The simulator allows to model a simple frequency offset by multiplying the waveforms with a complex exponential.

Thermal noise is modelled as white Gaussian noise and added in a last step after all other impairments. The noise power is determined based on the requested SNR and an estimation of the impaired signal's power. If the non-linearity has led to a significant transfer of signal power outside of the signalling band, the resulting SNR might actually be lower after the matched filters. This could lead to

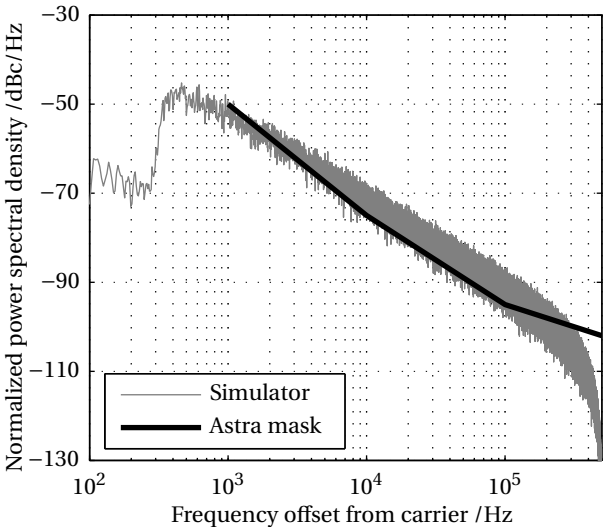


Figure A.4.: Phase noise mask and simulated power spectral density

slightly pessimistic estimates in high-distortion scenarios. For non-pathological cases, this effect can be considered secondary.

Following noise addition, the signal is filtered by the matched filters and downsampled at the optimal instant. The optimal instant is hard-coded into the simulator, as the occurring delays are caused by filtering only and are therefore determined a priori. This is an idealization as a real receiver would need to detect the symbol timing by itself. After downsampling, the detection and error counting is executed in the same way as for the symbol-level simulation.

A.3. Analysis of results

In addition to the SER estimate, which is the simulation's primary output, the generated waveforms are saved after each processing step. Essential characteristics, like average and peak power, or the SNR are estimated. A couple of tools are proposed to the user to inspect and analyse the saved waveforms.

- The complete sampled time series can be viewed.
- The power spectral density can be estimated and displayed
- An eye-diagram can be viewed.

- The symbol sequence, before up- or after downsampling, can be viewed in a constellation diagram, either in in-phase/quadrature representation or in Stokes-space.
- The estimated probability mass function of a signal's amplitude and power distribution can be displayed.

B. Modem implementation

A brief overview on the structure and functionality of the demonstrator modems is given in Chapter 3 and Chapter 4. This appendix completes these overviews with a more detailed presentation.

First, the overall system parameters are defined. Second, the FPGA development board as a platform for the implementation is presented and the design methodology is briefly exposed.

Then, both the PolSK transmitter and the receiver are presented by looking at each of the important processing steps in the signal chain in detail. Increased focus is put on the more unusual or specially developed parts. After the PolSK modem, the 4-D modem is described by only presenting the relevant changes with respect to the first.

Horizontal and vertical polarization are sometimes denoted X and Y, sometimes H and V. X and Y are used for digital signals, while H and V is reserved for the physical signals.

B.1. Hardware and system design

The design goal for the demonstration system is to build a satellite modem able to transmit and receive on two polarizations at once. The target application is a K_u-band transmission on orbiting spacecraft. The system should use the polarization shift keying modulation and be able to cope with the incoherent frequency conversion on the satellite.

B.1.1. System parameters

Satellite modems usually do not operate in K_u-band themselves but use an intermediate frequency in the L-band or at 70 MHz. A single-channel transmitter will usually work at 70 MHz IF. In a consumer system, the satellite signal is, after reception by the dish, amplified and downconverted in the LNB and fed to the reception tuner as an L-band signal. In this way, a large part of the transmission band can be provided to the user (e.g. television spectator) at once.

For the present demonstration there is no need to receive more than one channel, so that the intermediate frequency is fixed at 70 MHz for both the transmitter and the receiver. Lower frequencies are easier to master and equipment is

cheaper. In addition, the availability of various IF equipment (noise generators, impedance and frequency converters etc.) in the SES and University laboratories made this a straightforward choice.

The nominal bandwidth of a satellite transponder is 36 MHz (this is called TPE, transponder equivalent bandwidth). The bandwidth use of a bandpass signal is given by the symbol rate including the roll-off α:

$$B_{\mathrm{u}} = (1 + \alpha)/T \tag{B.1}$$

The careful decision was made to transmit at a symbol rate of 1 MBd. This should avoid any of the aforementioned difficulties with non-equal path length in the ground station and transponders. In addition, it preserves the possibility to signal on transponders with a channel arrangement in echelon and keeps the implementation complexity within reasonable limits. A similarly conservative choice was made for the roll-off, which was fixed to

$$\alpha = 0{,}35 \tag{B.2}$$

Digital satellite receivers usually comprise of a front-end tuner-chip selecting the channel and downconverting the received signal from L-band into complex baseband. This baseband signal is then sampled and processed by the actual digital receiver. Due to the low IF in the present case, it is possible to directly sample the signal and to do all processing in the digital domain.

B.1.2. Platform

There are two main possibilities to prototype digital modems. Using powerful *digital signal processors* (DSPs) with a couple of sequential processing chains or *field programmable gate arrays* (FPGAs), programmable logic devices possessing the parallel processing behaviour of a microchip. The availability of very suitable development platforms as well as my personal interest geared the choice towards an implementation on FPGAs. Reports on related previous achievements encouraged the decision [17, 83].

A readily packaged hardware solution for the task at hand is the *DSP development kit* from Avnet. It contains the *Xilinx KC705* development board [91] with a *Kintex-7* FPGA [89] and a number of buttons, LEDs and switches for user interaction. It is configured using a personal computer and a JTAG-to-USB interface device. The JTAG interface [45] also allows the online monitoring of FPGA internal register states, which is useful for debugging. A USB-based serial port is available to transmit data between a workstation and the FPGA. Onboard devices are configured by the FPGA using the I^2C interface [60].

The required digital-to-analogue and analogue-to-digital conversion capability is provided by an expansion card (FPGA mezzanine card, FMC), the *4DSP FMC150* [1]. It has a dual, 16-bit DAC and a dual 14-bit ADC. Both operate on a sampling frequency of 245,76 MHz, which is largely sufficient for a 70 MHz IF signal. The mezzanine card contains a clock chip to synchronize both converters and to provide a sample and data clock to the FPGA. A monitor chip supervises all supply voltages. The peripherals on the mezzanine card are configured over the SPI-bus [58]. Data transfers are carried out over high-speed DDR lanes.

In addition to the hardware capabilities, a software license for the corresponding design tools was included with the board. These tools provide the interfaces to the FPGA's internal functionalities, like clock management [88], I/O registers [90] and prefabricated building blocks for commonly used functionality like FIR-filters [94] or NCOs [93]. The modems are programmed in VHDL [46, 13] using the `Xilinx ISE` integrated development environment.

B.1.3. Design process and methodology

The design of the transmitter closely follows the model in the simulation software. The basic processing steps are the generation of random data, symbol selection, then pulse-shaping and upconversion into the target band. The implementation was accomplished respecting VHDL programming principles; the overall design was divided into many distinct functional units. Those were implemented separately and simulated using so-called *test beds*. Only after successful testing they have been integrated into the larger system design.

The completed transmitter was simulated with the `ISim` simulation software [95]. The resulting waveforms have been validated by comparing them to those created by the `Matlab` simulator. The transmitter is simpler to design than the receiver; its implementation allowed to familiarize myself with the programming language and to concentrate on the proper connectivity to the external ADC and DAC interfaces.

The receiver was first realized using a model in the block-based `Simulink` programming environment. It was fed with the transmitter's output waveforms yielded by the `ISim` simulations. The `Simulink` approach allowed a flexible design process that continued until the model was able to lock onto and demodulate the PolSK signal. As soon as satisfying reception quality was obtained, buildings blocks were gradually replaced by lower-level implementations. This process was reproduced down to a level of detail that allowed the transfer into digital logic.

Both the transmitter and the receiver consist, apart from their main functional unit, of several other parts. Bus-masters using the I^2C and the SPI standards are included to configure the microchips on the development board and the expansion card. A DDR register interface registers initiates and performs the

high-speed data transfers from the ADC and to the DAC chip [90]. A third unit implements the serial link to the PC for the monitoring of internal states. These different parts are controlled and synchronized by a central finite state machine (FSM), which is briefly described in Section B.2.4.

B.2. The PolSK transmitter

Figure B.1 shows a simplified block diagram of the transmitter. The design follows the textbook structure introduced in Section 2.6, except that it has one processing chain for each polarization. Each of the next sections discusses one step in the path of the raw information bits towards the PolSK waveform at the transmitter's output interface.

B.2.1. Symbol selection and pulse shaping

Random data from a linear feedback shift register (LFSR) [41] is used as a data source. It creates a repeating bit sequence with noise-like properties (i.e. the sequence's autocorrelation function is zero except in the origin). The 32 bit long shift register is defined by the following polynomial:

$$q(x) = x^{32} + x^{22} + x^2 + x + 1 \tag{B.3}$$

This polynomial leads to a maximum-length sequence [41], repeating only after $2^{32} - 1$ bits.

The symbol clock is fixed to $f_{\text{symbol}} = 1{,}024\,\text{MHz}$ such as to be an integer divisor of the DAC's sample clock. The LFSR runs four times faster than the symbol clock and every four bits are merged into a into a data word from 0 to 15. Consequently, the LFSR sequence repeats after 17 minutes. The 4-bit data words address a lookup table containing the precomputed in-phase and quadrature components of the Stokes-vectors. The symbol table changes after every symbol, such as to alternate the *fixed* carrier between horizontal and vertical polarization; X_{Q} will be zero for even numbered symbols and Y_{Q} zero for odd symbols.

The resulting four baseband components are upconverted by the factor ×30 and filtered using a root-raised cosine FIR filter with a group delay of $\tau_g = 5T$. This leads to a filter with 300 taps. It is used in time-multiplex to pulse-shape all four baseband components.

B.2.2. 8$^{\text{th}}$-band filter

The block diagram in Figure B.1 depicts a second upsampling and filtering stage. Its purpose is to bring the signal up to the DAC's sample rate. This function is

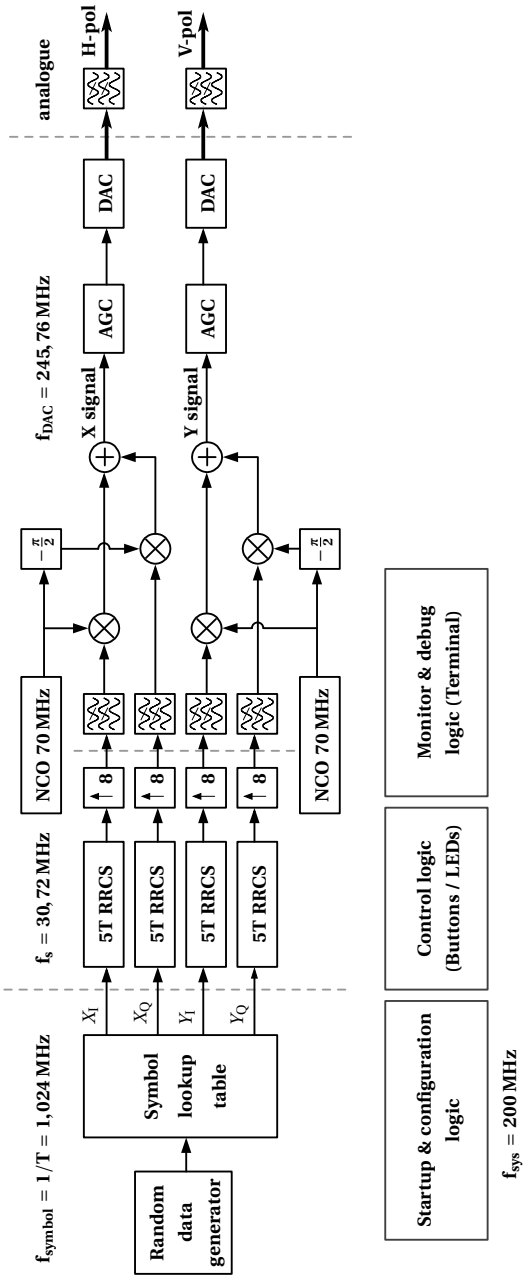


Figure B.1.: Simplified transmitter block diagram

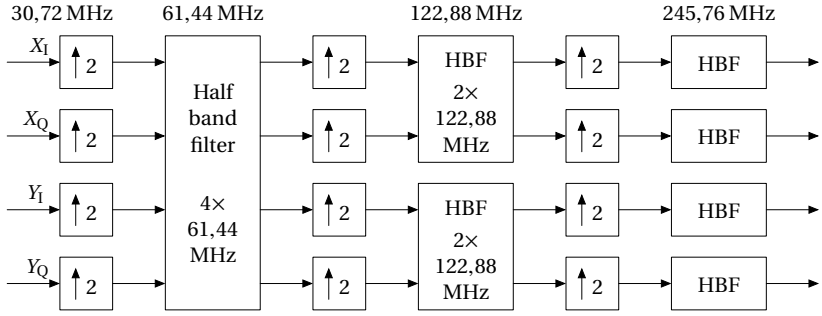


Figure B.2.: Filter structure of eighth-band filter. All filter blocks run at $f_{\max}$. For the first and second stages the data rate is slower, so multipliers can be reused in time-multiplex.

realized through an eighth-band filter. This very efficient filter structure consists of three cascaded stages of successive ×2 upsampling and half-band filtering. The filter's inner structure is displayed in Figure B.2.

Digital half-band filters [87] are an efficient tool in digital signal processing. They are FIR filters with a symmetrical frequency response around $f_s/4$, thereby filtering away *half* of the band. Half of the coefficients are zero and do not require explicit multiplication. Used in an upsampling battery as shown in Figure B.2, the first stages in the cascade can run on a slower sampling rate and be used in time-multiplex. In this way, even more multipliers are saved. From this structure, a resource-efficient FIR filter with excellent rejection and transition properties can be created.

The filter shown in Figure B.2 uses 96 multipliers (First stage: 1×52, second stage: 2×10, third stage: 4×6). A classic FIR filter structure with equivalent frequency response would require $4 \times 391 = 1564$ multipliers, a prohibitive amount.

B.2.3. Upconversion and output to hardware interface

With the sample rate adapted to the DAC frequency, the signal can be upconverted into the target 70 MHz band. Numerically controlled oscillators provide the required carrier which is multiplied onto the baseband signal.

Finally an automatic gain control (AGC) unit adjusts the average signal level according to a pre-specified target. It ensures that the DAC's dynamic range is optimally used while avoiding clipping errors. The AGC has a time-constant of $\tau = 8{,}3$ ms. Its detailed description is put back to the receiver section.

The final signal is fed to the board's digital-to-analogue converter interface.

The DAC's impulse response is rectangular and will lead to spectral spurs. The development board provides analogue low-pass filtering to suppress these spurs and faithfully reconstruct the original signal. A spectral plot of the final transmitter output on the X-polarization is displayed in Figure 3.4 in Chapter 3.

B.2.4. Startup sequence

Control and monitoring of the evaluation board and transmitter are carried out by a finite state machine (FSM). The block diagram in Figure B.3 briefly depicts the FSM's main functionality.

After start-up, the different peripheral hardware components are configured. These include the master clock chip, the DAC and ADC themselves and finally a small monitor chip watching the voltage levels on the expansion card.

After the subsystems' configuration, a test sequence is run on the DAC chip to verify proper connectivity. Upon successful completion, the transmitter and the DAC chip are activated and the state machine passes into its operational state. Different error conditions, reported from the DAC or monitor chip, but also the loss of any clock signals, trigger error sequences that require a system reset. The current FSM state or error condition can be read from an LED array on the development board.

Internal register states are communicated to a workstation using the JTAG-to-USB interface. This is used for debugging and online monitoring.

B.3. The PolSK receiver

The receiver's objective is to detect the original message from the captured signal. To this end it inverts one after one the processing steps in the transmitter. The signal is downconverted back into baseband, filtered and downsampled to the symbol clock. As essential quantities like the carrier frequencies and phases or the symbol clock are unknown to the receiver, it has the additional task to estimate those beforehand. The receiver's structure is depicted in Figure B.4.

B.3.1. Front-end

The receiver's front-end is the analogue-to-digital converter. The converter chip has an integrated amplifier, which is controlled by an AGC in the FPGA. It adjusts the amplifier gain in discrete steps to provide a strong, but undistorted signal. To avoid an unnecessary high number of adjustments, it works with a hysteresis and a time constant of 17 ms. It can optionally be disabled to provide a more reproducible reception environment.

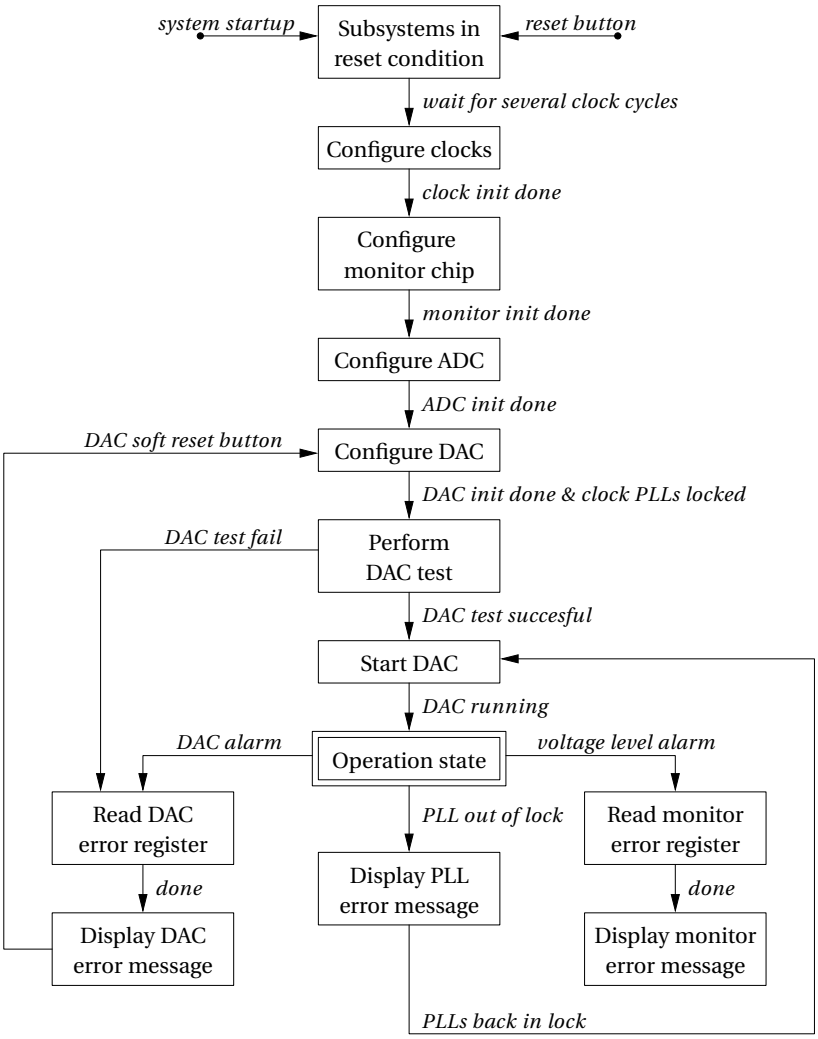


Figure B.3.: Main Finite State Machine Overview

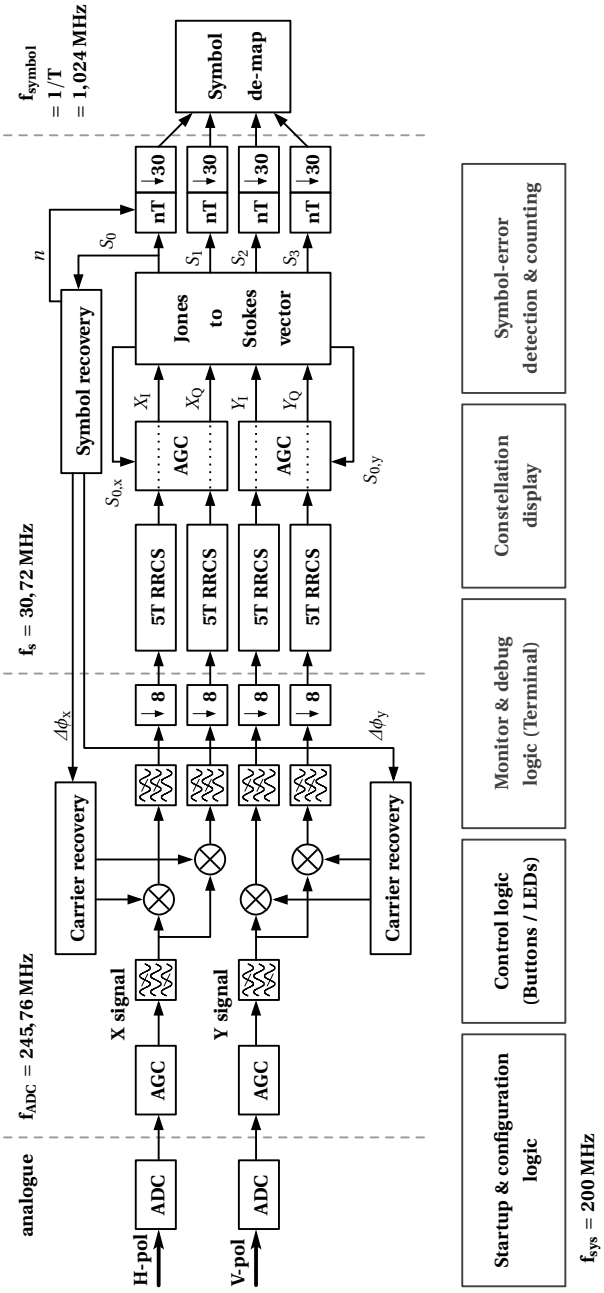


Figure B.4.: Simplified receiver block diagram

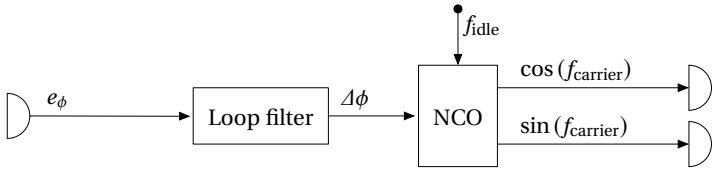


Figure B.5.: Block diagram of the carrier recovery system

A bandpass filter with $f_{\text{pass}} = 70 \pm 15$ MHz filters away out-of-band noise and conditions the data stream for downconversion and downsampling. The filter's frequency response is the same as that of the eighth-band filter in the transmitter, centred around 70 MHz. The advantageous implementation properties cannot be maintained to the full extent however.

B.3.2. Carrier recovery

According to the textbook example in Figure 2.6, the filtered signal is downconverted into baseband using the estimated cosine and sine signals from a carrier recovery unit.

The carrier recovery uses the quadrature samples at the symbol strobes as an estimate for the frequency error, knowing that the signal's phase is supposed to be zero. The carrier recovery system itself is a simple second-order type 2 phase-locked loop. The PLL's feedback path is shown in the block diagram in Figure B.5. The error samples are filtered and passed to the NCO. The NCO has a constant phase-increment for the centre LO frequency, which is adjusted using the filtered phase error samples.

The loop filter is a PI (proportional + integrating) control filter with a loop bandwidth set to 10 kHz. Its structure, based on the design recommendations in [33], is shown in the block diagram in Figure B.6. In addition to Gardner's book [39], the articles [56] and [18] have shaped the loop designs.

The error samples are provided with a frequency of $f_{\text{symbol}}/2$. As the RRCS filter is placed inside the loop (cf. Figure B.4), a delay of roughly $6T$ is accumulated between phase adjustment and measurement. The loop's time-constant is one order of magnitude below this delay, which should in consequence stay without impact [98].

B.3.3. Receive filtering

Receive filtering of the downconverted PolSK signal is achieved in two steps. First, a short (11 taps) FIR filter rejects the spectral image appearing at $f = 105{,}76$ MHz.

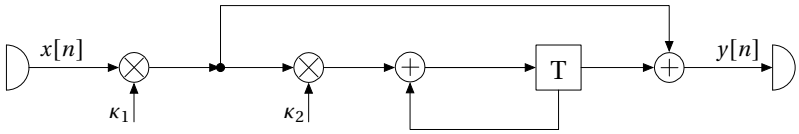


Figure B.6.: Carrier recovery PI loop filter

The remaining signal is then effectively band-limited and can safely be downsampled by the factor 8. The downsampled signal components are then fed to a root-raised cosine filter equal to the one in the transmitter. Thereby the matched filter condition is fulfilled.

B.3.4. Back-end automatic gain control

Depending on the amount of noise that was initially present in the received signal, the power level can greatly vary at this point in the reception chain. Therefore, a back-end AGC adjusts the signal level to a predetermined value ensuring a good dynamic range. The fixed signal level is also important to guarantee a constant loop gain in the carrier and symbol tracking PLLs.

The AGC gain is applied by multiplying the signal with an adjustable factor. As hinted in [7], the squared signal appearing as a by-product of the Stokes-vector conversion is used as a measure for the instantaneous power. Using S_0 to control AGC gain leads to a so-called *square-law* detector [65]. An error signal is created by subtracting the estimated power from the target value. This error signal is then smoothed out by a two-stage filter, resulting in a combined lowpass filter bandwidth of 120 Hz. The filter consists in a first CIC stage [44, 92] and a second, single-pol IIR filter. This filter is used in all AGC implementations for the present work, with varying time-constants. The filtered error signal is finally used to adjust the AGC-gain in the feedback loop. The following equations illustrate this process qualitatively:

$$y[n] = (g_{\mathrm{AGC}}[n]x[n])^2 \tag{B.4}$$

$$g_{\mathrm{AGC}}[n+1] = g_{\mathrm{AGC}}[n] + \beta \cdot \mathcal{LP}\left\{y_{\mathrm{target}} - y[n]\right\} \tag{B.5}$$

g_{AGC} is the AGC gain, the operator $\mathcal{LP}\{\cdot\}$ denotes low-pass filtering. The coefficient β is chosen small enough to obtain stable loop behaviour. The articles [65] and [84] have shaped the AGC design.

B.3.5. Jones-to-Stokes-vector transformation

The adjusted signal components in in-phase and quadrature representation are transformed into Stokes-vectors using the formula from equation (2.4).

All further processing is now done on the Stokes-parameters. In addition to the state of polarization, they serve as a measure for instantaneous power. The following intermediate results for computing S_0 are forwarded to the back-end AGCs, serving as $y[n]$ in equation (B.4) to compute the gain adjustment:

$$S_{0,y} = X_\mathrm{I}^2 + X_\mathrm{Q}^2 \tag{B.6}$$

$$S_{0,y} = Y_\mathrm{I}^2 + Y_\mathrm{Q}^2 \tag{B.7}$$

In addition, the S_0-parameter is fed to the symbol-recovery subsystem as a timing wave.

B.3.6. Symbol-clock recovery

The symbol-clock recovery has two functions. It indicates the best downsampling instant for symbol extraction and provides a divided symbol clock to sample the quadrature values X_Q and Y_Q.

The square timing-error detector from Section 2.5.1 is used. Although this technique is recommended for analogue detection, the high oversampling ratio in the receiver (×30) allows to use it in the present case[1]. In the PolSK signal, a large amount of energy goes into two discrete spectral components at ± half the symbol frequency around the carrier (cf. Section 3.1.5 and Figure 3.4). Upon squaring, these components amplify the detection capability by adding up to the envelope variations from the pulse-shape. In this way, the square timing-error detector has very good performance even in the case of relatively small roll-off values.

Functional description

A block diagram of the symbol timing recovery is shown in Figure B.7. It consists of three functional parts: the PLL, extracting and tracking the symbol clock from the S_0-parameter, a sample counter providing a sample number to be used for downsampling and finally two registers to divide the symbol clock by two.

The PLL uses a simple digital mixing principle by flipping the sign of the incoming S_0-parameter according to the sign of the binary clock signal (0 or 1). The loop filter smooths the down-mixed signal and filters away spurs. The PLL uses a type 2 PI filter, cascaded with a first-order IIR filter for additional low-pass

[1] A highly oversampled digital signal behaves a little bit like an analogue signal.

performance. This is necessary to remove the perturbing influence of the carrier residual in S_0. The filter block diagram along with its discrete-time frequency response is displayed in Figure B.8. The loop bandwidth is 100 Hz. Following the filter, the phase error $\Delta\Omega_i$ is added to the expected phase increment Ω_{idle}, corresponding to $T = 1/1{,}024$ MHz, and integrated in the phase accumulator. The most significant bit (MSB or sign bit) of the accumulated phase is then used as a clock signal.

An external sample-counter provides a common reference to the symbol-clock recovery and the downsampling block. It counts repeatedly from 0 to 29, to identify each sample during a symbol period. A rising edge on the symbol clock will save the current value on the sample counter and forward it to the downsampling unit. For instance, if the counter shows $n = 17$ when the symbol clock rises, the downsampling unit will choose sample #17 for downsampling and discard all other 29 samples.

Finally, a flip-flop is used to divide the symbol frequency by two. Thereby two clocks of opposing sign are created at half the symbol frequency. A rising edge on one of the divided clocks will save the current value of either X_{Q} or Y_{Q}. This value is then forwarded to the respective carrier recovery unit, as an estimate of the current phase error. During signal acquisition, it is not known which of these clocks belongs to which polarization. Therefore the routing is flipped until the carrier recovery can successfully lock.

B.3.7. Downconversion and de-mapping

At this stage, the stream of Stokes-vectors is oversampled by 30. To extract the transmitted symbols, the Stokes-parameters are downsampled to the symbol frequency at the instant of minimum inter-symbol interference (ISI) (or maximum eye-opening). The symbol recovery unit designates which of the samples is to be kept. This simple downsampling leads to a precision of $\pm T/60$, which might be a little crude, but should deliver working results[2].

Finally, symbols are de-mapped using the maximum likelihood criterion described in [5]. It picks the constellation point which was most likely transmitted when observing the received Stokes-vector. This is achieved by performing an inner product of the received Stokes-vector with all candidate vectors and choosing the candidate yielding the largest outcome.

[2] If higher precision is required, one needs to resort to interpolation [38, 26].

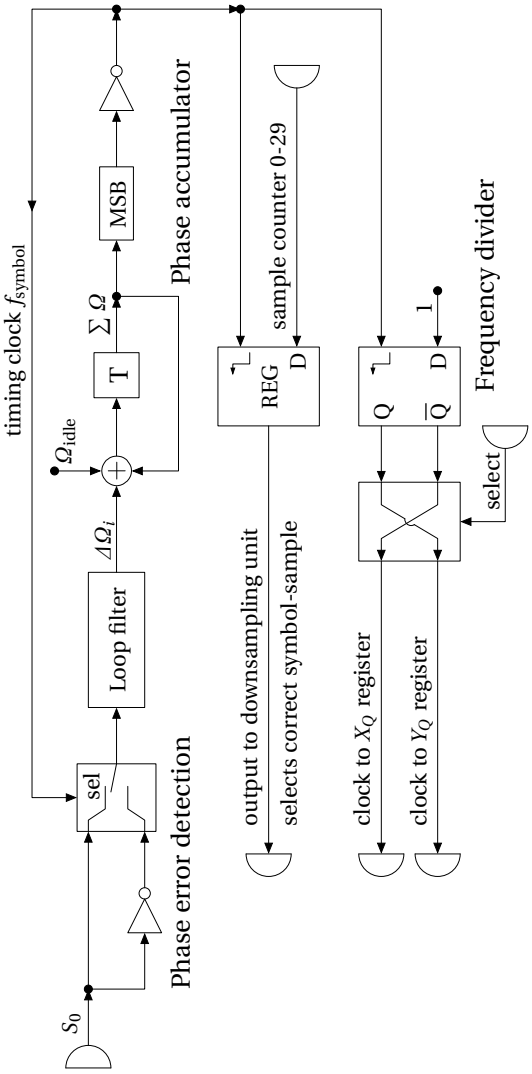


Figure B.7.: Block diagram of the timing recovery system

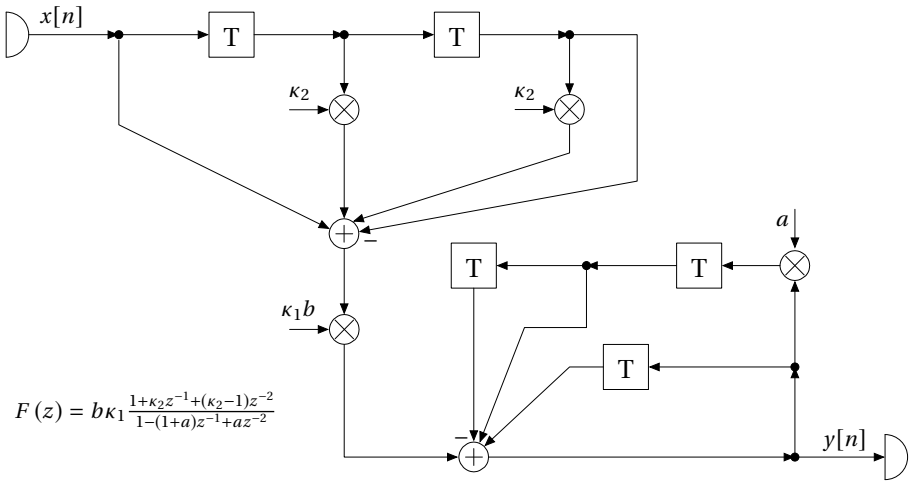


Figure B.8.: Loop filter used for symbol recovery

B.3.8. Additional functionality

The receiver contains the same linear feedback shift register as the transmitter. It can be synchronized by feeding it with a sequence of correctly detected symbols [14]. After synchronization, the register is able to predict all future incoming symbol and it is able to detect wrong symbol decisions. By counting the total number of symbols and the number of erroneous symbols, the symbol-error rate is measured.

The transmitter control FSM is reused for the receiver. Whereas the transmitter simply disables the ADC chip, the receiver contains additional logic for its configuration and to initiate the data transfers.

Improved debugging facilities are available in the receiver. The JTAG interface provides online information on the loop states, lock indicators, AGC coefficients, SER, et cetera. The received Stokes-vectors and I/Q values are transmitted to a workstation using a serial port and can be viewed life during reception.

B.4. Bi-orthogonal signalling

The demonstrator for the four-dimensional, bi-orthogonal signal is heavily based on the PolSK modem. All the hardware interfaces and many essential building blocks have been reused.

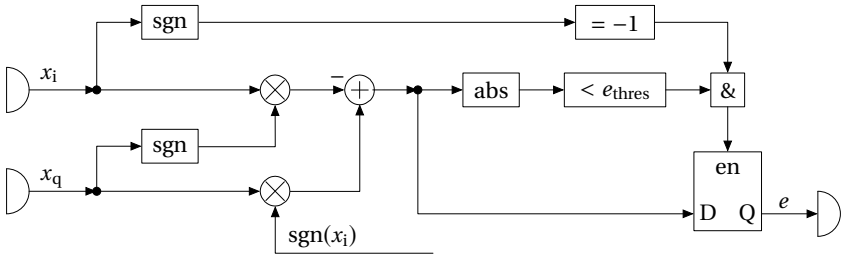


Figure B.9.: Sari-Moridi QPSK phase and frequency error detector

B.4.1. The transmitter

Only few modifications have been necessary to transform the PolSK into a bi-orthogonal transmitter. The symbol table was exchanged to contain the eight bi-orthogonal symbols, instead of the sixteen PolSK symbols. The linear feedback shift register used to pick the symbol value was modified to output only 3 bits at a time. To adapt the transmission format to a more contemporary form, the RRCS roll-off was reduced to 0,20.

B.4.2. The receiver

The receiver kept the same structure as well, except for the special PolSK carrier recovery system that could no longer be used. As the bi-orthogonal waveform looks very similar to QPSK, it was replaced by the Sari-Moridi carrier recovery for QPSK, introduced in Section 2.5.1. The concrete implementation of the phase and frequency detector is depicted in Figure B.9.

x_i and x_q are the sampled symbol values of the baseband quadratures, e is the estimated phase error. To increase the frequency locking capability, new phase-error values are only updated when they do not exceed the threshold of $\pi/7$. The detector's output is then low-pass filtered and controls a numerically controlled oscillator to downconvert the incoming signal. This is how the feedback loop is closed.

In QPSK carrier recovery without pilot tones, a certain phase ambiguity cannot be avoided, as the symbol looks identical when rotated by $\pi/2$ radians. This ambiguity is resolved in the receiver by adjusting the local oscillator phases in $\pi/2$ steps, until the error-counting shift register can properly synchronize. Maximum sixteen different states have to be checked (four for each polarization).

Besides this major change, the de-mapping routine and some thresholds were adapted. The square timing-recovery implemented for the PolSK modem worked surprisingly well in first tests, so that it was not replaced.

C. Unusual symbol constellations

Two rather exotic constellation designs were introduced to function as appropriate single-polarization references to some of the four-dimensional constellations. Both are presented here by giving the exact symbol coordinates, showing a constellation diagram and the formulas used for the computation of the symbol-error probability. The SER formulas have been validated by simulations.

C.1. The 8-point hexagonal constellation

The first constellation is the hexagonal 8-QAM, displayed in Figure C.1. The hexagonal structure is more dense then a plain rectangular base. Its overall shape is rectangular as well, but the side ratio is smaller than for a rectangular 8-QAM constellation. This increases the power efficiency beyond that achieved by the lattice choice alone.

The coordinates of the eight constellation points in complex notation are listed below:

$$\frac{d}{2} \cdot \left\{(-1),\ (1),\ \left(-\mathrm{j}\sqrt{3}\right),\ \left(\mathrm{j}\sqrt{3}\right),\right.$$
$$\left.\left(-2-\mathrm{j}\sqrt{3}\right),\ \left(-2+\mathrm{j}\sqrt{3}\right),\ \left(2-\mathrm{j}\sqrt{3}\right),\ \left(2+\mathrm{j}\sqrt{3}\right)\right\} \quad \text{(C.1)}$$

The coordinates are expressed as a function of the symbol distance d. The average symbol energy is:

$$E_\mathrm{s} = \frac{1}{2}\left(\frac{9}{8}d^2\right) \quad \text{(C.2)}$$

C.1.1. Computation of error probability

To compute the error probability of hexagonal 8-QAM, the I/Q constellation diagram is divided into 8 decision regions, as shown in Figure C.1. The regions are of 3 different shapes, denoted by A, B and C.

Rather than the error probability, the probability of a correct decision P_c is computed. For each symbol, selected with equal probability $P = 1/8$, a correct decision is made when the noisy signal lands within its decision region. The total probability for a correct decision is:

$$P_\mathrm{c} = \frac{1}{8}\left(4P_\mathrm{A} + 2P_\mathrm{B} + 2P_\mathrm{C}\right) \quad \text{(C.3)}$$

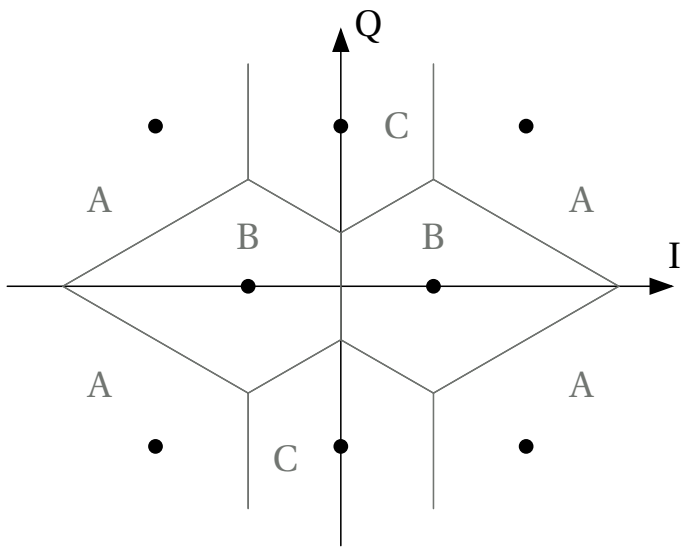


Figure C.1.: Hexagonal 8-QAM constellation with decision regions

For all three regions A, B and C, the integration of the noise probability density function is performed with the symbol at the origin. The corresponding integrand is:

$$f_{X,Y}(x,y) = \frac{1}{\pi\sigma^2} e^{-\frac{1}{\sigma^2}(x^2+y^2)} \tag{C.4}$$

with the complex noise variance $\sigma^2 = 2N_0$. Using symmetries in the decision regions, the probabilities can be calculated as follows:

$$\begin{aligned} P_\mathrm{A} = &\int_{-d/2}^{d/2} \int_{\frac{-\sqrt{3}}{3}(x+d)}^{\frac{-\sqrt{3}}{6}d} f_{X,Y}(x,y)\,dy\,dx + \int_{-d/2}^{d/2} \int_{\frac{-\sqrt{3}}{6}d}^{+\infty} f_{X,Y}(x,y)\,dy\,dx \\ &+ \int_{d/2}^{+\infty} \int_{\frac{-\sqrt{3}}{2}d}^{+\infty} f_{X,Y}(x,y)\,dy\,dx \end{aligned} \tag{C.5}$$

$$P_\mathrm{B} = 4 \int_{0}^{d/2} \int_{0}^{\frac{\sqrt{3}}{3}(d-x)} f_{X,Y}(x,y)\,dy\,dx + 2 \int_{-d}^{-d/2} \int_{0}^{\frac{\sqrt{3}}{3}(d+x)} f_{X,Y}(x,y)\,dy\,dx \tag{C.6}$$

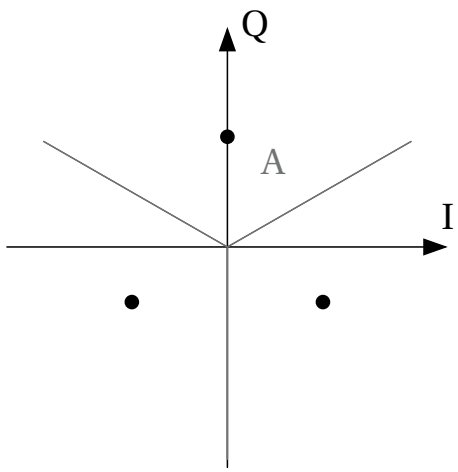


Figure C.2.: 3-point PSK constellation with decision regions

$$P_{\mathrm{C}} = 2 \int_{0}^{d/2} \int_{\frac{\sqrt{3}}{3}(x-d)}^{\frac{-\sqrt{3}}{6}d} f_{X,Y}(x,y)\,dy\,dx + 2 \int_{0}^{d/2} \int_{\frac{-\sqrt{3}}{6}d}^{+\infty} f_{X,Y}(x,y)\,dy\,dx \tag{C.7}$$

The integral in equation (C.4) can be factorized. For the numerical computation, the expressions are split into integrals over x and y and evaluated with the error function. This way, numerical integration over infinite boundaries is avoided. The integrals containing a dependency on x in the boundaries are evaluated using the `Matlab` function `quad2d`. Its algorithm is described in [71].

C.2. The 3-point PSK constellation

The second custom constellation is a 3-point PSK, with symbols placed on the vertices of an equilateral triangle. It is used as a two-dimensional modulation type with a spectral efficiency close to that of the four-dimensional bi-orthogonal signal. The constellation is shown in Figure C.2, together with the decision region boundaries. The coordinates are expressed directly in dependence of the symbol energy E_{s}:

$$\sqrt{2E_{\mathrm{s}}} \cdot \left\{ (\mathrm{j}),\ \left(\tfrac{-\sqrt{3}-\mathrm{j}}{2}\right),\ \left(\tfrac{\sqrt{3}-\mathrm{j}}{2}\right) \right\} \tag{C.8}$$

C.2.1. Computation of error probability

As the decision regions are congruent, the error probability can be computed from the symbol in the upper half-plane. The noise probability density function is integrated over the right half of this decision region, illustrated by the letter A in Figure C.2. The symbol-error probability is then:

$$P_\mathrm{e} = 1 - 2P_\mathrm{A} \tag{C.9}$$

with

$$P_\mathrm{A} = \int_0^{+\infty} \int_{x/\sqrt{3}}^{+\infty} f_{X,Y}(x, y - \sqrt{2E_\mathrm{s}})\, dy\, dx \tag{C.10}$$

$$= \frac{1}{2\sigma\sqrt{\pi}} \int_0^{+\infty} e^{\frac{-x^2}{\sigma^2}} \cdot \operatorname{erfc}\left[\frac{x}{\sigma\sqrt{3}} - \sqrt{\frac{2E_\mathrm{s}}{\sigma^2}}\right] dx \tag{C.11}$$

The final integral is computed with the `Matlab` function `quadgk` [72], which is able to handle an infinite boundary.

D. Notation and acronyms

D.1. Mathematical notation

Scalars are in normal script, while vectorial quantities are written bold:

x : Scalar

$\mathbf{x}$: Algebraic vector

$\vec{x}$: Vector with an explicitly spatial interpretation

For vectors, individual components are denoted by subscripts. Usually a subscripted x denotes the first, y the second component and so forth. Discrete-time indices are indicated using square brackets $[\cdot]$ and sometimes, for convenience, with a subscript. Continuous-time is indicated with round braces $(\cdot)$.

Various operations on a complex scalar z or vector $\mathbf{z}$:

$\Re\mathfrak{e}\{z\}$: Real part

$\Im\mathfrak{m}\{z\}$: Imaginary part

$|z|$: Absolute value (magnitude)

z^* : Complex conjugated

$\mathbf{z}^{\mathrm{T}}$: Transposed

$\mathbf{z}^{\mathrm{H}}$: Hermitian $= (\cdot)^{\mathrm{T}}$ and $(\cdot)^*$

Statistical operators on a random variable A (and Y):

$p(a)$: Probability of the event $A = a$

$p_\lambda(a)$: Probability of the event $A = a$ for given λ

$\mathrm{E}[A]$: Expected value

$\mathrm{E}_A[x(A)]$: Expected value of $x(A)$ over A

$\mathrm{Var}[A]$: Variance

$\mathrm{Cov}[A; Y]$: Covariance

The operator for finding the argument a that maximizes an expression $Q(a)$:

$$a_{\max} = \arg\max_a Q(a)$$

Operations on a (discrete or continuous) time function s[i] or s(t):

$\mathcal{LP}\{s\}$: Low-pass filter
$\mathcal{F}\{s\}$: Fourier transform

Special functions:

$\mathrm{erfc}(x)$: Error function
$\delta(x)$: Dirac delta-function
$\delta[x]$: Unit pulse
$\mathrm{sgn}(x)$: Sign extraction

D.2. Symbols

Special symbols:

j : Imaginary unit
e : Base of natural logarithm

Symbol nomenclature:

$\vec{E}$: Electric field
ω : Angular frequency
$\Delta\omega$: Angular frequency offset
$\vec{k}$: Wave number
A : Amplitude
ϕ : Phase
ϕ_0 : Phase at the time origin
τ : Time offset
Δt : Delay difference

$\vec{J}$: Jones-vector
X_I : In-phase component of X
X_Q : Quadrature component of X
$\vec{S}$: Stokes-vector
δ : Phase difference between polarizations

c_i : Complex scalar symbol with time index i
$\mathbf{c}_\mathrm{i}$: Complex vector symbol with time index i
$\mathbf{c}$: Vector containing a certain number of symbols
$s(t)$: Modulated signal on one polarization
$\mathbf{s}(t)$: Dual-polarization modulated signal
ω_c : Angular carrier frequency
$g(t)$: Pulse-shape
α : RRCS roll-off value
T : Duration of a symbol
f_s : Sampling frequency
P_s : Power of signal s
T_m : Measurement duration
$n(t)$: Continuous-time noise process
$r(t)$: Noisy copy of $s(t)$

E_s : Energy per symbol
E_s' : Energy per symbol per polarization
E_b : Energy per bit
b : Bits per symbol
N_0 : Single-sided noise power spectral density
B : Signal bandwidth

m_0 : Mean amplitude of PolSK constellation symbols
E_R' : Helper variable for PolSK energy allocation
$\overline{m_c}$: Average mean value of sequence c
R_{cc} : Autocorrelation function of sequence c
$\overline{R_{cc}}$: Average autocorrelation function of sequence c
$\overline{S_{cc}}$: Average power spectral density of sequence c

$z(t)$: Timing wave
s_{bb} : Baseband component of a signal s
$\Delta\epsilon$: Loop error signal
κ_{p} : Phase detector gain
κ_{v} : NCO scaling coefficient
κ : Loop gain
B_{N} : Loop noise bandwidth
T_0 : Observation length
ν : Loop damping coefficient

$\hat{\tau}$: Average number of nearest neighbours
d_{min} : Distance between nearest neighbours
σ^2 : Variance of white Gaussian random process
$\mathbf{u}$: Vector of signal parameters not to be estimated
ξ : Pulse-shape dependency in the MCRB

Λ : Likelihood function
λ : Signal parameter to be estimated
$\mathbf{u}$: Vector of signal parameters not to be estimated
ξ : Pulse-shape dependency in the MCRB

D.3. List of acronyms

2-D Two-dimensional

4-D Four-dimensional

ADC Analogue-to-digital conversion (also A/D)

AGC Automatic gain control

AM Amplitude modulation

APSK Amplitude and phase shift keying

AWGN Additive white Gaussian noise

BPSK Binary phase shift keying

CIC Cascaded integrator-comb (filter)

DAC Digital-to-analogue conversion (also D/A)

D/C Down-conversion

DDR Double data rate

DFT Discrete Fourier-transform

DPSK Differential phase shift keying

DQPSK Differential quaternary phase shift keying

DSP Digital signal processor (also: processing)

DVB-S(2) Digital Video Broadcasting over satellite (second generation)

EIRP Equivalent isotropically radiated power

FIR Finite impulse response (filter)

FMC FPGA mezzanine card

FPGA Field programmable gate array

FSM Finite state machine

GED Gardner timing-error detector

HBF Half-band filter

HPA High-power amplifier

I2C Inter-integrated circuit (bus)

IF Intermediate frequency

IIR Infinite impulse response (filter)

IMUX Input multiplexing (filter)

I/O Input/output

I/Q In-phase/quadrature

ISI Inter-symbol interference

ITU International telecommunication union

JTAG Joint test action group

LAM Lattice amplitude modulation

LED Light-emitting diode

LFSR Linear feedback shift register

LNA Low-noise amplifier

LNB Low-noise block converter

LO Local oscillator

LPF Low-pass filter

MCRB Modified Cramer-Rao bound

MSB Most significant bit

NCO Numerically controlled oscillator

OBO Output back-off

OMUX Output multiplexing (filter)

PAPR Peak-to-average power ratio

PDM Polarization division multiplexing

PFD Phase and frequency detector

PLL Phase-locked loop

PM Phase modulation

PN Phase noise

PolSK Polarization Shift Keying

PSK Phase shift keying

QAM Quadrature amplitude modulation

QPSK Quaternary phase shift keying

RF Radio-frequency

RRCS Root-raised cosine

SER Symbol-error rate

SES Société Européenne des Satellites

SNR Signal-to-noise ratio

SoP State of polarization

SPI Serial peripheral interface (bus)

TEM Transversal electromagnetic (wave)

TPE Transponder equivalent bandwidth

TWT(A) Travelling wave tube (amplifier)

UB Union bound

U/C Up-conversion

USB Universal serial bus

VCO Voltage-controlled oscillator

VHDL Very high-speed integrated circuit hardware description language

Bibliography

[1] 4DSP. *FMC150 User Manual*. Product data sheet. Version r1.7. 2012. URL: support@4dsp.com.

[2] Erik Agrell and Magnus Karlsson. "Power-Efficient Modulation Formats in Coherent Transmission Systems". In: *Journal of Lightwave Technology* 27 (22 2009), pp. 5115–5126. ISSN: 0733-8724. DOI: 10.1109/JLT.2009.2029064.

[3] Philippe C. Becker, N. Anders Olsson and Jay R. Simpson. *Erbium-doped fiber amplifiers: Fundamentals and technology*. Optics and photonics. San Diego: Academic Press, 1999. ISBN: 0120845903.

[4] Sergio Benedetto, Roberto Gaudino and Pierluigi Poggiolini. "Direct detection of optical digital transmission based on polarization shift keying modulation". In: *IEEE Journal on Selected Areas in Communications* 13.3 (1995), pp. 531–542. ISSN: 0733-8716. DOI: 10.1109/49.372412.

[5] Sergio Benedetto and Pierluigi Poggiolini. "Multilevel Polarization Shift Keying: Optimum receiver structure and performance evaluation". In: *IEEE Transactions on Communications* 42 (2/3/4 1994), pp. 1174–1186.

[6] Sergio Benedetto and Pierluigi Poggiolini. "Theory of Polarization Shift Keying modulation". In: *IEEE Transactions on Communications* 40 (4 1992), pp. 708–721.

[7] Silvello Betti, Giancarlo De Marchis and Eugenio Iannone. "Polarization modulated direct detection optical transmission systems". In: *Journal of Lightwave Technology* 10 (12 1992), pp. 1985–1997. DOI: 10.1109/50.202809.

[8] Silvello Betti et al. "A novel multilevel coherent optical system: 4-Quadrature signaling". In: *Journal of Lightwave Technology* 9 (4 1991), pp. 514–523. DOI: 10.1109/50.76666.

[9] Silvello Betti et al. "Multilevel coherent optical system based on Stokes parameters modulation". In: *Journal of Lightwave Technology* 8 (7 1990), pp. 1127–1136. ISSN: 0733-8724. DOI: 10.1109/50.56417.

[10] S. Betti et al. “State of polarisation and phase noise independent coherent optical transmission system based on Stokes parameter detection”. In: *Electronics Letters* 24 (23 1988), pp. 1460–1461. ISSN: 0013-5194.

[11] Lotfollah Beygi et al. “Coded Modulation for Fiber-Optic Networks. Toward better tradeoff between signal processing complexity and optical transport reach”. In: *IEEE Signal Processing Magazine* 31 (2 2014), pp. 93–103.

[12] Erwin Biebl. *Optische Übertragungstechnik*. Vorlesungsunterlagen. München: Technische Universität München, 2006.

[13] Stephen D. Brown and Zvonko G. Vranesic. *Fundamentals of digital logic with VHDL design*. Boston: McGraw-Hill, 2000. 1 computer laser optical disc (4 3/4 in.) ISBN: 0070125910.

[14] Luis Miguel Brugarolas. “Simple BER meter is easy to build”. In: *EDN Design Ideas* (March, 2 2000), pp. 115–116. URL: www.edn.com.

[15] R. Calvani, R. Caponi and F. Cisternino. “Polarisation phase-shift keying: a coherent transmission technique with differential heterodyne detection”. In: *Electronics Letters* 24 (10 1988), pp. 642–643. ISSN: 0013-5194.

[16] Enzo A. Candreva, Giovanni E. Corazza and Alessandro Vanelli-Coralli. “On the Optimization of Signal Constellations for Satellite Channels”. In: *International Workshop on Satellite and Space Communications (IWSSC)*. 2007, pp. 17–21. DOI: 10.1109/IWSSC.2007.4409383.

[17] F. Cardells-Tormo et al. “Design of a DVB-S receiver in FPGA”. In: *IEEE Workshop on Signal Processing Systems (SIPS)*. 2003, pp. 6–11. ISBN: 1520-6130. DOI: 10.1109/SIPS.2003.1235635.

[18] Alfonso Carlosena and Antonio Mànuel-Lázaro. “Design of High-Order Phase-Lock Loops”. In: *IEEE Transactions on Circuits and Systems II: Express Briefs* 54 (1 2007), pp. 9–13. ISSN: 1549-7747. DOI: 10.1109/TCSII.2006.883205.

[19] John H. Conway and Neil J. A. Sloane. *Sphere packings, lattices, and groups*. 3rd ed. Vol. 290. Grundlehren der mathematischen Wissenschaften. New York: Springer, 1999. 703 pp. ISBN: 0387985859.

[20] John Philip Dakin. “Improvements relating to optical fibre communication systems”. United Kingdom patent. GB2153176A. 1985.

[21] A. N. D'Andrea and Marco Luise. "Optimization of symbol timing recovery for QAM data demodulators". In: *IEEE Transactions on Communications* 44 (3 1996), pp. 399–406. ISSN: 0090-6778. DOI: 10.1109/26.486334.

[22] Aldo N. D'Andrea, Umberto Mengali and Ruggero Reggiannini. "The modified Cramer-Rao bound and its application to synchronization problems". In: *IEEE Transactions on Communications* 42 (234 1994), pp. 1391–1399. ISSN: 0090-6778. DOI: 10.1109/TCOMM.1994.580247.

[23] E. Dietrich et al. "Heterodyne transmission of a 560 Mbit/s optical signal by means of polarisation shift keying". In: *Electronics Letters* 23 (8 1987), pp. 421–422. ISSN: 0013-5194. DOI: 10.1049/el:19870305.

[24] DIMITRI PFEIFFER. "Polarization as modulation dimension". Département d'ingénierie mathématique. Mém.de mast. Louvain-la-Neuve : Université catholique de Louvain, 2012.

[25] Ivan B. Djordjevic, Lilorad Cvijevic and Changyu Lin. "Multidimensional Signaling and Coding Enabling Multi-Tb/s Optical Transport and Networking. Multidimensional aspects of coded modulation". In: *IEEE Signal Processing Magazine* 31 (2 2014), pp. 104–117.

[26] Lars Erup, Floyd. M. Gardner and Robert A. Harris. "Interpolation in digital modems. II. Implementation and performance". In: *IEEE Transactions on Communications* 41 (6 1993), pp. 998–1008. ISSN: 0090-6778. DOI: 10.1109/26.231921.

[27] European Telecommunications Standards Institute, ed. *Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for 11/12 GHz satellite services*. EN 300 421. Aug. 1997.

[28] European Telecommunications Standards Institute, ed. *Digital Video Broadcasting (DVB); Second generation framing structure, channel coding and modulation systems for Broadcasting, Interactive Services, News Gathering and other broadband satellite applications (DVB-S2)*. EN 302 307. Aug. 2009.

[29] Goulven Eynard et al. "Efficient phase noise modeling of a PLL-based frequency synthesizer". In: *4th International Symposium on Communications, Control and Signal Processing (ISCCSP)*. 2010, pp. 1–4. DOI: 10.1109/ISCCSP.2010.5463478.

[30] G. David Forney Jr. "Coset codes. I. Introduction and geometrical classification". In: *IEEE Transactions on Information Theory* 34 (5 1988), pp. 1123–1151. ISSN: 0018-9448. DOI: 10.1109/18.21245.

[31] G. David Forney Jr. and G. Ungerboeck. “Modulation and coding for linear Gaussian channels”. In: *IEEE Transactions on Information Theory* 44 (6 1998), pp. 2384–2415. DOI: 10.1109/18.720542.

[32] G. David Forney Jr. and Lee-Fang Wei. “Multidimensional constellations. I. Introduction, figures of merit, and generalized cross constellations”. In: *IEEE Journal on Selected Areas in Communications* 7 (6 1989), pp. 877–892. DOI: 10.1109/49.29611.

[33] Michael Francis. *Infinite Impulse Response Filter Structures in Xilinx FPGAs*. White Paper: Spartan-3A DSP, Virtex-5/Virtex-4 FPGAs, LogiCORE IP. Tech. rep. WP330 (v1.2). Xilinx, Inc., 2009.

[34] George A. Franco. “Polarization Modulation Data Transmission System”. United States patent. 2992427. 1961.

[35] L. Franks. “Carrier and Bit Synchronization in Data Communication – A Tutorial Review”. In: *IEEE Transactions on Communications* 28 (8 1980), pp. 1107–1121. ISSN: 0090-6778. DOI: 10.1109/TCOM.1980.1094775.

[36] L. Franks and J. Bubrouski. “Statistical Properties of Timing Jitter in a PAM Timing Recovery Scheme”. In: *IEEE Transactions on Communications* 22 (7 1974), pp. 913–920. ISSN: 0090-6778. DOI: 10.1109/TCOM.1974.1092318.

[37] Floyd Martin Gardner. “A BPSK/QPSK Timing-Error Detector for Sampled Receivers”. In: *IEEE Transactions on Communications* 34 (5 1986), pp. 423–429. ISSN: 0090-6778. DOI: 10.1109/TCOM.1986.1096561.

[38] Floyd Martin Gardner. “Interpolation in digital modems. I. Fundamentals”. In: *IEEE Transactions on Communications* 41 (3 1993), pp. 501–507. ISSN: 0090-6778. DOI: 10.1109/26.221081.

[39] Floyd Martin Gardner. *Phaselock techniques*. 3rd ed. Hoboken NJ: John Wiley, 2005. xxii, 425. ISBN: 9780471430636.

[40] Claude Gasquet and Patrick Witomski. *Fourier analysis and applications: Filtering, numerical computation, wavelets*. Vol. 30. Texts in applied mathematics. New York: Springer, 1999. ISBN: 0387984852.

[41] Maria George and Peter Alfke. *Linear Feedback Shift Registers in Virtex Devices*. Tech. rep. XAPP 210. Xilinx, Inc., 1999.

[42] Y. Han and G. Li. “Experimental demonstration of direct-detection quaternary differential polarisation-phase-shift keying with electrical multilevel decision”. In: *Electronics Letters* 42 (2 2006), pp. 109–111. ISSN: 0013-5194. DOI: 10.1049/el:20063534.

[43] Eugene Hecht. *Optics*. 3rd ed. Addison-Wesley, 1998. ISBN: 0201838877.

[44] Eugene B. Hogenauer. "An economical class of digital filters for decimation and interpolation". In: *IEEE Transactions on Acoustics, Speech and Signal Processing*, 29 (2 1981), pp. 155–162. ISSN: 0096-3518. DOI: 10.1109/TASSP.1981.1163535.

[45] Institute of Electrical and Electronics Engineers, ed. *IEEE Standard for Reduced-Pin and Enhanced-Functionality Test Access Port and Boundary-Scan Architecture*. Std. 1149.7-2009. 2010, pp. c1–985. DOI: 10.1109/IEEESTD.2010.5412866.

[46] Institute of Electrical and Electronics Engineers, ed. *IEEE Standard VHDL Language Reference Manual*. Std. 1076-1987. 1988. DOI: 10.1109/IEEESTD.1988.122645.

[47] International Telecommunication Union, Radiocommunication sector, ed. *Estimation of polarization discrimination in calculations of interference between geostationary-satellite networks in the fixed-satellite service*. 05.1997. S.736-3. 1997. URL: http://www.itu.int/rec/R-REC-S.736-3-199705-I/en.

[48] Michel T. Ivrlač and Josef A. Nossek. "Toward a Circuit Theory of Communication". In: *IEEE Transactions on Circuits and Systems I: Regular Papers* 57.7 (July 2010), pp. 1663–1683. ISSN: 1549-8328. DOI: 10.1109/TCSI.2010.2043994.

[49] Michel C. Jeruchim, Philip Balaban and K. Sam Shanmugan. *Simulation of communication systems. Modeling, methodology, and techniques*. 2nd ed. New York: Kluwer Academic/Plenum Publishers, 2000. xxvii, 907. ISBN: 0306462672.

[50] Magnus Karlsson and Erik Agrell. "Spectrally efficient four-dimensional modulation". In: *Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference (OFC/NFOEC)*. 2012, pp. 1–3.

[51] Magnus Karlsson and Erik Agrell. "Which is the most power-efficient modulation format in optical links?" In: *Optics Express* 17 (13 2009), pp. 10814–10819.

[52] Kiyoshi Kobayashi et al. "Poly-polarization Multiplexing Scheme for Satellite Communications". In: *30th AIAA International Communications Satellite Systems Conference*. (Ottawa). Vol. 30. 2012.

[53] A. Kupiszewski. "The Gyrotron: A High-Frequency Microwave Amplifier". In: *Deep Space Network Progress Report 42-52, NASA Jet Propulsion Laboratory* (1979), pp. 8–12.

[54] Rajiv Laroia, Nariman Farvardin and Steven A. Tretter. “On optimal shaping of multidimensional constellations”. In: *IEEE Transactions on Information Theory* 40 (4 1994), pp. 1044–1056.

[55] Paul Leopardi. “Distributing points on the sphere”. School of Mathematics and Statistics. PhD thesis. University of New South Wales, 2007.

[56] Yair Linn. “Efficient Structures for PLL’s Loop Filter Design in FPGAs in High-Datarate Wireless Receivers – Theory and Case Study”. In: *Wireless Technology. Applications, Management, and Security*. Ed. by Steven Powell and J.P Shim. Vol. 44. Lecture Notes in Electrical Engineering. Springer US, 2009, pp. 115–132. ISBN: 978-0-387-71787-6. DOI: 10.1007/978-0-387-71787-6_8.

[57] Umberto Mengali and Aldo N. D’Andrea. *Synchronization techniques for digital receivers*. Applications of communications theory. New York: Plenum Press, 1997. xiii, 520. ISBN: 0-306-45725-3.

[58] Motorola, Inc. *SPI Block Guide*. S12SPIV3/D V03.06. Motorola, Inc. 2000. URL: www.motorola.com.

[59] Walter K. Niblack and Edwin H. Wolf. “Polarization Modulation and Demodulation”. United States patent. 3284632. 1966.

[60] NXP Semiconductors. *I^2C-bus specification and user manual*. 6th ed. UM10204. 4th Apr. 2014.

[61] Martin Oerder and Heinrich Meyr. “Digital filter and square timing recovery”. In: *IEEE Transactions on Communications* 36 (5 1988), pp. 605–612. DOI: 10.1109/26.1476.

[62] Алехандр Степанович Попов. “Прибор для обнаружения и регистрирования электрических колебаний”. В: *Журнал Русского физико-химического общества* (том XXVIII, часть физическая, 1896), с. 1—14.

[63] John G. Proakis and Masoud Salehi. *Digital communications*. In collab. with Masoud Salehi. 5th ed. Boston: McGraw-Hill, 2008. ISBN: 978-007-126378-8.

[64] Lennart Råde. *Springers mathematische Formeln: Taschenbuch für Ingenieure, Naturwissenschaftler, Wirtschaftswissenschaftler*. 2. Aufl. Berlin und Heidelberg: Springer, 1997. ISBN: 3540628290.

[65] Iulian Rosu. *Automatic Gain Control (AGC) in Receivers*. 2013. URL: www.qsl.net/va3iul/ (visited on 01/02/2013).

[66] Debabrata Saha and Theodore G. Birdsall. “Quadrature-quadrature phase-shift keying”. In: *IEEE Transactions on Communications* 37 (5 1989), pp. 437–448. DOI: 10.1109/26.24595.

[67] Adel A. M. Saleh. “Frequency-Independent and Frequency-Dependent Nonlinear Models of TWT Amplifiers”. In: *IEEE Transactions on Communications* 29 (11 1981), pp. 1715–1720.

[68] Hikmet Sari and Saïd Moridi. “New phase and frequency detectors for carrier recovery in PSK and QAM systems”. In: *IEEE Transactions on Communications* 36 (9 1988), pp. 1035–1043. ISSN: 0090-6778. DOI: 10.1109/26.7515.

[69] SES ASTRA. *ASTRA Universal Single LNB. Technical Recommendations*. 1.0. Sept. 2007.

[70] SES ASTRA. *ASTRA Universal Twin LNB. Technical Recommendations*. 1.1. Sept. 2007.

[71] L. F. Shampine. “Matlab program for quadrature in 2D”. In: *Journal of Computational and Applied Mathematics* 202(1) (2008), pp. 266–274.

[72] L. F. Shampine. “Vectorized Adaptive Quadrature in MATLAB”. In: *Journal of Computational and Applied Mathematics* 211 (2008), pp. 131–140.

[73] Claude E. Shannon. “A mathematical theory of communication”. In: *Bell System Technical Journal* 27 (1948), pp. 379–423.

[74] Károly Simonyi. *Theoretische Elektrotechnik*. 10. Aufl. Leipzig: Barth Verlagsgesellschaft, 1993. 973 S. ISBN: 3335003756.

[75] David Slepian. “Permutation modulation”. In: *Proceedings of the IEEE* 53 (3 1965), pp. 228–236. ISSN: 0018-9219. DOI: 10.1109/PROC.1965.3680.

[76] Neil. J. A. Sloane. *Nice arrangements of points on a sphere in various dimensions. Tables of Spherical Codes*. Part 1. Tables of putatively optimal packings in 3, 4 and 5 dimensions with n = 4, ..., 130 points. Information Sciences Research, AT&T Shannon Lab. URL: http://neilsloane.com/packings/dim3/ (visited on 2012).

[77] Xuan Tang. “Polarisation Shift Keying modulated free-space optical communication systems”. School of Computing, Engineering and Information Sciences. PhD thesis. Newcastle: University of Northumbria, 2012.

[78] X. Tang et al. “Experimental demonstration of polarisation shift keying in the free space optical turbulence channel”. In: *IEEE Conference on Communications in China Workshops ICCC*. 2012, pp. 31–36. DOI: 10.1109/ICCCW.2012.6316470.

[79] Giorgio Taricco, Ezio Biglieri and Valentino Castellani. “Applicability of four-dimensional modulations to digital satellites: a simulation study”. In: *IEEE Global Telecommunications Conference (GLOBECOM)*. including a Communications Theory Mini-Conference. Technical Program Conference Record. Vol. 4. 1993, pp. 28–34. DOI: 10.1109/GLOCOM.1993.318423.

[80] The Mathworks Inc. *Matlab Documentation*. "berawgn" function. Version R2012b. Closed-form formulas are given in the "berawgn" documentation. 2012.

[81] Wolfgang Utschick. *Statistische Signalverarbeitung*. Vorlesungsunterlagen. München: Technische Universität München, 19. Apr. 2010.

[82] Harry L. Van Trees. *Detection, estimation, and modulation theory*. New York: Wiley, 2001. ISBN: 9780471095170.

[83] Vinay K. Velkuru and Abhay Samant. “A design for software defined M-PSK radio on FPGA for low SNRs and symbol rates upto 10MS/s”. In: *IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*. 2011, pp. 574–578. DOI: 10.1109/ICSIPA.2011.6144169.

[84] Mladen Vucic and Marko Butorac. “All-digital high-dynamic automatic gain control”. In: *IEEE International Symposium on : Circuits and Systems (ISCAS)*. 2009, pp. 1032–1035. DOI: 10.1109/ISCAS.2009.5117935.

[85] Julian Webber et al. “Experimental Evaluation of a Poly-Polarization Multiplexing System with Timing/Frequency Recovery for Satellite Communications”. In: *31st AIAA International Communications Satellite Systems Conference (ICSSC)*. 2013.

[86] George R. Welti and Jhong S. Lee. “Digital transmission with coherent four-dimensional modulation”. In: *IEEE Transactions on Information Theory* 20 (4 1974), pp. 497–502. DOI: 10.1109/TIT.1974.1055247.

[87] Dietmar Wenzel. *Multirate Signal Processing*. Lecture notes. München: Technische Universität München, 2004–2005.

[88] Xilinx, Inc. *7 Series FPGAs Clocking Resources User Guide*. UG472. Version 1.6. 2nd Oct. 2012. URL: www.xilinx.com.

[89] Xilinx, Inc. *7 Series FPGAs Overview*. DS180. Version 1.12. 15th Oct. 2012. URL: www.xilinx.com.

[90] Xilinx, Inc. *7 Series FPGAs SelectIO Resources User Guide*. UG471. Version 1.3. 31st Oct. 2012. URL: www.xilinx.com.

[91] Xilinx, Inc. *KC705 Evaluation Board for the Kintex-7 FPGA User Guide*. UG810. Version 1.1. 2012. URL: www.xilinx.com.

[92] Xilinx, Inc. *LogiCORE IP CIC Compiler*. Product Specification. DS845. Version 3.0. 2011. URL: www.xilinx.com.

[93] Xilinx, Inc. *LogiCORE IP DDS Compiler*. Product Specification. DS794. Version 5.0. 2011. URL: www.xilinx.com.

[94] Xilinx, Inc. *LogiCORE IP FIR Compiler*. Product Specification. DS795. Version 6.3. 2011. URL: www.xilinx.com.

[95] Xilinx, Inc. *Synthesis and Simulation Design Guide*. UG626. Version 13.3. 2011.

[96] Kazuichi Yamamoto and Masayoshi Tanaka. "50W class multi-port amplifier for multi-beam satellite communications". In: *IEEE MTT-S International Microwave Symposium Digest*. 1989, 1281–1284 vol.3. DOI: 10.1109/MWSYM.1989.38961.

[97] Fumihiro Yamashita et al. "Variable Polarization/Frequency Division Multiplexing (VPFDM) for Satellite Communications". In: *IEEE 64th Vehicular Technology Conference (VTC)*. 2006, pp. 1–5. DOI: 10.1109/VTCF.2006.572.

[98] Oded Yaniv and Dan Raphaeli. "Near-optimal PLL design for decision-feedback carrier and timing recovery". In: *IEEE Transactions on Communications* 49 (9 2001), pp. 1669–1678. ISSN: 0090-6778. DOI: 10.1109/26.950353.

[99] Masanori Yofune et al. "Optimization of Signal Design for Poly-Polarization Multiplexing in Satellite Communications". In: *IEEE Communication Letters* 17 (11 2013), pp. 2017–2020. DOI: 10.1109/LCOMM.2013.100713.131618.

[100] Tsuyoshi Yoshido et al. "Cycle slip compensation with polarization block coding for coherent optical transmission". In: *IEEE Signal Processing Magazine* 31.2 (Mar. 2014), pp. 57–69.

[101] Xiang Zhou. "Efficient clock and carrier recovery algortihms for single-carrier coherent optical systems". In: *IEEE Signal Processing Magazine* 31.2 (2014), pp. 35–45.