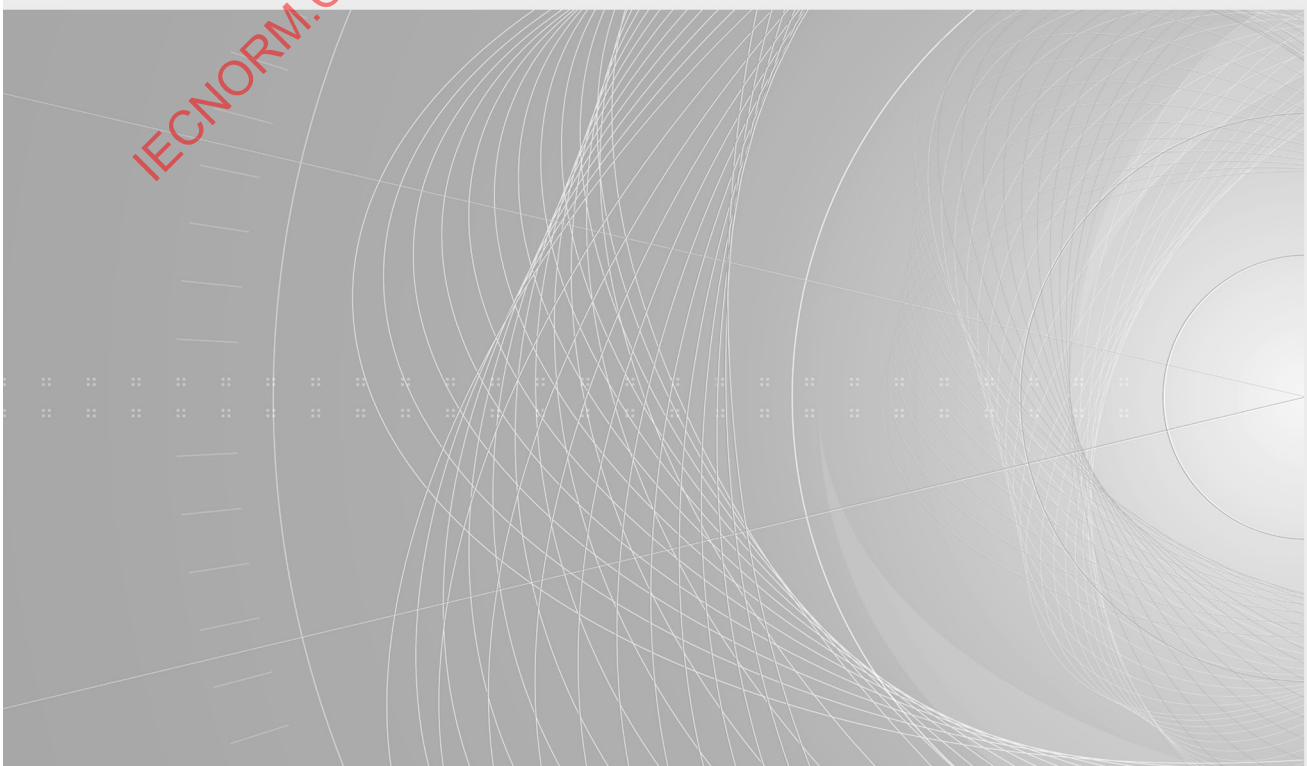
IEC TR 61282-16

Edition 1.0 2022-05

TECHNICAL REPORT

Fibre optic communication system design guidelines – Part 16: Coherent receivers and transmitters with high-speed digital signal processing



IEC TR 61282-16:2022-05(en)

IEC Secretariat
3, rue de Varembé
CH-1211 Geneva 20
Switzerland

Tel.: +41 22 919 02 11
info@iec.ch
www.iec.ch

About the IEC
The International Electrotechnical Commission (IEC) is the leading global organization that prepares and publishes International Standards for all electrical, electronic and related technologies.

About IEC publications
The technical content of IEC publications is kept under constant review by the IEC. Please make sure that you have the latest edition, a corrigendum or an amendment might have been published.

IEC publications search - webstore.iec.ch/advsearchform
The advanced search enables to find IEC publications by a variety of criteria (reference number, text, technical committee, …). It also gives information on projects, replaced and withdrawn publications.

IEC Just Published - webstore.iec.ch/justpublished
Stay up to date on all new IEC publications. Just Published details all new publications released. Available online and once a month by email.

IEC Customer Service Centre - webstore.iec.ch/csc
If you wish to give us your feedback on this publication or need further assistance, please contact the Customer Service Centre: sales@iec.ch.

IEC Products & Services Portal - products.iec.ch
Discover our powerful search engine and read freely all the publications previews. With a subscription you will always have access to up to date content tailored to your needs.

Electropedia - www.electropedia.org
The world's leading online dictionary on electrotechnology, containing more than 22 300 terminological entries in English and French, with equivalent terms in 19 additional languages. Also known as the International Electrotechnical Vocabulary (IEV) online.

IEC TR 61282-16

Edition 1.0 2022-05

TECHNICAL REPORT

Fibre optic communication system design guidelines – Part 16: Coherent receivers and transmitters with high-speed digital signal processing

INTERNATIONAL ELECTROTECHNICAL COMMISSION

ICS 33.180.01

ISBN 978-2-8322-0094-0

CONTENTS

INTERNATIONAL ELECTROTECHNICAL COMMISSION

_____

FIBRE OPTIC COMMUNICATION SYSTEM DESIGN GUIDELINES –

Part 16: Coherent receivers and transmitters with high-speed digital signal processing

FOREWORD

1) The International Electrotechnical Commission (IEC) is a worldwide organization for standardization comprising all national electrotechnical committees (IEC National Committees). The object of IEC is to promote international co-operation on all questions concerning standardization in the electrical and electronic fields. To this end and in addition to other activities, IEC publishes International Standards, Technical Specifications, Technical Reports, Publicly Available Specifications (PAS) and Guides (hereafter referred to as "IEC Publication(s)"). Their preparation is entrusted to technical committees; any IEC National Committee interested in the subject dealt with may participate in this preparatory work. International, governmental and non-governmental organizations liaising with the IEC also participate in this preparation. IEC collaborates closely with the International Organization for Standardization (ISO) in accordance with conditions determined by agreement between the two organizations.
2) The formal decisions or agreements of IEC on technical matters express, as nearly as possible, an international consensus of opinion on the relevant subjects since each technical committee has representation from all interested IEC National Committees.
3) IEC Publications have the form of recommendations for international use and are accepted by IEC National Committees in that sense. While all reasonable efforts are made to ensure that the technical content of IEC Publications is accurate, IEC cannot be held responsible for the way in which they are used or for any misinterpretation by any end user.
4) In order to promote international uniformity, IEC National Committees undertake to apply IEC Publications transparently to the maximum extent possible in their national and regional publications. Any divergence between any IEC Publication and the corresponding national or regional publication shall be clearly indicated in the latter.
5) IEC itself does not provide any attestation of conformity. Independent certification bodies provide conformity assessment services and, in some areas, access to IEC marks of conformity. IEC is not responsible for any services carried out by independent certification bodies.
6) All users should ensure that they have the latest edition of this publication.
7) No liability shall attach to IEC or its directors, employees, servants or agents including individual experts and members of its technical committees and IEC National Committees for any personal injury, property damage or other damage of any nature whatsoever, whether direct or indirect, or for costs (including legal fees) and expenses arising out of the publication, use of, or reliance upon, this IEC Publication or any other IEC Publications.
8) Attention is drawn to the Normative references cited in this publication. Use of the referenced publications is indispensable for the correct application of this publication.
9) Attention is drawn to the possibility that some of the elements of this IEC Publication may be the subject of patent rights. IEC shall not be held responsible for identifying any or all such patent rights.

IEC 61282-16 has been prepared by subcommittee 86C: Fibre optic systems and active devices, of IEC technical committee 86: Fibre optics. It is a Technical Report.

The text of this Technical Report is based on the following documents:

Draft	Report on voting
86C/1776/DTR	86C/1782/RVDTR

Full information on the voting for its approval can be found in the report on voting indicated in the above table.

The language used for the development of this Technical Report is English.

This document was drafted in accordance with ISO/IEC Directives, Part 2, and developed in accordance with ISO/IEC Directives, Part 1 and ISO/IEC Directives, IEC Supplement, available at www.iec.ch/members_experts/refdocs. The main document types developed by IEC are described in greater detail at www.iec.ch/standardsdev/publications.

A list of all parts in the IEC 61282 series, published under the general title *Fibre optic communication system design guidelines*, can be found on the IEC website.

Future documents in this series will carry the new general title as cited above. Titles of existing documents in this series will be updated at the time of the next edition.

The committee has decided that the contents of this document will remain unchanged until the stability date indicated on the IEC website under "http://webstore.iec.ch" in the data related to the specific document. At this date, the document will be

- reconfirmed,
- withdrawn,
- replaced by a revised edition, or
- amended.

INTRODUCTION

Coherent optical receivers are widely used in long-haul fibre optic communication systems, especially in systems that transmit optical carriers at data rates of 100 Gbit/s and higher. While the principle of coherent detection is very similar to that of super-heterodyne (or homodyne) detection in radio and microwave receivers, its implementation is significantly more challenging. The main reason is that optical frequencies are substantially higher than radio frequencies, so it becomes more difficult to match the local oscillator frequency in the coherent receiver to the frequency of the transmitted optical signal. Furthermore, optical signals tend to be highly polarized, which means that the amplitude of a coherently received signal can be substantially reduced or even vanish if the polarization state of the local oscillator light does not match the polarization state of the received optical signal. This polarization matching is particularly difficult to achieve in fibre optic communications systems, which usually do not preserve the launched state of polarization of the transmitted signal. To overcome these problems, modern coherent receivers typically consist of four parallel coherent optical mixers that provide phase and polarization diversity, and they rely on high-speed digital signal processors to retrieve the transmitted data from the four received electrical signals.

This rather complex coherent receiver architecture is further justified by the fact that it allows the receiver to mitigate various types of signal impairments introduced in the fibre optic link (or in the receiver itself) simply by means of additional electronic signal processing. Most notably, it is possible to substantially reduce the signal distortions caused by polarization-mode dispersion (PMD) or uncompensated chromatic dispersion (CD) in the fibre link, without requiring additional optical PMD or CD compensators. For this reason, coherent optical communication systems generally allow signal transmission at much higher data rates than communication systems using direct-detection receivers. Furthermore, coherent detection with subsequent digital signal processing facilitates the decoding of complex vector-modulated signals, such as quadrature-amplitude modulated signals (QAM) and polarization-multiplexed (PM) signals, and thereby the transmission of higher data rates.

Aside from fibre optic communications systems, coherent optical receivers are also used in various test and measurement instrumentation. Most notable examples are optical modulation analysers (OMAs) and high-resolution optical spectrum analysers (HR-OSAs). Optical modulation analysers are high-performance optical reference receivers and are used to assess the signal quality of complex vector-modulated optical signals. They are typically composed of a carefully calibrated coherent receiver and a high-speed real-time digitizing oscilloscope to record the coherently received signals, which are then analysed with the help of a software-based digital signal processor.

High-resolution optical spectrum analysers are often used to analyse narrowband features of the optical spectrum of a modulated signal, such as a residual optical carrier or other spectral lines. In contrast to OMAs, they typically employ a low-speed coherent receiver, so as to utilize the frequency-selectivity of coherent detection. The key component in these instruments is a continuously tuneable local oscillator, which is scanned over the frequency range of the signal to be analysed while the total power of the coherently received signal is recorded. The spectral resolution of these instruments can be of the order of a few MHz. Other examples of coherent optical test instruments include in-service PMD analysers and in-band optical signal-to-noise ratio analysers for polarization-multiplexed signals.

FIBRE OPTIC COMMUNICATION SYSTEM DESIGN GUIDELINES –

Part 16: Coherent receivers and transmitters with high-speed digital signal processing

1 Scope

This part of IEC 61282 is a technical report on coherent optical receiver and transmitter technologies that are employed in fibre optic communication systems as well as in optical test and measurement equipment. This document describes the principle of operation and functional capabilities of coherent optical receivers as well as the operation of optical transmitters used to generate complex vector-modulated signals. It is intended to serve as a technical foundation for other IEC documents and standards related to coherent optical transmission techniques.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

IEC 60050-731, *International Electrotechnical Vocabulary (IEV) – Part 731: Optical fibre communication* (available at www.electropedia.org)

IEC TR 61931, *Fibre optic – Terminology*

3 Terms, definitions, and abbreviated terms

3.1 Terms and definitions

For the purposes of this document, the terms and definitions given in IEC 60050-731 and IEC TR 61931 apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

- IEC Electropedia: available at http://www.electropedia.org/
- ISO Online browsing platform: available at http://www.iso.org/obp

3.2 Abbreviated terms

ABC	adaptive bias control
ADC	analogue-to-digital converter
AM	amplitude modulation
ASIC	application-specific integrated circuit
ASK	amplitude-shift keying
BER	bit-error ratio
BPSK	binary phase-shift keying
CD	chromatic dispersion
CFP	C form-factor pluggable
CMA	constant modulus algorithm

CMRR	common-mode-rejection ratio
COSA	coherent optical subassembly
CW	continuous wave
DAC	digital-to-analogue converter
DD-LMS	decision-directed least-mean-square
DGD	differential group delay
DSP	digital signal processor
DWDM	dense wavelength-division multiplexing
EDFA	erbium-doped fibre amplifier
ENOB	effective number of bits
EVM	error-vector magnitude
ER	extinction ratio
EQL	equalizer filter
FEC	forward error correction
FFT	fast Fourier transform
FIR	finite impulse response
FSE	fractionally spaced equalizer
FWM	four-wave mixing
GS	geometrically shaped
GVD	group velocity dispersion
HB-CDM	high-bandwidth coherent driver modulator
HR-OSA	high-resolution optical spectrum analyser
LO	local oscillator
LPF	low-pass filter
MAP	maximum a posteriori
MMA	multi-modulus algorithm
MMI	multi-mode interference
MZM	Mach-Zehnder modulator
NLFT	nonlinear Fourier transform
NLPN	nonlinear phase noise
I	in-phase component
IA	implementation agreement
ICR	intradyne coherent receiver
IC-TROSA	integrated coherent transmit-receive optical subassembly
iFFT	inverse fast Fourier transform
Im	imaginary part
InP	indium phosphide
ITLA	integrable tuneable laser assembly
OFDM	orthogonal frequency-division multiplexing
OMA	optical modulation analyser
OOK	on-off keying
OSA	optical spectrum analyser
OSNR	optical signal-to-noise ratio

OTN	optical transport network
PAM	pulse-amplitude modulation
PBC	polarization beam combiner
PBS	polarization beam splitter
PC	polarization controller
PD	photodiodes, photo detectors
PDL	polarization-dependent loss
PLC	planar lightwave circuit
PMD	polarization-mode dispersion
PM	polarization multiplexed
PM-QPSK	polarization multiplexed QPSK
PR	polarization rotator
PRBS	pseudo-random binary sequence
PS-QPSK	polarization switched QPSK
PSK	phase-shift keying
Q	quadrature component
QAM	quadrature amplitude modulation
QASK	quaternary amplitude-shift keying
QPSK	quaternary phase-shift keying
RAM	random access memory
RDE	radius-directed equalizer
Re	real part
RF	radio frequency
RMS	root-mean-square
ROADM	reconfigurable optical add-drop multiplexer
RRC	root-raised-cosine
SiPh	silicon photonics
SNR	signal-to-noise ratio
SPM	self-phase modulation
SSB	single-sideband
TDE	time-domain equalizer
TEF	transversal electrical filter
TIA	transimpedance amplifier
TROSA	transmit-receive optical subassembly
VCO	voltage-controlled oscillator
WDM	wavelength-division multiplexing
X	X-polarized component
XI	in-phase component of X-polarized signal
XPM	cross-phase modulation
XQ	quadrature component of X-polarized signal
Y	Y-polarized component
YI	in-phase component of Y-polarized signal
YQ	quadrature component of Y-polarized signal

4 Background

The development of coherent optical receivers started in the late 1980s and was driven by the desire to improve the sensitivity of optical receivers in order to increase the reach of fibre-optic communication systems [1][1]. In addition, coherent receivers were expected to facilitate decoding of phase-shift-keyed signals (PSKs) and to serve as narrow-band frequency selectors in wavelength-division multiplexed systems. However, the interest in coherent receivers diminished rapidly after practical Erbium-doped fibre amplifiers (EDFAs) became available just a few years later, because these optical amplifiers enabled repeater-less transmission of optical signals over very long distances without the need for coherent detection. Besides, coherent receivers at that time were still not mature enough to be deployed in practical fibre optic communication systems, because coherent detection was very sensitive to the polarization state of the received signals and, in addition, required transmitter and local-oscillator lasers with very high phase stability [2].

The interest in coherent optical receiver technology revived in the late 2000s, when high-speed analogue-to-digital converters (ADCs) and high-speed digital signal processors (DSPs) became available. These high-speed ADCs sample the received electrical signals at or above the symbol rate of the transmitted signals and deliver the digitized signals to a DSP, which then processes the sampled electrical signals in real time to retrieve the transmitted digital or analogue information [3]. It turned out that the combination of linear signal detection in coherent receivers and signal processing in high-speed DSPs not only solved the technical problems of early-generation coherent receivers, but also enabled electronic mitigation of linear (and even nonlinear) signal distortions encountered in the optical communication link [4]. Most importantly, coherent reception with subsequent digital signal processing allows adaptive mitigation of undesired impairments caused by polarization-mode dispersion (PMD) and uncompensated chromatic dispersion (CD) in the transmission fibre, either of which can severely distort the waveforms of optical signals at data rates of 10 Gbit/s or higher [6]. Modern coherent receivers are often capable of accommodating large amounts of PMD and CD and, as a result, allow signal transmission at much higher modulation rates and/or over much longer distances than direct-detection receivers [5]. Moreover, the DSP used in coherent receivers can even mitigate signal distortions resulting from non-ideal optical and electronic components in the coherent receiver itself, such as signal skew, amplitude imbalance, and other deviations from the desired frequency response, as described in more detail in Clause 5 of this document.

Coherent receivers also facilitate the decoding of complex vector-modulated signals, such as M-ary PSK or quadrature-amplitude modulation (M-QAM), which can carry several times the information of binary on-off-keyed (OOK) signals of the same symbol rate [7]. Moreover, modern coherent receivers are capable of decoding polarization-multiplexed (PM) signals, which can carry up to twice the information of corresponding single-polarized signals [7], [8]. In fact, the first commercially deployed fibre optic communication systems with coherent detection transmitted polarization-multiplexed quaternary phase-shift-keyed (PM-QPSK) signals at symbol rates of 10 GBd, yielding a total transmission rate of 40 Gbit/s per optical carrier [9]. Just a few years later, a second generation of coherent receivers was introduced allowing transmission of single-carrier PM-QPSK signals at 100 Gbit/s [10]. Subsequent development of transmitters with linear modulation response then enabled transmission of polarization-multiplexed 16QAM signals at 200 Gbit/s and even 400 Gbit/s [11]. Moreover, transmission of 32QAM and 64QAM signals under practical conditions has been demonstrated, and the feasibility of transmitting of even higher-order QAM signals is under investigation [13], [14], [15].

The generation of higher-order optical QAM signals usually requires linear modulators and driver amplifiers as well as careful equalization of non-ideal transmitter characteristics introduced by the optical and electronic transmitter components, including frequency roll-off, skew, and nonlinear response. This equalization can be achieved by employing digital signal processing in the transmitter to facilitate adjustable pre-distortion of the electrical signals applied to the optical modulator [2]. Therefore, as soon as fast high-speed digital-to-analogue

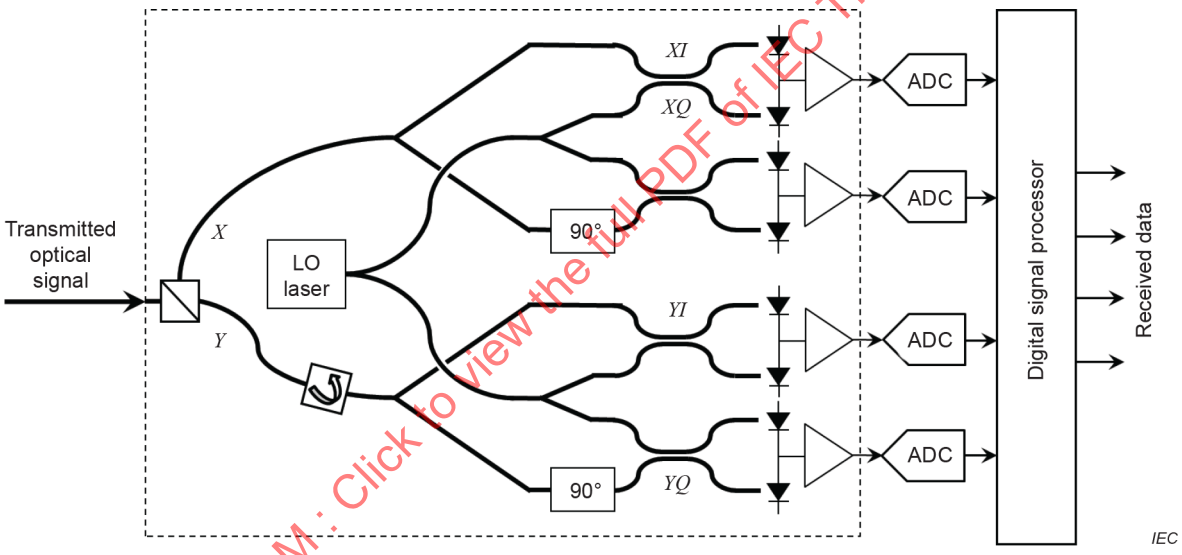
[1] Numbers in square brackets refer to the Bibliography.

converters (DACs) became commercially available, high-speed DSPs were also introduced in optical transmitters. In addition to equalization, the transmitter DSPs are frequently used to spectrally shape the electrical drive signals applied to the optical modulator in order to generate optical signals with the desired optical spectrum [12]. Transmitter DSPs also facilitate electronic pre-compensation of chromatic dispersion in the transmission fibre [6], [15]. Clause 8 describes the generation of optical signals with complex modulation formats and the functions of transmitter DSPs.

5 Coherent transmission of vector-modulated signals

5.1 Typical receiver architecture

As mentioned in Clause 4, modern coherent receivers for fibre optic communication systems usually employ four parallel coherent mixers that are configured in such a way that they detect the optical amplitude, phase, and polarization state of the received signal. Figure 1 displays a typical arrangement of these four mixers, which are labelled XI, XQ, YI, and YQ [16]. The mixers thus generate four analogue electrical signals, which are then converted to digital signals via high-speed analogue to digital converters (ADCs) and further processed in a high-speed digital signal processor (DSP).



Key

ADC Analogue-to-digital converter

LO Local oscillator

X X-polarized signal

Y Y-polarized signal

XI In-phase component of X-polarized signal

XQ Quadrature component of X-polarized signal

YI In-phase component of Y-polarized signal

YQ Quadrature component of Y-polarized signal

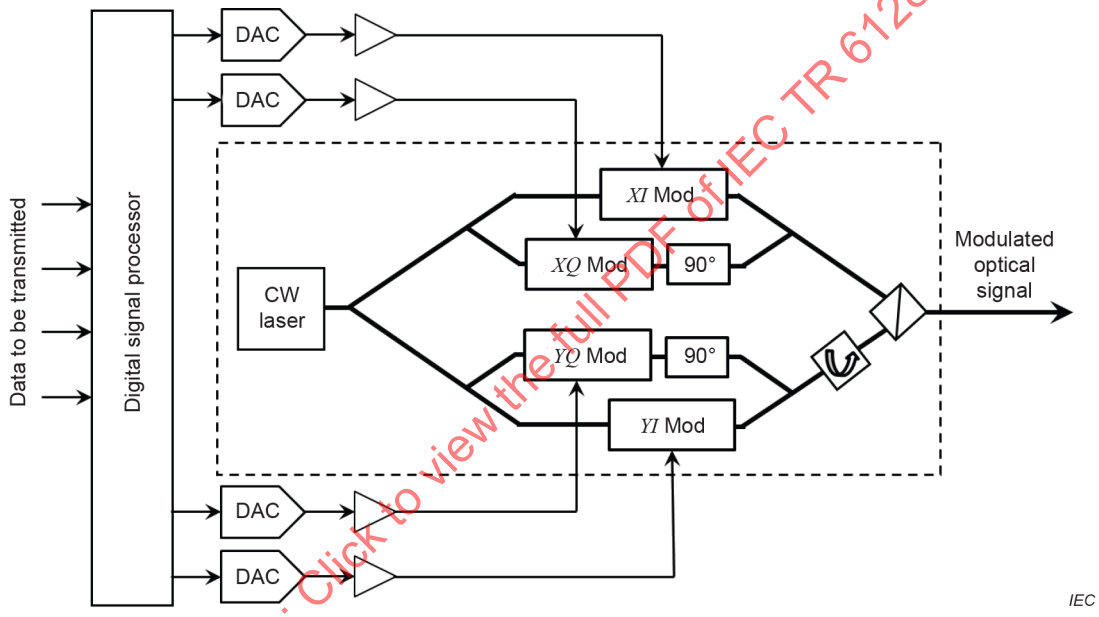
Figure 1 – Coherent optical receiver

The operation of these mixers and their capabilities will be discussed in more detail in Clause 6. It is important to note that the arrangement shown in Figure 1 makes the coherent receiver insensitive to time-varying fluctuations in the optical phase and polarization state of the received signal, which is a highly desirable feature for practical applications in fibre optic communication systems [1]. Moreover, it also allows the receiver to decode complex vector-modulated signals in which amplitude, phase, and polarization state are independently modulated [8], [17]. As described in more detail in 5.3, these vector-modulated signals can carry several times more

information than conventional on-off-keyed (OOK) signals transmitted at the same symbol rate. Hence, transmission of information at a given data rate with vector-modulated signals can be accomplished at much lower symbol rates, and hence with much lower electrical and optical bandwidth, than transmission with conventional OOK signals, which is also a highly desirable feature. However, decoding of these vector-modulated signals in the coherent receiver is a fairly complex process, which is usually accomplished by employing a high-speed DSP, as discussed in more detail in Clause 7.

5.2 Typical transmitter architecture

Coherent transmission systems frequently employ optical transmitters that are capable of generating complex vector-modulated signals. An example of such a transmitter is displayed in Figure 2, where the transmitter comprises four parallel modulators, which are labelled XI, XQ, YI, and YQ. These four modulators are typically employed in a nested Mach-Zehnder configuration in such a way that they allow independent modulation of the optical amplitude, phase, and polarization state of the transmitted signal, as discussed in more detail in Clause 8 [18].



Key

90° Optical phase shifter
CW Continuous wave
DAC Digital-to-analogue converter
Mod Optical modulator
XI In-phase component of X-polarized signal
XQ Quadrature component of X-polarized signal
YI In-phase component of Y-polarized signal
YQ Quadrature component of Y-polarized signal

Figure 2 – Optical transmitter for coherent transmission

The labels XI, XQ, YI, and YQ (or IX, QX, IY, and QY) are widely used in the industry to characterize the various signal components in vector-modulated signals, both for the optical transmitter and the optical receiver [16], [18]. However, it is very important to note that the signal components generated in the transmitter by the XI, XQ, YI, and YQ modulators are in general different from the signal components generated by the XI, XQ, YI, and YQ mixers in the coherent receiver. In fact, the received signal components in XI, XQ, YI, and YQ are related to the transmitted signal components in XI, XQ, YI, and YQ by a complex and often frequency-dependent function (see 7.1). Therefore, a DSP is often employed to de-convolute the transmitted signals from the four received signal components.

NOTE The use of the same labels for the transmitter signal components and the receiver signal components can lead to considerable confusion, because there usually is no simple relationship between the transmitter and receiver components. For example, if a transmitter generates signals using only the *XI* and *YI* modulators (e.g. to generate polarization-multiplexed binary phase-shift-keyed signals), it is sometimes believed that the coherent receiver would only need *XI* and *YI* mixers. As discussed in Clause 6, this is only the case for homodyne receivers with a phase-locked oscillator laser. However, modern intradyne receivers need all four *XI*, *XQ*, *YI*, and *YQ* mixers to recover the transmitted signal components from the received signal.

Frequently, DSPs are also employed in the transmitter in conjunction with high-speed digital-to-analogue converters (DACs). As discussed in more detail in Clause 9, the DSP is used here to map the information data into the transmitted signal components and to pre-distort the digital drive signals for the modulators.

Unlike conventional direct-detection receivers, a coherent receiver requires a continuous-wave (CW) laser to serve as a local oscillator (LO), with which the received signal components are mixed, as shown in Figure 1. This laser typically emits unmodulated light at a single (but tuneable) optical frequency, just like the transmitter lasers in conventional non-coherent communication systems [19]. Coherent detection, however, requires transmitter and LO lasers with relatively low optical phase noise. It should be noted that the phase noise requirements for LO lasers can be different from those for the transmitter lasers, as discussed in more detail in 6.3.3, 6.4.5, and 6.4.6.

5.3 Vector-modulated signals

5.3.1 Mathematical description

In this document, the term vector-modulated signals is used for optical signals that carry information data in their amplitude, phase, and/or polarization state [8], [17]. These signals can be described by a two-dimensional complex vector $\vec{A}_\mathrm{s}(t)$ of the general form shown in Formula (1).

$$\vec{A}_\mathrm{s}(t) = \sqrt{P_\mathrm{S}}\begin{pmatrix} A_\mathrm{X}(t) \\ A_\mathrm{Y}(t) \end{pmatrix} \mathrm{e}^{\,\mathrm{j}\,\Phi_\mathrm{s}(t)} = \sqrt{P_\mathrm{S}}\begin{pmatrix} A_\mathrm{XI}(t) + \mathrm{j}\,A_\mathrm{XQ}(t) \\ A_\mathrm{YI}(t) + \mathrm{j}\,A_\mathrm{YQ}(t) \end{pmatrix} \mathrm{e}^{\,\mathrm{j}\left[\omega_\mathrm{s}t + \varphi_\mathrm{s}(t) + \hat{\varphi}_\mathrm{s}(t)\right]} \tag{1}$$

where

P_S is the average optical power of the signal;

$A_\mathrm{X}(t)$ and $A_\mathrm{Y}(t)$ are normalized complex amplitudes with $\left|A_\mathrm{X}(t)\right|^2 + \left|A_\mathrm{Y}(t)\right|^2 = 1$, which describe the modulated signal components in two mutually orthogonal polarization states;

$A_\mathrm{XI}(t)$, $A_\mathrm{XQ}(t)$, $A_\mathrm{YI}(t)$, and $A_\mathrm{YQ}(t)$ are real-valued amplitudes, which represent the in-phase and quadrature signal components of $A_\mathrm{X}(t)$ and $A_\mathrm{Y}(t)$, respectively;

$\Phi_\mathrm{s}(t) = \omega_\mathrm{s}t + \varphi_\mathrm{s}(t) + \hat{\varphi}_\mathrm{s}(t)$ is a time-varying common phase;

$\omega_\mathrm{s}=2\pi f_\mathrm{s}$ is the angular frequency of the modulated optical carrier;

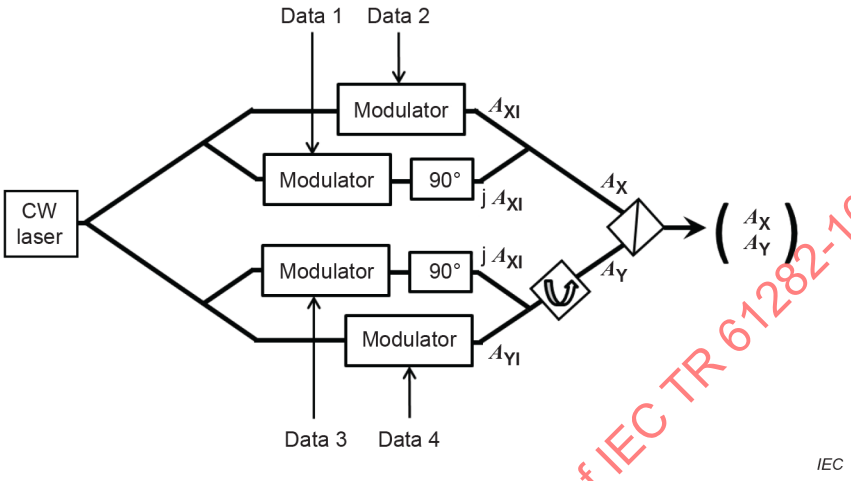
$\varphi_\mathrm{s}(t)$ represents optical phase variations;

$\hat{\varphi}_\mathrm{s}(t)$ is a random variable representing laser phase noise from the transmitter laser;

e is Euler's number, approximately equal to 2,71828; and

$\mathrm{j} = \sqrt{-1}$ is the imaginary unit.

The four signal components A_{XI}, A_{XQ}, A_{YI}, and A_{YQ} usually carry different information (i.e. different payload data). As briefly discussed in 5.2, they are typically generated by four parallel amplitude modulators, which are arranged in a nested Mach-Zehnder structure. This arrangement includes 90° optical phase shifters to generate the quadrature components $A_{XQ}(t)$ and $A_{YQ}(t)$ as well as a polarization rotator to generate the orthogonally polarized $A_Y(t)$ signal component, as shown schematically in Figure 3 and described in more detail in Clause 8.



Key

90° Optical phase shifter

CW Continuous wave

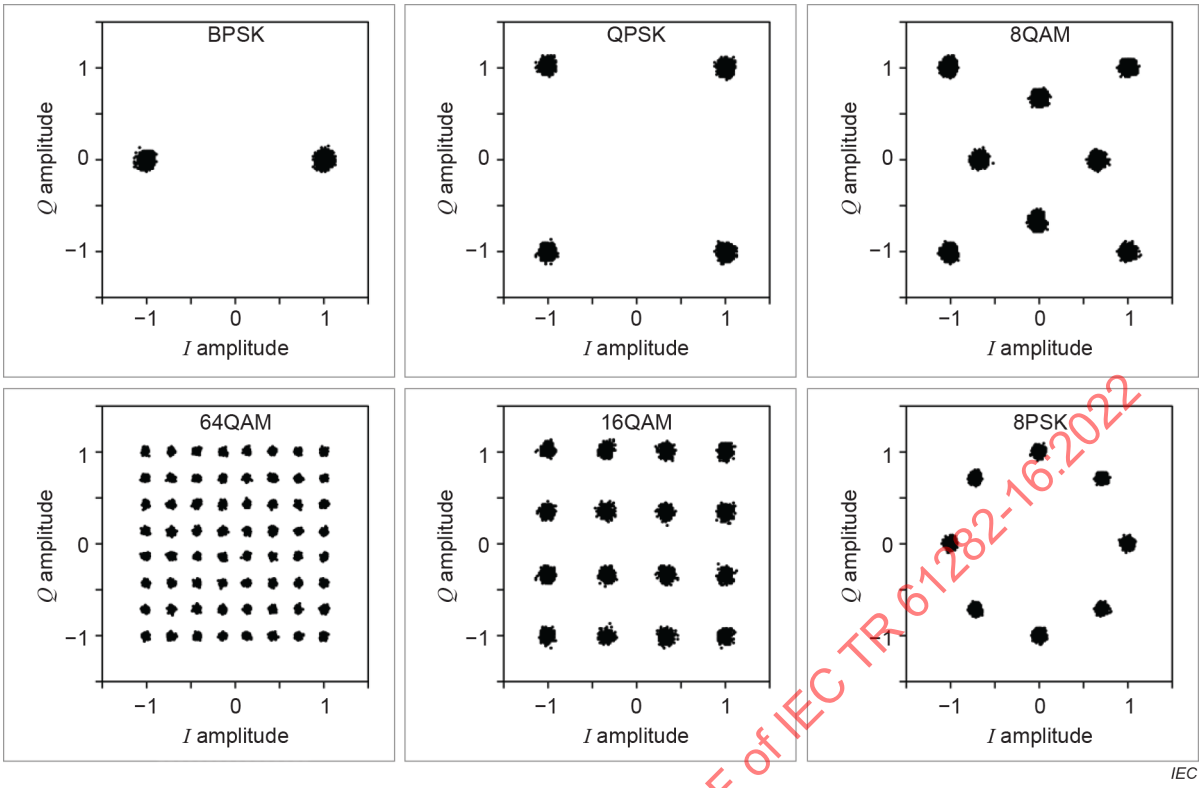
Figure 3 – Generation of vector modulated signals

The orthogonal polarization states represented by A_X and A_Y are often a pair of linearly polarized states when the signal is generated in the optical transmitter. However, this pair of orthogonal polarization states is generally transformed in the transmission fibre into another pair of orthogonal polarization states before the signal arrives at the receiver.

5.3.2 Binary amplitude and phase modulation

Conventional non-coherent communication systems typically transmit binary on-off-keyed (OOK) signals in a single polarization state [1]. In these signals, only $A_{XI}(t)$ is modulated (i.e. toggled between 0 and 1) while $A_{XQ} \equiv 0$ and $A_Y \equiv 0$ in Formula (1). Single-polarized binary OOK signals thus carry one bit of information per transmitted symbol and can be readily detected non-coherently with a single photodetector. OOK signals are sometimes generated either by an optical amplitude modulator (e.g. of the type discussed in 8.2), or alternatively, by a simpler intensity modulator.

Single-polarized binary phase-shift-keyed (BPSK) signals are typically generated by an optical amplitude modulator, which toggles $A_{XI}(t)$ between −1 and 1 while keeping $A_{XQ} \equiv 0$ and $A_Y \equiv 0$, as shown in the constellation diagram of Figure 4. Single-polarized BPSK signals can carry only one bit of information per transmitted symbol, but their detection requires a receiver that is sensitive to the phase of the received signal [20]. Single-polarized BPSK signals can be detected non-coherently with a properly tuned differential delay interferometer and a single photodetector. In principle, it is also possible to detect these signals with a coherent receiver having only a single mixer, as described in 6.2. However, in practical applications it is far more convenient and easier to detect these signals with a coherent receiver of the type shown in Figure 1, which has four parallel mixers, providing phase- and polarization-diversity, so that the receiver can tolerate phase and frequency variations in the transmission and reference lasers as well as time-varying polarization transformations in the transmission fibre.



Key

BPSK Binary phase-shift keying

I In-phase component

PSK Phase-shift keying

Q Quadrature component

QAM Quadrature-amplitude modulation

QPSK Quaternary phase-shift keying

Figure 4 – Examples of modulation formats for coherent communication

5.3.3 Quadrature amplitude modulation

The number of information bits carried by each transmitted symbol can be increased by also modulating the quadrature component A_{XQ} in addition to and independently of the in-phase component A_{XI}. Single-polarized signals using quaternary phase-shift-keying (QPSK), for example, carry two bits of information per transmitted symbol, since A_{XI} and A_{XQ} are both toggled independently between −1 and +1, as shown in Figure 4 [8]. QPSK signals and higher order quadrature amplitude modulated signals can be generated by two parallel amplitude modulators, as discussed in 8.3.

Even higher bit rates can be transmitted by using multi-level phase-shift keying (PSK) or quadrature amplitude modulation (QAM), such as 8PSK or 8QAM, which carry three bits per symbol. Moreover, 16QAM signals carry four bits per symbol and 64QAM signals even six bits per symbol. Single-polarized QPSK signals and multi-level PSK or QAM signals can be detected with coherent receivers that employ at least two parallel mixers with phase diversity, as discussed in 6.3. However, in practice it is far more convenient to detect these signals with a phase- and polarization-diversity coherent receiver, because of time-varying polarization transformations in the transmission fibre.

5.3.4 Polarization multiplexing

For any given modulation format, the number of bits per transmitted symbol can be increased further by transmitting polarization-multiplexed (PM) or dual-polarization (DP) signals, in which both polarized components, $A_X(t)$ and A_Y, are independently modulated. Polarization-multiplexed QPSK (PM-QPSK) signals, for example, carry four bits of information per symbol [8]. Polarization-multiplexed signals can readily be detected with phase- and polarization-diversity coherent receivers, which typically employ four parallel mixers, as described in 6.4. Coherent receivers thus facilitate the use of polarization multiplexing and higher order modulation formats, which allow transmission at higher data rates without requiring higher symbol rates. This is important because the electrical bandwidth requirements for the various electrical and optical components in the optical transmitter and receiver predominantly depend on the symbol rate.

5.3.5 Higher-dimensional coding and constellation shaping

It should be noted that higher-order QAM and PSK signals are more sensitive to optical noise than binary ASK or OOK signals and, hence, usually do not have the same transmission reach as binary signals [8]. In general, the more bits of information encoded in a signal, the more sensitive the signal is to optical noise and to other linear and nonlinear distortions. To make the transmission of higher-order QAM and PSK signals more robust to such impairments, new modulation techniques have been introduced in optical communication systems, which include multi-dimensional coding as well as geometrically and probabilistically shaped modulation formats [21], [22], [23].

In four-dimensional (4D) coding, the two polarization tributaries are not independently modulated, as described above, but instead are varied in such a way that the constellation points of the combined optical signal are distributed in a four-dimensional constellation space (whereas conventional polarization-multiplexed signals can be represented in a three-dimensional constellation space) [21]. A simple example of a 4D modulation format is polarization-switched QPSK (sometimes referred to as PS-QPSK), where only one polarization tributary is transmitted at any given time. Thus, when a QPSK symbol is transmitted in the X polarization in time slot n, then no signal is transmitted in the Y polarization, so that $A_{XC}(t_n) \neq 0$ and $A_{YC}(t_n) \equiv 0$; whereas, in another time slot m, a QPSK symbol can be transmitted in the Y polarization and no signal is transmitted in the X polarization, so that $A_{XC}(t_m) \equiv 0$ and $A_{YC}(t_m) \neq 0$. In PS-QPSK signals, the alternating occurrence of symbols and zeros in the two polarization tributaries is digitally modulated (i.e. non-periodic), so that an additional bit of information can be encoded in each transmitted QPSK symbol, yielding a total information rate of 3 bits per symbol (versus 4 bits per symbol for PM-QPSK). Hence, PS-QPSK signals carry less information than PM-QPSK signals, but when transmitted with the same average optical power, they are more resilient to optical noise than PM-QPSK signals [21]. This concept has been extended to modulation formats of higher cardinality, in which case the two polarization tributaries alternatingly transmit orthogonal subsets of constellation points, or even to eight-dimensional (8D) coding [21], [24].

In geometrically shaped (GS) modulation formats, the constellation points in either polarization tributary are arranged in such a way that the smallest Euclidian distance between them is as large as possible; whereas in probabilistically shaped modulation formats, the digital information is mapped into the constellation space in such a way that the frequency of occurrence of each constellation point is not equiprobable, as in conventional modulation formats, but is distributed, on average, according to a predetermined probability function, so that symbols carrying the innermost constellation points are more frequently transmitted than symbols carrying the outermost constellation points. Hence, for a given number of constellation points, probabilistic shaping reduces the information rate that can be transmitted, which is not the case for geometric shaping. To carry the same information rate as a conventional signal with equiprobable constellation points, a probabilistically shaped signal generally requires a modulation format of higher cardinality. For example, a probabilistically shaped 64QAM signal could transmit the same information rate as conventional 16QAM or 32QAM signals, and a probabilistically shaped 1024QAM signal the same information rate as a conventional 256QAM, but with higher tolerance to optical noise and with reduced optical bandwidth [25]. The higher

noise tolerance of probabilistically shaped signals arises from the fact that the innermost constellation points are more frequently transmitted than the outermost constellation points. Hence, at the same average optical signal power, the symbols of a probabilistically shaped signal carry more optical power than the same symbols of a conventional signal of the same modulation format and cardinality but with equiprobable constellation points. In terms of optical amplitude, therefore, the smallest Euclidian distance between the constellation points of a probabilistically shaped signal is effectively larger than that of a conventional signal of the same modulation format and cardinality.

The above described techniques of higher-dimensional coding, geometrical shaping, and probabilistic shaping are not mutually exclusive, but can in fact be employed simultaneously [22]. However, a more detailed description of these new coding schemes is beyond the scope of this document. Application of these advanced coding techniques requires additional signal processing at the transmitter and receiver, but can be accomplished with the same coherent receivers and optical transmitters as used for conventional vector-modulated formats (see 6.4 and 8.4).

Likewise, the same type of coherent receivers can be used for multi-carrier modulation formats, which are signals that are composed of independently modulated sub-carriers, like signals that employ orthogonal frequency-division multiplexing (OFDM) [26].

6 Coherent receiver architectures and functional capabilities

6.1 Basic principle of coherent detection

6.1.1 General

A coherent optical receiver mixes the transmitted optical signal with an optical reference signal from an unmodulated local-oscillator (LO) laser, which generates a beat signal that is proportional to the product of the optical amplitudes of the two signals [1], [2]. The two optical signals are usually mixed in one or more high-speed photodiodes (PDs), utilizing the nonlinear (i.e. quadratic) response of the photodetectors to generate the desired beat signal. The resulting photocurrents are proportional to the time-varying amplitude and phase of the transmitted optical signal, as described in more detail in 6.2, 6.3, and 6.4. In fact, the operation of a coherent receiver is very similar to that of super-heterodyne (or super-homodyne) receivers, which are often used in radio and microwave transmission systems to increase receiver sensitivity and to enhance the frequency selectivity of the receiver.

6.1.2 Homodyne and heterodyne detection

Coherent receivers can be operated as homodyne, heterodyne or intradyne detectors [2], [20]. In the case of homodyne detection, the reference signal has exactly the same optical frequency as the carrier frequency of the transmitted optical signal, as shown schematically in Figure 5 [26].

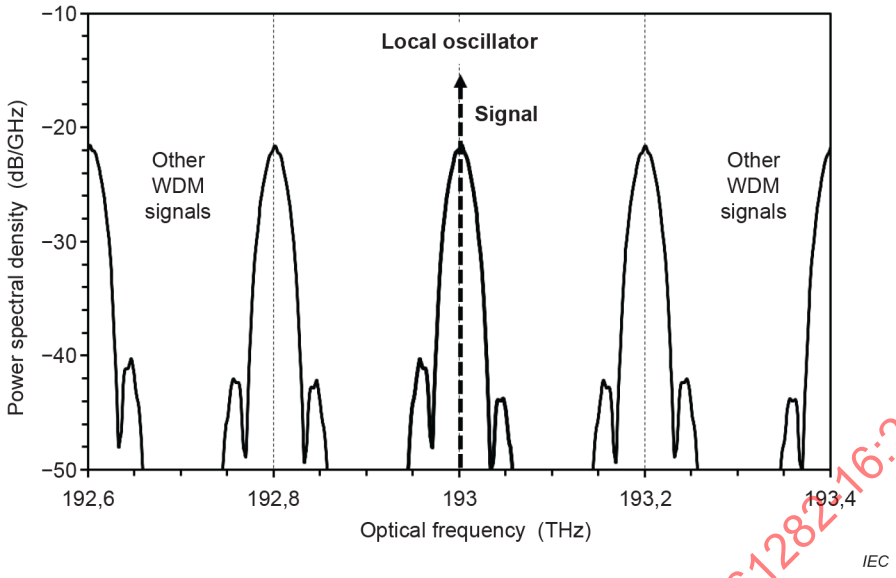


Figure 5 – Signal and local oscillator frequencies for homodyne detection

In the case of heterodyne detection, the reference frequency and the carrier frequency are offset by an amount that is slightly larger than the spectral width of the transmitted signal, as shown in Figure 6 [26]. Intradyne detection is very similar to homodyne detection, but it allows for relatively small frequency offsets between the reference signal and the carrier frequency of the transmitted optical signal [2]. Coherent fibre optic communication systems as well as optical modulation analysers typically use intradyne detection [9].

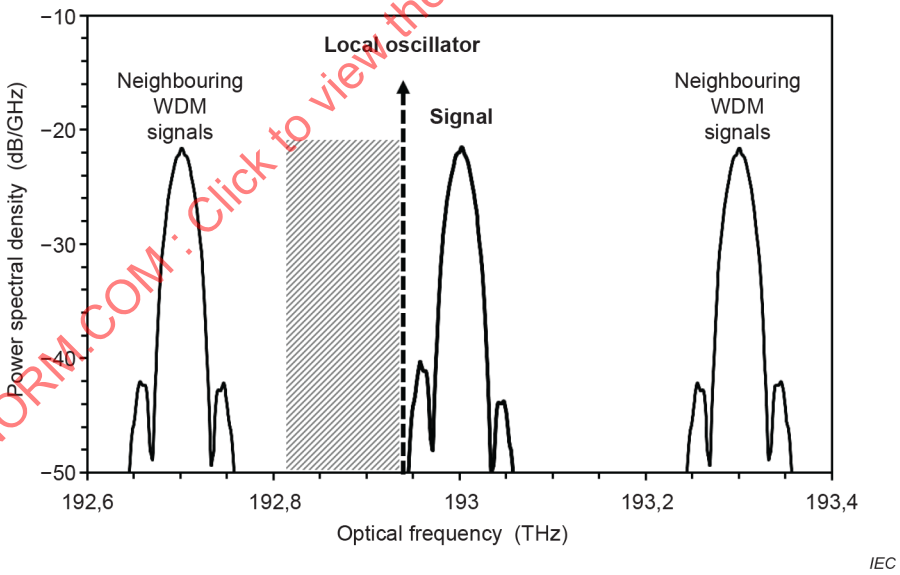


Figure 6 – Signal and local oscillator frequencies for heterodyne detection

In either case, mixing of the reference signal with the transmitted signal in the PD generates a photocurrent in the electrical baseband, having a single-sided spectrum, as shown schematically in Figure 7 for homodyne and heterodyne detection [26].

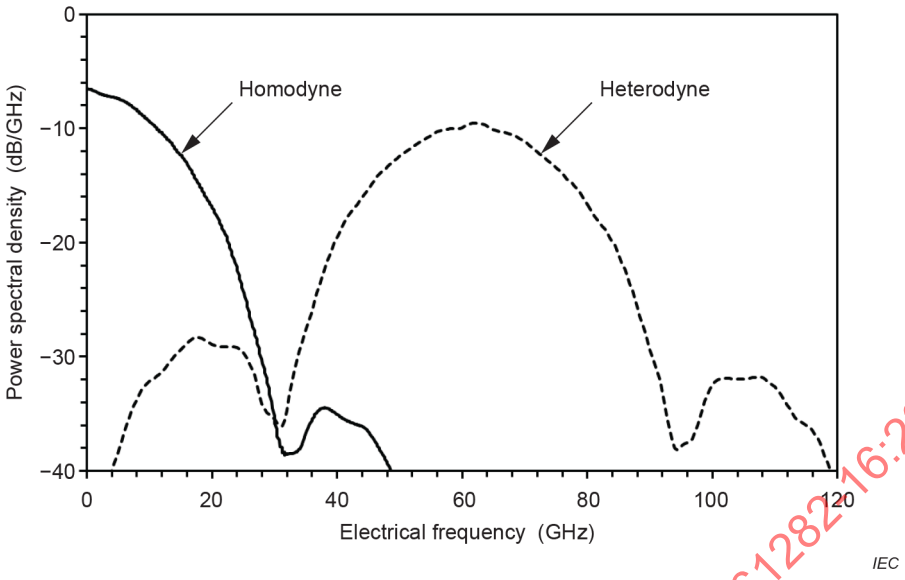


Figure 7 – Electrical spectra of homodyne and heterodyne beat signals

In the case of heterodyne detection, the carrier frequency of the transmitted signal is translated to an intermediate electrical frequency $|f_r - f_c|$, where f_r denotes the reference frequency and f_c the carrier frequency of the transmitted signal. If the two optical frequencies are sufficiently different, the entire (two-sided) spectrum of the transmitted optical signal can be translated to the single-sided electrical spectrum, as shown in the example of Figure 7, thus preserving the information contained in the upper and lower modulation sidebands of the transmitted signal.

6.1.3 Intradyne detection

In the case of homodyne detection, the reference frequency is exactly equal to the carrier frequency of the transmitted signal, that is $f_r = f_c$; whereas for intradyne detection, the reference frequency is only approximately equal to the transmitted carrier frequency, i.e. $f_r \approx f_c$ [2]. In both cases, the upper and lower modulation sidebands of the transmitted signal are folded into the single sideband of the electrical spectrum, which means that each frequency component of the resulting electrical spectrum is composed of two mixing products, one of which results from mixing of the LO laser with an optical frequency component in the lower modulation sideband of the transmitted signal at frequency f_l, and the other one from mixing of the LO laser with an optical frequency component in the upper modulation sideband of the signal at frequency f_u, with $f_r - f_l = f_u - f_r$.

In general, this spectral folding of the mixing products does not preserve all phase information contained in the transmitted optical signal. Hence, the decoding capabilities of a single coherent mixer are limited. To avoid these restrictions, modern coherent communication systems employ multiple intradyne mixers in the receiver, which generate mixing products at different optical phases between the LO laser output and the transmitted signal, as discussed in more detail in 6.3. This dual-mixer architecture is commonly referred to as a phase diversity receiver [27]. It should be noted that heterodyne receivers do not have these limitations but require photodiodes with twice the bandwidth of those used in intradyne receivers [26].

6.1.4 Polarization dependence

Heterodyne as well as homodyne mixers are sensitive to the polarization state of the transmitted signal. This effect causes fading of the mixing product when LO laser light and the received signal are in different polarization states. Such polarization fading can be avoided by employing additional mixers to generate mixing products between the LO laser and copies of the transmitted signal in different polarization states, which is known as polarization diversity [28]. Polarization-diversity receivers are discussed in more detail in 6.4 and 6.5.

It should be noted that heterodyne receivers typically require at least twice the electrical bandwidth of homodyne receivers, as can be clearly seen from Figure 7. While the bandwidth requirements for homodyne receivers are similar to those for direct-detection receivers (typically less than the symbol rate of the transmitted signal), heterodyne receivers need to have an electrical bandwidth of at least the entire width of the optical spectrum. Typically, the bandwidth of heterodyne receivers is substantially wider than twice the symbol rate of the transmitted signal, so as to avoid undesired signals distortion due to spectral folding of the mixing products [30]. In the example of Figure 7, the 3 dB bandwidth of the heterodyne receiver is about 3 times that of the homodyne receiver.

6.1.5 Frequency dependence

Coherent receivers in general are highly frequency dependent, just like radio or microwave homo- or heterodyne receivers. They hence need to be tuned to the frequency of the DWDM channel to be detected. The optical bandwidth of a coherent receiver is typically only as wide as the optical bandwidth of the transmitted signal [10].

Coherent receivers can thus serve as wavelength demultiplexers in DWDM systems. This feature is exploited, for example, in colourless drop ports of reconfigurable optical add-drop multiplexers (ROADMs) to select the wavelength channel of the signal to be received [30], which is simply accomplished by tuning the frequency of the LO laser to the desired WDM channel. To avoid interference from neighbouring DWDM signals, the channel spacing of the DWDM signals has to be larger than twice the bandwidth of the coherent receiver. Thus, coherent receivers based on homodyne detection allow much smaller WDM channel spacing than those using heterodyne detection, because the latter require unoccupied optical spectrum of the same width as the selected signal on the other side of the reference signal, as indicated by the hashed area in Figure 6. Hence, with heterodyne receivers, the number of useable wavelength channels is typically smaller than with homodyne receivers, thus limiting the transmission capacity of fibre optic DWDM communication systems [8].

6.1.6 Phase and polarization diversity

Early versions of coherent receivers employed only a single optical mixer to generate the desired beat signal. These simple receivers were very sensitive to polarization fluctuations in the transmitted signal and, in the case of homodyne receivers, required phase locking of the optical reference signal [1], [2]. These limitations made it quite difficult to use these receivers in commercial fibre-optic communication systems, where the fibre link usually does not preserve the state of polarization of the transmitted signals. However, it was shown that these problems can be circumvented by employing multiple mixers in a coherent receiver, as described in more detail in 6.3 and 6.4 [20].

Modern coherent receivers typically use intradyne detection and comprise four balanced coherent mixers to generate a set of four complementary beat signals, which carry information on the amplitude, phase, and polarization state of the transmitted signal [9]. These receivers are commonly referred to as phase- and polarization-diversity receivers [29]. It is important to note that these receivers facilitate the detection of complex vector-modulated signals, such as multi-level phase-shift keyed (PSK) or quadrature-amplitude modulated (QAM) signals, as well as the decoding of polarization-multiplexed signals, which can carry twice the information than conventional single-polarized signals without requiring additional optical bandwidth, as discussed in 5.3 [32]. Thus, coherent receivers have enabled a rapid increase in the transmission capacity of fibre optic communication systems [8].

However, a considerable amount of signal processing is needed to extract the transmitted information from the four beat signals, which is usually accomplished with the help of high-speed digital signal processors (DSPs), as discussed in more detail in Clause 7. The various types of coherent receivers are described in more detail in 6.2 to 6.4, beginning with single coherent mixers and ending with modern quadruple coherent mixers.

6.2 Single coherent mixer with balanced detection

6.2.1 Principle of operation

A single coherent mixer typically employs a symmetric 2 × 2 optical waveguide coupler (i.e. a 3 dB coupler), which combines the transmitted signal $\vec{A}_s(t)$ with the reference signal $\vec{A}_r(t)$, as shown schematically in Figure 8. The two output signals of the coupler are of the form $\left[\vec{A}_s(t) - j\,\vec{A}_r(t)\right]/\sqrt{2}$ and $\left[\vec{A}_r(t) - j\,\vec{A}_s(t)\right]/\sqrt{2}$, where $\vec{A}_s(t)$ is of the general form of Formula (1), and $\vec{A}_r(t)$ describes the optical amplitude of the reference signal whose polarization state is assumed to be fixed, as described by Formula (2).

$$\vec{A}_r(t) = \sqrt{P_{LO}}\begin{pmatrix}1\\0\end{pmatrix} e^{\,j\,\Phi_r(t)} = \sqrt{P_{LO}}\begin{pmatrix}1\\0\end{pmatrix} e^{\,j\,[\omega_r t + \varphi_r(t) + \hat{\varphi}_r(t)]} \tag{2}$$

where

P_{LO} is the optical power of the reference signal,

$\omega_r = 2\pi f_r$ the angular frequency of the reference signal,

$\varphi_r(t)$ is the optical phase of the reference signal, and

$\hat{\varphi}_r(t)$ is a random variable representing laser phase noise from the transmitter laser.

Subsequent mixing of the two output signals of the coupler in two separate high-speed photodiodes (PDs) generates the two photocurrents $i_1(t)$ and $i_2(t)$ given by Formula (3) [33].

$$\begin{aligned} i_1(t) &= R_1\left\{\left|\vec{A}_r(t)\right|^2 + \left|\vec{A}_s(t)\right|^2 + 2\,\mathrm{Im}\left[\vec{A}_r(t)\cdot\vec{A}_s^{\,*}(t)\right]\right\} + \hat{i}_1(t) \\ i_2(t) &= R_2\left\{\left|\vec{A}_r(t)\right|^2 + \left|\vec{A}_s(t)\right|^2 - 2\,\mathrm{Im}\left[\vec{A}_r(t)\cdot\vec{A}_s^{\,*}(t)\right]\right\} + \hat{i}_2(t) \end{aligned} \tag{3}$$

where

R_1 and R_2 represent the responsivities of the photodiodes, in units of A/W,

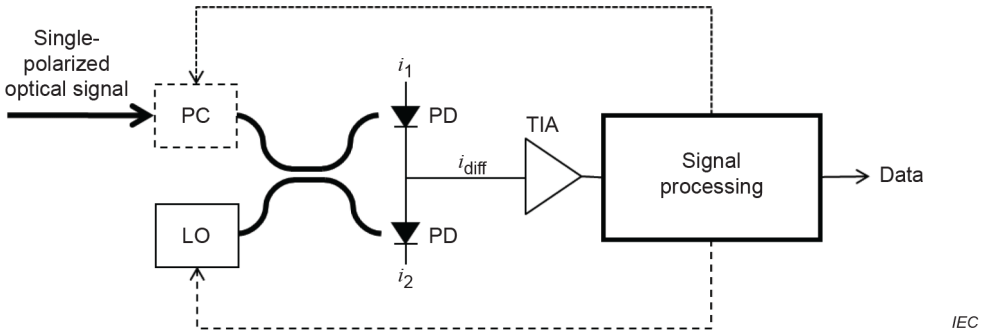
$\hat{i}_1(t)$ and $\hat{i}_2(t)$ are random variables that represent receiver noise,

$(\,)^*$ denotes the complex conjugate, and

$(\,)\cdot(\,)$ indicates the scalar vector product.

Note that R_1 and R_2 include here not only the photo-detection efficiency of the PDs but also the various optical splitting losses in the 3 dB coupler as well as optical insertion losses and attenuation in the receiver, so that the photocurrents are directly related to the input power of the received and reference signals [29].

NOTE In practice, the responsivity of each PD can be significantly different for light entering the signal input port of the mixer versus light entering the LO input port of the mixer, because of asymmetric insertion losses and additional optical elements (like optical attenuators and monitor PDs) in the signal input port. For the sake of simplicity, this difference is ignored in Formula (3).



Key

LO Local oscillator laser

PC Polarization controller

PD Photodiode

TIA Transimpedance amplifier

Figure 8 – Single balanced mixer for coherent reception

The first two terms on the right side of Formula (3) are undesired by-products of the mixing process, which result from non-coherent detection (also known as direct detection) of the reference signal and the received signal, whereas the third term represents the desired coherent beat signal. The random variables $\hat{i}_1(t)$ and $\hat{i}_2(t)$ represent receiver noise generated in the photo-detection process, including thermal and shot noise as well as input-referred amplifier noise [2], [20], [29], [34].

6.2.2 Common-mode rejection

With a perfectly balanced mixer, i.e. with $R_1 = R_2$, one can eliminate the undesired direct-detection products by subtracting the two photocurrents from each other to obtain the differential current $i_{diff}(t)$, as described by Formula (4) [20].

$$i_{diff}(t) = i_1(t) - i_2(t) = 2R_1 \ \mathrm{Im}\left[\vec{A}_r(t) \cdot \vec{A}_s^{\,*}(t)\right] + \hat{i}_{tot}(t) \tag{4}$$

This differential signal includes only the coherent beat signal and the receiver noise. The suppression of the direct detection products is known as common-mode rejection, since the direct-detection products have the same sign and magnitude in both photocurrents (common mode), whereas the coherent beat signals have opposite sign (differential mode). In practice, the common-mode rejection is limited, because of small receiver imbalance, $R_1 \neq R_2$, and differential time delays between the two photocurrents, which is known as skew. The amount of common-mode suppression in a balanced mixer is characterized by a common-mode-rejection ratio (CMRR), R_{CMR}, which is usually expressed in dB and calculated from Formula (5) [33].

$$R_{CMR} = 20 \log_{10}\left[\frac{i_1(t) - i_2(t)}{i_1(t) + i_2(t)}\right] \tag{5}$$

The CMRR can be measured by injecting light into only one of the two input ports of the mixer. In general, the signal input and the reference input yield different CMMR values, because of variations in the PD responsivities. High CMRR of the received signal is especially important for applications in colour-less ROADM ports, as discussed in 6.1.5, where the receiver sees a multitude of DWDM signals, of which only one is selected for coherent reception. In this case, all DWDM signals contribute to the direct-detection products in the photocurrents $i_1(t)$ and $i_2(t)$, as described by Formula (6).

$$i_1(t) = R_1\left[\left|\vec{A}_r(t)\right|^2 + \left|\vec{A}_s^{\lambda_1}(t)\right|^2 + \left|\vec{A}_s^{\lambda_2}(t)\right|^2 + \ldots + \left|\vec{A}_s^{\lambda_n}(t)\right|^2 + 2\,\mathrm{Im}\left\{\vec{A}_r(t)\cdot\vec{A}_s^{*\lambda_1}(t)\right\}\right] + \hat{i}_1(t)$$
$$i_2(t) = R_2\left[\left|\vec{A}_r(t)\right|^2 + \left|\vec{A}_s^{\lambda_1}(t)\right|^2 + \left|\vec{A}_s^{\lambda_2}(t)\right|^2 + \ldots + \left|\vec{A}_s^{\lambda_n}(t)\right|^2 - 2\,\mathrm{Im}\left\{\vec{A}_r(t)\cdot\vec{A}_s^{*\lambda_1}(t)\right\}\right] + \hat{i}_2(t) \quad (6)$$

where

n is the total number of DWDM signals,

$\vec{A}_s^{\lambda_m}(t)$ is the optical amplitude of the DWDM signal at wavelength λ_m, m = 1 … n, and

λ_1 is the wavelength to be coherently received.

The subtraction of the two photocurrents can be accomplished by connecting the two photodiodes in series and tapping the differential current at midpoint between the photodiodes, as shown in Figure 8. In practice, a differential transimpedance amplifier is often used to subtract the two photocurrents, as shown in Figure 9. In both cases, the losses and path lengths in the two mixer arms should be well matched, since there are usually no means for compensation of receiver imbalance or skew between the signals.

Imbalance between the two mixer arms degrades the CMRR at all frequencies, while skew introduces frequency dependence in the CMRR, degrading it more at higher frequencies than at lower frequencies. For this reason, the CMRR is often specified for different frequencies. Typical limits for the CMRR are 16 dB at low frequencies and 10 dB at high frequencies. Receivers designed to operate at symbol rates of around 25 GBd, for example, should have less than 2 ps differential time delay (or skew) between the two photodetector signals $i_1(t)$ and $i_2(t)$, which corresponds to a path length difference of less than 0,4 mm. These tight tolerances are very difficult to meet when using fibre-pigtailed optical couplers. Typically, coherent receivers for fibre-optic communication systems employ compact planar waveguide circuits for the optical splitter(s) and monolithically or hybrid integrated photodiodes that are directly connected to transimpedance amplifiers.

Furthermore, the signal-to-noise ratio (SNR) of balanced receivers is 3 dB higher than that of equivalent single-ended receivers using only one photodiode, because the amplitude of the beat signal in Formula (4) is twice as large as that in Formula (3), whereas the RMS amplitude of the total noise current in Formula (4) is only 41 % higher than that of the individual noise current in Formula (3). Assuming $\hat{i}_1(t) \approx \hat{i}_2(t)$, the total noise in the differential current is given by Formula (7).

$$\hat{i}_{\text{tot rms}} = \sqrt{\left\langle \hat{i}_1^{\,2}(t)\right\rangle + \left\langle \hat{i}_2^{\,2}(t)\right\rangle} \approx \sqrt{2}\,\hat{i}_{1\,\text{rms}} \quad (7)$$

where the brackets ‹…› indicate time averaging.

The optical power of the reference signal is typically in the range of 10 mW to 40 mW and thus much larger than the average power of the received signal, which can range from 0,01 mW to 1 mW, so that the reference signal amplifies the coherently received beat signal. As a result, coherent receivers usually have significantly higher sensitivity than direct detection receivers. In many cases, the power of the reference signal is high enough, so that the receiver noise is dominated by shot noise resulting from direct detection of the reference signal. The RMS amplitude of this shot noise, $\hat{i}_{\text{tot rms}}$, is given by Formula (8) [20].

$$\hat{i}_{\text{tot rms}} \approx \sqrt{2\,q\,R\,P_{\text{LO}}\,B} \tag{8}$$

where

q ≈1,60 × 10^{-19} C is the elementary charge,

B is the electrical bandwidth of the photodiode (in Hz), and

P_{LO} is the optical power of the reference signal from the LO laser (in W).

When the receiver noise is dominated by shot noise, the signal-to-noise ratio of the coherently received beat signal becomes independent of P_{LO}.

6.2.3 Polarization dependence

As discussed in 6.1.4, the magnitude of the coherent beat signal depends strongly on the polarization state of the received signal. It is maximal only when the received signal is in the same polarization state as the reference signal from the LO laser. The beat signal vanishes when the two signals are in orthogonal polarization states. This strong polarization dependence is very undesirable in fibre-optic communication systems, because the transmission fibres usually do not maintain the state of polarization of the transmitted signal [28], [32]. Hence, a single-polarized optical signal usually arrives at the receiver in an arbitrary polarization state, characterized by a polarization angle θ and a differential phase ψ, as described by Formula (9), where the phase parameters θ and ψ can both fluctuate randomly with time.

$$\vec{A}_{\text{s}}(t) = \sqrt{P_{\text{S}}} \begin{pmatrix} \cos(\theta) \\ \sin(\theta)\, \text{e}^{\,\text{j}\psi} \end{pmatrix} a_{\text{s}}(t)\, \text{e}^{\,\text{j}\,\Phi_{\text{s}}(t)} \tag{9}$$

The reference signal from the local oscillator laser is usually inserted into the coupler via polarization-maintaining fibre, so that it is always in a fixed and well-defined polarization state, as described by Formula (2).

For pure amplitude modulation with real-valued $a_{\text{s}}(t)$, Formula (4) becomes Formula (10).

$$i_{\text{diff}}(t) = R\,\sqrt{P_{\text{LO}}\,P_{\text{S}}}\,\cos(\theta)\; a_{\text{s}}(t)\,\sin\left[\Phi_{\text{r}}(t) - \Phi_{\text{s}}(t)\right] + \hat{i}_{\text{tot}}(t) \tag{10}$$

It is readily seen from Formula (10) that the amplitude of the coherent beat signal is proportional to cos(θ) and, hence, completely vanishes when θ = ±90°. This effect is known as polarization-dependent signal fading. It can be avoided by placing an optical polarization controller (PC) in front of the coherent mixer, as indicated in Figure 8, and by continuously adjusting this controller to maximize the amplitude of the coherent beat signal [35]. Such automatic optical polarization controllers would add considerable complexity to the coherent receiver. Instead, as described in more detail in 6.4, modern coherent receivers employ four balanced coherent mixers which are operated in parallel to detect four different polarization components simultaneously [9]. The desired beat signal can then be reconstructed from these four signal components with the help of digital signal processing. It should also be noted that single coherent mixers cannot decode polarization-multiplexed signals, which are often used to transmit signals at line rates of 100 Gbit/s and above [32].

6.2.4 Homodyne detection

Single mixers can be operated either as homodyne or as heterodyne receivers. In the case of homodyne detection, the optical frequency f_r of the reference signal is exactly equal to the carrier frequency f_s of the transmitted optical signal, so that its optical phase is locked to that of the carrier frequency [1], [2]. With homodyne detection, the simple coherent receiver of Figure 8 can detect only signals that carry either pure amplitude (or intensity) modulation (AM) or, alternatively, pure binary phase modulation, but not quadrature amplitude modulation (QAM) or other combinations of amplitude and phase modulation. In the case of pure AM, having real-valued amplitude variations $a_s(t)$ but constant phase φ_s (e.g. $\varphi_s \equiv 0$), Formula (10) reduces to Formula (11).

$$i_{\text{diff}}(t) = R\sqrt{P_{\text{LO}}\,P_{\text{S}}}\cos(\theta)\,a_{\text{s}}(t)\sin\left[\varphi_{\text{r}} + \hat{\varphi}_{\text{tot}}(t)\right] + \hat{i}_{\text{tot}}(t) \tag{11}$$

where

$\hat{\varphi}_{\text{tot}}(t) = \hat{\varphi}_{\text{r}}(t) - \hat{\varphi}_{\text{s}}(t)$ is the combined phase noise of the transmitter and local oscillator lasers.

The beat signal of Formula (11) is proportional to the time-varying amplitude of the transmitted signal, as desired, but it is also proportional to sin (φ_r), so that its magnitude depends on the relative optical phase between the reference signal and the optical carrier of the transmitted signal. The beat signal is maximal only when φ_r = 90° (or alternatively –90°) and vanishes when φ_r = 0° or 180°. Thus, homodyne detection with a single balanced mixer requires active locking of the optical phase of the local oscillator phase, which is quite difficult to accomplish in practice. However, the single balanced mixer is capable of detecting arbitrary AM, in particular higher-order pulse-amplitude modulation (PAM).

The transmitter and local oscillator lasers used in a coherent transmission system should exhibit low phase noise, so that $\left|\hat{\varphi}_{\text{tot}}(t)\right|$ is sufficiently small and does not introduce significant additional amplitude variations in the beat signal, which would distort the received signal [36]. In general, coherent detection requires transmitter and local oscillator lasers with narrow linewidth [10]. The amount of acceptable phase noise depends on the type and cardinality of the modulation format as well as on the symbol rate of the transmitted signal [37]. In general, the higher the cardinality of the modulation format, the less phase noise from the transmitter and LO lasers can be tolerated. Moreover, the higher the symbol rate, the more phase noise can be tolerated from the transmitter laser, but not necessarily from the LO laser [38].

NOTE There are additional (and often stricter) requirements for the phase noise of the LO laser when chromatic dispersion is compensated in the coherent receiver, as described in more detail in 6.3.3 and 6.4.6. As a result, the tolerance to LO laser phase noise typically decreases with increasing symbol rate.

Similar considerations apply to purely phase modulated signals, which are of the form shown in Formula (12).

$$\vec{A}_{\text{s}}(t) = \sqrt{P_{\text{S}}}\begin{pmatrix}\cos(\theta)\\ \sin(\theta)\,e^{j\psi}\end{pmatrix} \mathrm{e}^{\,\mathrm{j}\,\varphi_{\text{s}}(t)}\,\mathrm{e}^{\,\mathrm{j}\left[\omega_{\text{s}}t + \hat{\varphi}_{\text{s}}(t)\right]} \tag{12}$$

In this case Formula (10) becomes Formula (13).

$$i_{\text{diff}}(t) = R\sqrt{P_{\text{LO}}\,P_{\text{S}}}\cos(\theta)\sin\left[\varphi_{\text{s}}(t) - \varphi_{\text{r}} + \hat{\varphi}_{\text{tot}}(t)\right] + \hat{i}_{\text{tot}}(t) \tag{13}$$

In contrast to the case of AM, the phase modulation cannot be arbitrary. A single balanced mixer is capable of detecting binary phase-shift keying (BPSK), where $\varphi_s(t)$ is toggled between only two phases, for example 0° and 180° (assuming φ_r = 90°, as above), but it cannot decode quaternary phase-shift keying (QPSK), where $\varphi_s(t)$ is toggled between the four phases 0°, 90°, 180° and 270°, since $i_{diff}(t)$ would be zero when $\varphi_s(t)$ is 90° and 270°. This is a direct result of the spectral folding that occurs in single homodyne mixers (see 6.1.2).

6.2.5 Heterodyne detection

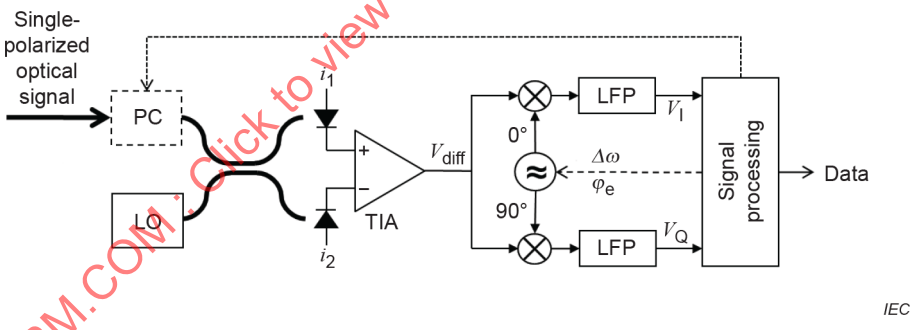
In the case of heterodyne detection with $\omega_s \neq \omega_r$, the transmitted signal can carry arbitrary AM and/or PM, and thus arbitrary QAM, so that the received differential photocurrent is given by Formula (14).

$$i_{diff}(t) = R\sqrt{P_{LO}\,P_S}\,\cos(\theta)\,a_s(t)\,\sin\left[\Delta\omega\,t + \varphi_r - \varphi_s(t) + \hat{\varphi}_{tot}(t)\right] + \hat{i}_{tot}(t) \quad (14)$$

where

$\Delta\omega = \omega_r - \omega_s$ denotes the frequency offset between the reference signal and the carrier frequency of the transmitted signal.

The phase and amplitude information encoded in the transmitted signal can be recovered, for example, by down-mixing the output signal of the transimpedance amplifier, $V_{diff}(t) = Z\,i_{diff}(t)$, simultaneously with the in-phase and quadrature outputs of a sinusoidal oscillator at frequency $\Delta\omega$, i.e. by down-mixing with the electrical signals $\sin(\Delta\omega t + \varphi_e)$ and $\cos(\Delta\omega t + \varphi_e)$, as shown schematically in Figure 9 [1].



Key

LO Local oscillator laser
LPF Low-pass filter
PC Polarization controller
TIA Transimpedance amplifier

Figure 9 – Balanced heterodyne mixer with electrical down-mixing

With $\varphi_e = \varphi_r - 90°$ and after low-pass filtering, to reject all frequency components above $\Delta\omega$, one obtains the quadrature signals $V_I(t)$ and $V_Q(t)$ described by Formula (15).

$$V_{\mathrm{I}}(t) = \frac{Z\,R\,\sqrt{P_{\mathrm{LO}}\,P_{\mathrm{S}}}\,\cos(\theta)}{2}\,a_{\mathrm{s}}(t)\,\sin\left[\,\varphi_{\mathrm{s}}(t) - \hat{\varphi}_{\mathrm{tot}}(t)\right] + \hat{V}_1(t)$$
$$V_{\mathrm{Q}}(t) = \frac{Z\,R\,\sqrt{P_{\mathrm{LO}}\,P_{\mathrm{S}}}\,\cos(\theta)}{2}\,a_{\mathrm{s}}(t)\,\cos\left[\,\varphi_{\mathrm{s}}(t) - \hat{\varphi}_{\mathrm{tot}}(t)\right] + \hat{V}_2(t) \qquad (15)$$

where

$\hat{V}_1(t)$ and $\hat{V}_2(t)$ represent the receiver noise after down-mixing and filtering, and

Z is the transimpedance of the TIA (in units of V/A or Ω).

The two signals $V_{\mathrm{I}}(t)$ and $V_{\mathrm{Q}}(t)$ comprise all information on the amplitude and phase modulation in the transmitted signals, which can be seen by combining the two signals into a complex vector as shown in Formula (16).

$$V_{\mathrm{I}}(t) + \mathrm{j}V_{\mathrm{Q}}(t) = \frac{Z\,R\,\sqrt{P_{\mathrm{LO}}\,P_{\mathrm{S}}}\,\cos(\theta)}{2}\,a_{\mathrm{s}}(t)\,\mathrm{e}^{\,\mathrm{j}\left[\varphi_{\mathrm{s}}(t) - \hat{\varphi}_{\mathrm{tot}}(t)\right]} + \hat{V}_{\mathrm{noise}}(t)\,\mathrm{e}^{\,\mathrm{j}\,\hat{\varphi}_{\mathrm{noise}}(t)} \qquad (16)$$

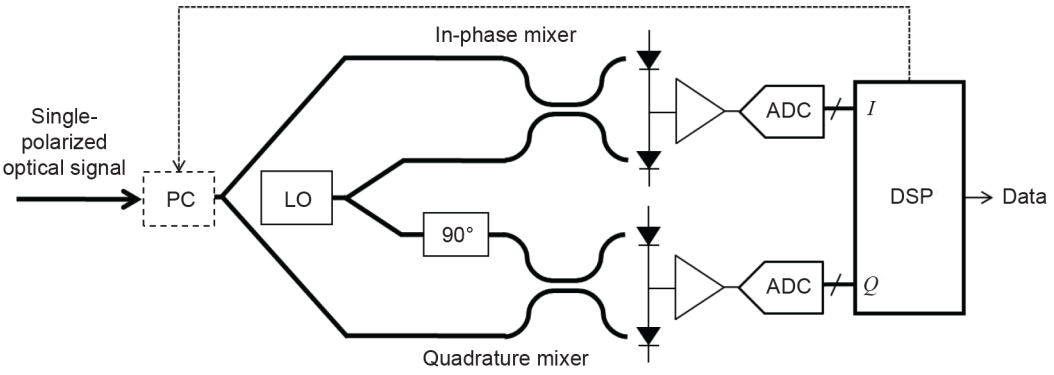
Aside from the additional noise terms on the rightmost side, Formula (16) resembles the complex phasor of the transmitted signal given in Formula (9). The down-mixing and filtering can be accomplished with the help of analogue electronics, as shown in Figure 9, or equivalently in a high-speed digital signal processor, after the analogue voltage $V_{\mathrm{diff}}(t)$ is converted into a digital signal by a high-speed analogue-to-digital converter (ADC).

Hence, the simple heterodyne receiver of Figure 9 allows decoding of arbitrary M-QAM signals, without requiring phase-locking of the LO laser. Any drift in the frequency and phase of the transmitter and LO lasers can be compensated for by adaptive adjustment of the frequency $\Delta\omega$ and phase φ_{e} of the electrical oscillator in the receiver. However, heterodyne receivers require PDs and TIAs with at least twice the electrical bandwidth of a comparable homodyne receiver, which can limit the symbol rate of the transmitted signal [30]. In most applications, a polarization controller is required to avoid polarization-induced signal fading in the heterodyne receiver [35]. However, it should be noted that polarization-fading can also be avoided by encoding the information data into a dual-polarization signal, using for example Alamouti coding, in such a way that the coherently received beat signal does not depend on the polarization state of the transmitted signal at the receiver [39].

6.3 Dual coherent mixer with phase diversity

6.3.1 Principle of operation

The capabilities of coherent receivers can be greatly enhanced by employing multiple mixers in parallel, so as to recover more properties of the received optical signal. As a first step, in 6.3 a coherent receiver with phase diversity is considered. Such phase-diversity receivers typically employ two balanced mixers, an “in-phase mixer” and a “quadrature” mixer, which are operated in parallel [20], [29]. Either of the two mixers is similar to the single balanced mixer of Figure 8, but the quadrature mixer includes an additional optical phase shifter that introduces a 90° optical phase shift in the reference signal, as shown in Figure 10. This 90° phase shift can be generated, for example, by adding a suitable length of additional optical waveguide in the second mixer, so as to delay the reference signal by an amount that corresponds to a 90° shift in optical phase.



Key

90°	Optical phase shifter
ADC	Analogue-to-digital converter
DSP	Digital signal processor
LO	Local oscillator laser
PC	Polarization controller

Figure 10 – Dual coherent mixer with phase diversity

The transmitted signal and the reference signal from the LO laser are both split into two equal copies, which are then fed symmetrically into two parallel mixers, as shown in Figure 10. Hence, the in-phase mixer combines the transmitted signal with the 0°-shifted copy of the reference signal, whereas the quadrature mixer combines the transmitted signal with the 90° phase-shifted copy of the reference signal, thus creating a coherent beat signal whose electrical phase is always shifted by 90° relative to the beat signal created by the in-phase mixer (phase quadrature). This combination of in-phase and quadrature mixer is often referred to as a 90° hybrid mixer, a term that was adopted from RF and microwave technologies [20].

It should be noted that the arrangement shown in Figure 10 is not the only way to implement a phase-diversity receiver. If the transmitted signal and the reference signal are combined in a three-branch equal-split directional coupler, for example, the three output ports of this coupler produce beat signals with relative phase shifts of 0°, 120°, and 240° [20], [29]. However, most practical applications employ balanced 90° hybrid mixers.

Assuming homodyne reception with $\Delta\omega$ = 0 and φ_r = 0, as well as otherwise identical path lengths in the two mixers, the in-phase and quadrature mixers in Figure 10 generate two phase shifted photocurrents $i_I(t)$ and $i_Q(t)$, which are given in Formula (17) [1], [29].

$$\begin{aligned} i_I(t) &= R\sqrt{P_{LO}P_S}\cos(\theta)\ a_s(t)\ \sin\left[-\varphi_s(t)+\hat{\varphi}_{tot}(t)\right]+\hat{i}_{tot}(t) \\ i_Q(t) &= R\sqrt{P_{LO}P_S}\cos(\theta)\ a_s(t)\ \cos\left[-\varphi_s(t)+\hat{\varphi}_{tot}(t)\right]+\hat{i}_{tot}(t) \end{aligned} \quad (17)$$

These two photocurrents have the same form as the two down-mixed heterodyne signals of Formula (15). Therefore, a dual homodyne mixer with phase diversity provides the same functionality as a single heterodyne mixer. Hence, it allows detection of arbitrary M-QAM signals (including QPSK) without needing the wide electrical bandwidth of a heterodyne receiver. However, splitting of the reference and transmitted signals between the two mixers reduces the effective responsivity R by 50 %, compared with Formula (10), thus yielding substantially smaller photocurrents in each mixer than for the single mixer in Formula (10). The amplitude of each photocurrent can be raised to the signal level of a single mixer by increasing the power of the reference signal and the transmitted signal both by 3 dB or, alternatively, by increasing the

power of only the reference signal by 6 dB, where the latter would increase the signal-to-noise ratio (SNR) of the two signals in Formula (17).

When used with ASK or BPSK signals, the phase-diversity mixer of Figure 10 introduces more noise in the photocurrents than the single balanced mixer of Figure 8 [27], [29]. This can be readily seen in the case of homodyne detection with a phase-locked LO laser, as discussed in 6.2.4. If the reference phase φ_r is adjusted in such a way that only $i_I(t)$ carries the desired beat signal, then $i_Q(t)$ only carries receiver noise (e.g. shot noise), which is an undesired addition to the received signal. However, this disadvantage vanishes when quadrature modulated signals (such as QPSK or M-QAM) are received [27].

However, the phase-diversity mixer of Figure 10 is much more versatile than a single mixer, because it allows homodyne detection with small frequency offsets between the reference signal and the transmitted signal as well as compensation of I-Q skew and chromatic dispersion, as explained in 6.3.2 through 6.3.5.

6.3.2 Intradyne detection with frequency offset removal

In contrast to the single mixer discussed in 6.2, the phase-diversity mixer of Figure 10 allows homodyne detection without requiring optical phase locking of the LO laser to the optical carrier of the transmitted signal [29]. In fact, the LO laser does not even have to be operated at exactly the same frequency as the transmitter laser, because the two signals $i_I(t)$ and $i_Q(t)$ carry the information needed to remove undesired frequency and phase offsets [2]. This mode of operation is commonly referred to as intradyne reception.

To demonstrate this capability of the phase-diversity mixer, the parameters $\Delta\omega$ and φ_r are used to describe the frequency and phase offsets between the reference signal and the transmitted signal, so that the two photocurrents of Formula (17) become $\tilde{i}_I$ and $\tilde{i}_Q$ as described by Formula (18).

$$\begin{aligned}\tilde{i}_I(t) &= R\sqrt{P_{LO}P_S}\cos(\theta)\,a_s(t)\;\sin\left[\Delta\omega t+\varphi_r-\varphi_s(t)+\hat{\varphi}_{tot}(t)\right]+\hat{i}_{tot}(t)\\ \tilde{i}_Q(t) &= R\sqrt{P_{LO}P_S}\cos(\theta)\,a_s(t)\;\cos\left[\Delta\omega t+\varphi_r-\varphi_s(t)+\hat{\varphi}_{tot}(t)\right]+\hat{i}_{tot}(t)\end{aligned} \tag{18}$$

If the frequency offset $\Delta\omega$ and the phase offset φ_r are known or can be measured, the beat signals of Formula (18) can be transformed into the form of Formula (17) by applying the matrix multiplication described by Formula (19).

$$\begin{bmatrix} i_I(t)\\ i_Q(t)\end{bmatrix}=\begin{bmatrix}\sin\left(\Delta\omega t+\varphi_r\right), & \cos\left(\Delta\omega t+\varphi_r\right)\\ \cos\left(\Delta\omega t+\varphi_r\right), & -\sin\left(\Delta\omega t+\varphi_r\right)\end{bmatrix}\times\begin{bmatrix}\tilde{i}_I(t)\\ \tilde{i}_Q(t)\end{bmatrix} \tag{19}$$

This operation can be performed in real time, using either analogue multipliers or, alternatively, a high-speed digital signal processor (DSP) when the two beat signals are digitized via high-speed ADCs [40]. However, since $\Delta\omega$ and φ_r can fluctuate with time, especially when independent lasers are used for the transmitter and local oscillator, it is usually necessary to continuously adjust the parameters in Formula (19) to completely remove the frequency and phase offsets [40].

The frequency and phase offsets are usually not known a priori, but they can be estimated in the DSP using various methods [41], [42]. Relatively large frequency offsets, for example, can be estimated from a spectral analysis of the two photocurrents $\tilde{i}_{\mathrm{I}}(t)$ and $\tilde{i}_{\mathrm{Q}}(t)$ by Fourier transforming the complex signal $\tilde{i}_{\mathrm{c}}(t) = \tilde{i}_{\mathrm{I}}(t) + \mathrm{j}\,\tilde{i}_{\mathrm{Q}}(t)$, which yields a two-sided electrical spectrum $\tilde{S}_{\mathrm{c}}(\omega)$ of the coherent beat signal.

Such a Fourier transform is often calculated in DSPs when chromatic dispersion is to be compensated, as discussed in more detail in 6.3.3. The frequency offset estimation can thus be performed in the same signal processing step. However, compensation of large frequency offsets should be performed before chromatic dispersion is compensated (see 7.2).

For most modulation formats, the transmitted signal spectrum is symmetric and centred at zero when $\Delta\omega$ = 0. Hence, if a frequency offset $\Delta\omega$ is present, it can be estimated from the position of the central frequency of the spectrum, as shown schematically in Figure 11 [43]. Another method for coarse frequency offset estimation involves calculation of the time-varying phase changes in $\tilde{i}_{\mathrm{c}}(t)$ [41].

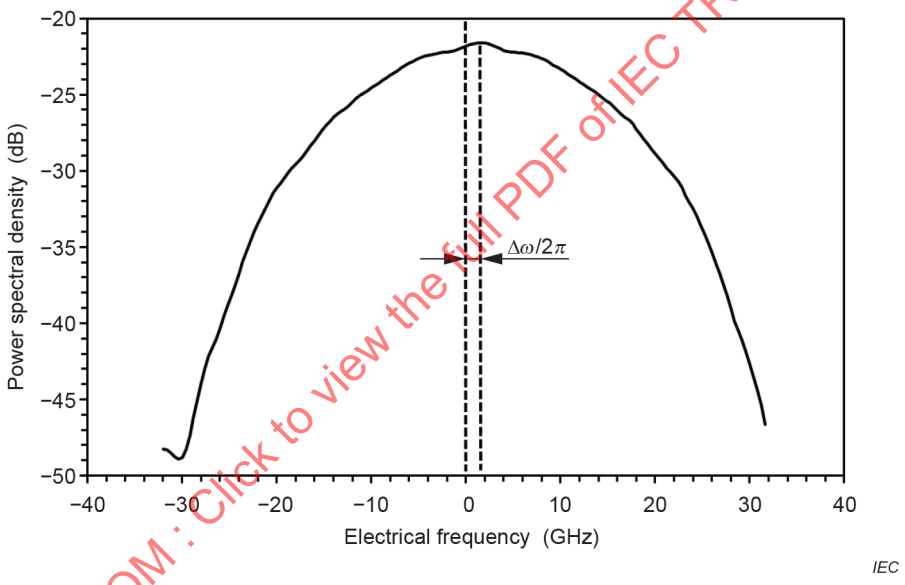


Figure 11 – Intradyne beat spectrum with 2 GHz frequency offset

Likewise, there are several methods to estimate time-varying phase offsets in $\tilde{i}_{\mathrm{c}}(t)$ (including small frequency offsets), most of which strongly depend on the particular modulation format used for signal transmission [42]. Phase offset removal is usually referred to as carrier phase recovery [44]. An exemplary method for carrier phase recovery is described in 6.4.5.

With phase-diversity receivers, it is thus possible to use independently frequency-stabilized lasers for the transmitter and local oscillator, like integrable tuneable laser assemblies (ITLAs), which are frequently employed in coherent transmission systems [19]. These lasers typically have a frequency accuracy of better than ±2,5 GHz, so that the worst-case frequency difference between the transmitter and local oscillator lasers will not exceed 5 GHz. This offset is small compared with the total bandwidth of a 25 GBd optical signal.

6.3.3 Compensation of chromatic dispersion

Fibre optic communication links often exhibit considerable amounts of chromatic dispersion (CD), especially those employing dispersion-unshifted single-mode fibres, such as B-652, B-654, and B-657 fibres [45]. Chromatic dispersion causes optical signals at different frequencies to propagate at different speeds through the fibre optic link. When the amount of group velocity dispersion (GVD) is large enough, the frequency-dependent propagation speed can introduce significant differential time delays between the various frequency components of a modulated single-carrier optical signal, which in turn can cause severe waveform distortions in the transmitted signal [46]. Formula (20) shows how to calculate the CD-induced differential time delay Δt_{CD} between two frequency components of angular frequency difference $\Delta\omega_s$ [46].

$$\Delta t_{CD}\left(\Delta\omega_s\right) = \frac{2\pi D_L \, c \, \Delta\omega_s}{{\omega_s}^2} \tag{20}$$

where

D_L denotes the total GVD which the signal has experienced (in units of s/m),

c is the speed of light in vacuum (approximately 299 792 458 m/s), and

ω_s is the angular carrier frequency of the signal (in units of rad/s).

When the frequency-dependent differential time delays Δt_{CD} are large enough, they introduce undesired differential phase shifts across the signal spectrum. This is shown in Figure 12 for the example of D_L = 3 000 ps/nm, which is approximately the GVD in about 180 km of dispersion-unshifted B-652 fibre at 1,5 µm wavelength [45]. For comparison, the grey curve in Figure 12 shows the spectral shape of a 128 Gb/s PM-QPSK signal that has been filtered with a root-raised-cosine function having a roll-off factor 0,5 [12]. It is clearly seen from Figure 12 that the differential phase shifts at the edges of the optical spectrum increase rapidly with increasing bandwidth of the optical signal spectrum and, consequently, with increasing modulation rate of the transmitted signal.

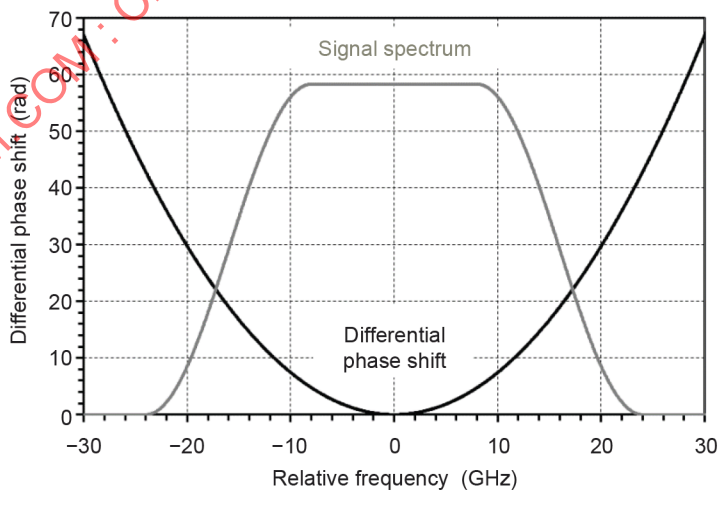


Figure 12 – Differential phase shifts introduced by 3 000 ps/nm GVD

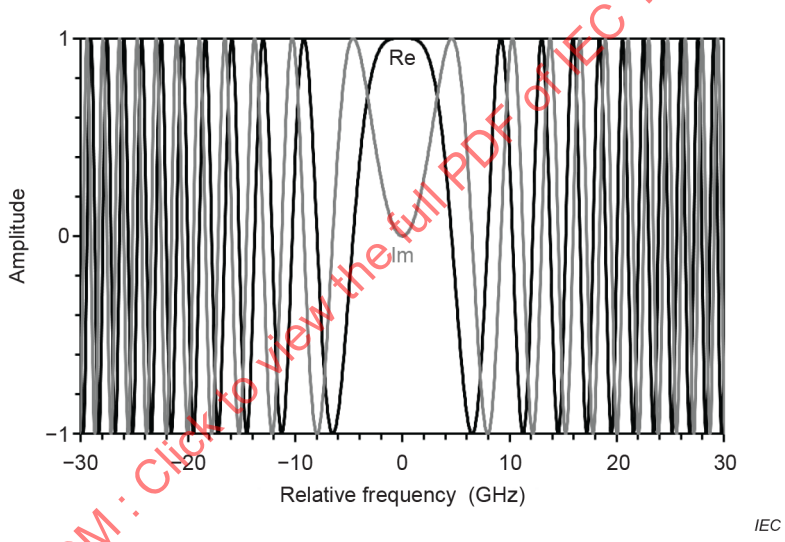
The effects of CD on a modulated signal can be described in the frequency domain by a complex transfer function $H(\omega_s + \Delta\omega_s)$, which is given in Formula (21) [48].

$$H\left(\omega_s + \Delta\omega_s\right) = \exp\left\{ j \frac{\pi D_L c \, \Delta\omega_s^2}{\omega_s^2} \right\} \tag{21}$$

where

$\Delta\omega_s$ is the angular frequency of a spectral component relative to the carrier frequency ω_s

The transfer function $H(\omega_s + \Delta\omega_s)$ is a rapidly oscillating function, whose oscillation frequency increases with increasing frequency offset $\Delta\omega_s$ as well as with increasing D_L [10]. These oscillations are shown in Figure 13 for the example of D_L = 3 000 ps/nm around an optical frequency of 195 THz. The frequency of these oscillations increases rapidly with increasing frequency offset $\Delta\omega_s$.



Key

Re Real part

Im Imaginary part

Figure 13 – Transfer function for GVD of 3 ns/nm in frequency domain

Since the differential phase shifts Δt_{CD} are largest at the edges of the optical signal spectrum (see Figure 12), the CD-induced waveform distortions increase rapidly with increasing modulation frequency.

However, with a phase-diversity mixer it is possible to remove these undesired differential phase shifts by introducing similar phase shifts of opposite sign in the complex beat signal $i_c(t) = i_I(t) + j\ i_Q(t)$. This can be accomplished by multiplying the spectral components $S_c(\Delta\omega_s)$ of $i_c(t)$ with the transfer function $H^{-1}(\Delta\omega_s)$ given in Formula (22), which is the inverse function of Formula (21) [48].

$$H^{-1}\left(\Delta\omega_{\mathrm{s}}\right)=\exp\left\{-\mathrm{j}\,\frac{\pi D_{\mathrm{L}}\,\mathrm{c}\,\Delta\omega_{\mathrm{s}}{}^{2}}{\omega_{\mathrm{s}}{}^{2}}\right\} \tag{22}$$

This procedure requires a Fourier transformation of $i_{\mathrm{c}}(t)$ into the frequency domain and, after multiplication with $H^{-1}(\Delta\omega_{\mathrm{s}})$, an inverse a Fourier transformation back into the time domain [4]. Alternatively, one can remove the undesired effects of CD in the time domain, where the spectral transfer function $H^{-1}(\omega)$ corresponds to a convolution of the complex beat signal $i_{\mathrm{c}}(t)$ with the transfer function $h^{-1}(t)$ given in Formula (23) [49].

$$h^{-1}(t)=\sqrt{\frac{\omega_{\mathrm{s}}{}^{2}}{2\pi\left|D_{\mathrm{L}}\right|\mathrm{c}}}\exp\left\{\mathrm{j}\,\frac{\omega_{\mathrm{s}}{}^{2}t^{2}}{4\pi D_{\mathrm{L}}\,\mathrm{c}}\right\} \tag{23}$$

The convolution of $i_{\mathrm{c}}(t)$ with $h^{-1}(t)$ is given by Formula (24).

$$\left[h^{-1}*i_{\mathrm{c}}\right](t)\equiv\int_{-\infty}^{\infty}h^{-1}(\tau)\,i_{\mathrm{c}}(t-\tau)\,d\tau\approx\int_{-\Delta t_{\max}}^{\Delta t_{\max}}h^{-1}(\tau)\,i_{\mathrm{c}}(t-\tau)\,d\tau \tag{24}$$

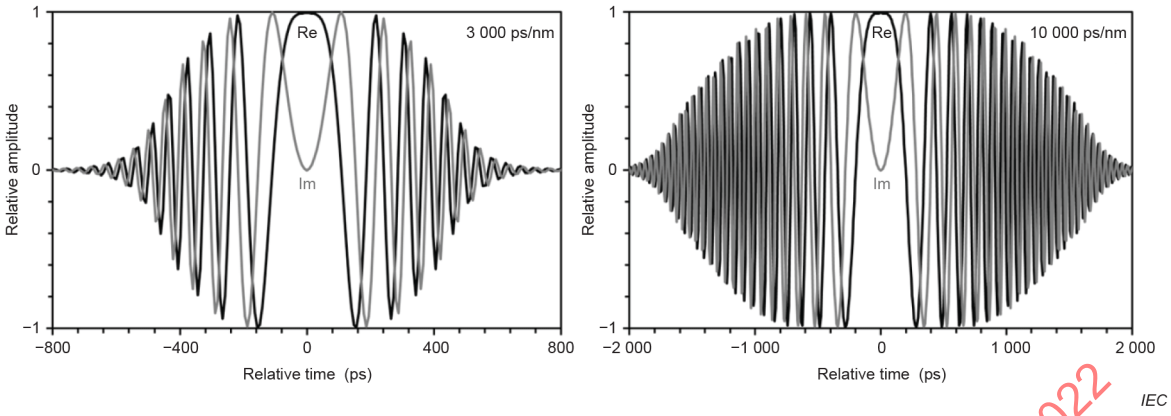
where

$\left[h^{-1}*i_{\mathrm{c}}\right](t)$ denotes the convolution of the function $h^{-1}(t)$ with the function $i_{\mathrm{c}}(t)$.

The complex transfer function $h^{-1}(t)$ also exhibits rapid oscillations, whose frequency increases rapidly with time, as depicted in Figure 14, but decreases with increasing GVD D_{L} [50]. The integration range of the convolution in Formula (24) should be confined to the maximal differential time delays that are introduced in the optical spectrum of the transmitted signal, i.e. to the interval $-\Delta t_{\max}$ to $+\Delta t_{\max}$, with $\Delta t_{\max} = D_{\mathrm{L}}\,(\lambda_{\max} - \lambda_{\min})/2$, where $\lambda_{\max}$ and $\lambda_{\min}$ denote the smallest and largest wavelengths at the ends of the signal spectrum [51].

Hence, the larger D_{L}, the longer $\Delta t_{\max}$ has to be, and the more transmitted symbols are involved in the convolution. This can be seen from Figure 14, which displays the real and imaginary parts of $h^{-1}(t)$ tailored to compensate 3 000 ps/nm and 10 000 ps/nm GVD in a 128 Gb/s PM-QPSK signal, which has the spectrum of Figure 12. In the case of 3 000 ps/nm GVD the correlation time is $2\Delta t_{\max} \approx 1\,200$ ps, whereas for 10 000 ps/nm it is $2\Delta t_{\max} \approx 4\,000$ ps. Thus, a GVD of 10 000 ps/nm correlates the amplitudes of about 128 symbols in the 128 Gb/s PM-QPSK signal.

Note that the number of correlated bits increases with the square of the symbol rate. In a similarly filtered 256 Gb/s PM-QPSK signal, therefore, a GVD of 10 000 ps/nm would correlate the amplitudes of about 512 symbols, which is four times the number of symbols correlated in a 128 Gb/s signal, because of a twofold increase in $2\Delta t_{\max}$ and a twofold increase in symbol rate.



Key

Re Real part

Im Imaginary part

Figure 14 – Inverse transfer functions for GVD of 3 ns/nm and 10 ns/nm in time domain

In Figure 14, the amplitudes of $h^{-1}(t)$ are gradually tapered towards zero at $-\Delta t_{max}$ and $+\Delta t_{max}$. This tapering (or apodization) is needed to avoid aliasing of the high-frequency oscillations in $h^{-1}(t)$ which would otherwise occur above the Nyquist frequency of the digitally sampled signals [49].

It is thus possible to compensate the CD-induced differential phase shifts electronically in the received beat signal [53]. However, this electronic dispersion compensation requires LO lasers with very low phase noise, especially when the amount of GVD is large, as explained in more detail in 6.4.6. In the presence of CD, the coherent beat signal $i_c(t)$ can be expressed as shown in Formula (25), where a transmitted signal of the form of Formula (9) is assumed as well as $\varphi_r = 90°$.

$$i_c(t) = R\sqrt{P_{LO}\,P_S}\,\cos(\theta)\,e^{j\left[\Delta\omega t+\hat{\varphi}_r(t)\right]}\int_{-\Delta t_{max}}^{\Delta t_{max}} h(\tau)\,a_s(t-\tau)\,e^{j\left[\varphi_s(t-\tau)+\hat{\varphi}_s(t-\tau)\right]}\,d\tau + \hat{i}_{tot}(t) \quad (25)$$

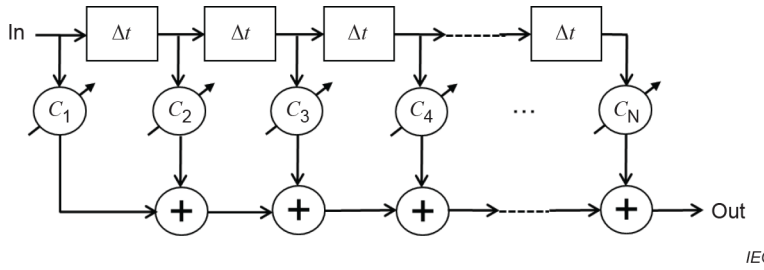
After removal of the frequency offset $\Delta\omega$ between the transmitter and LO lasers, as described in 6.3.2, and with sufficiently small LO laser phase noise $\hat{\varphi}_r(t) \approx 0$, the original transmitted signal can be restored by performing the convolution of Formula (24), which yields Formula (26).

$$i_{comp}(t) = \left[h^{-1} * i_c\right](t) = R\sqrt{P_{LO}\,P_S}\,\cos(\theta)\;a_s(t)\,e^{j\left[\varphi_s(t)+\hat{\varphi}_s(t)\right]} + \hat{i}_{tot}(t) \quad (26)$$

The compensated signal $i_{comp}(t)$ in Formula (26) is very similar to the undistorted signal described by Formula (17).

Random or periodic variations of the LO laser phase that occur during the integration time $2\Delta t_{max}$ in Formula (24) can distort the deconvolution of GVD in the signal and, hence, prevent complete compensation of the CD-induced differential phase shifts [38]. The larger the accumulated GVD in the signal, the larger $2\Delta t_{max}$ becomes and, hence, the longer the LO laser

phase has to remain stable. Since Δt_{max} also increases with increasing symbol rate of the transmitted signal, the maximal tolerable phase noise of the LO laser generally decreases with increasing symbol rate [52]. Hence, compensation of large amounts of GVD typically requires an LO laser with very low phase noise, as discussed in more detail in 6.4.6. It should be noted that phase noise in the transmitter laser does not affect electronic CD compensation, but could impact the signal decoding, as described in 6.4.5.



Key

Δt Time delay

$c_1 \ldots c_N$ Complex-valued tap coefficients

Figure 15 – Fractionally spaced equalizer with a tapped delay line

The convolution of Formula (24) can be accomplished with the help of a tapped delay line having a finite number of taps with variable complex-valued tap coefficients c_n, where n = 1, …, N [48]. This type of finite-impulse response (FIR) filter, shown in Figure 15, is well known from analogue and digital radio-frequency communication systems, where it is used to equalize frequency and phase response of the transmission system as well as frequency-dependent time delays, such as skew and dispersion. This equalizer is known as a transversal electrical filter (TEF) as well as a fractionally spaced equalizer (FSE), where the latter term is more frequently used in optical communication systems [48]. Sometimes it is simply referred to as a time-domain equalizer (TDE) or FIR filter [49], [51].

An FSE typically comprises an even number N of concatenated delay elements, each introducing the same time delay Δt, followed by variable signal taps, as shown in Figure 15. The various delayed copies of the input signal $S_{in}(t)$ are then weighted (i.e. attenuated or amplified) with coefficients c_m, m = 1, …, N + 1, which can be either real- or complex-valued, and finally combined into a single output signal of the form $S_{out}(t) = c_1 S_{in}(t) + c_2 S_{in}(t + \Delta t) + \ldots + c_N S_{in}(t + [N - 1] \Delta t)$. Such FSE can be implemented to have complex-valued coefficients c_m as well as complex-valued input signals, so that it can then emulate the operation of Formula (24), using $i_c(t)$ as the input signal.

For CD compensation, the complex-valued coefficients c_m for the FSE can be calculated analytically from Formula (27) [46], [47].

$$c_m = \frac{\Delta t}{2\pi} \int_{-\pi/\Delta t}^{\pi/\Delta t} H^{-1}(\Delta\omega_s)\, e^{-j\, n\, \Delta\omega_s \Delta t}\, d\,\Delta\omega_s \qquad (27)$$

where

$H^{-1}(\Delta\omega_s)$ is the inverse transfer function of CD from Formula (22), and

$n = m - N/2 - 1$.

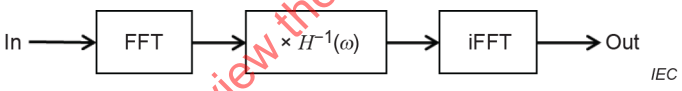
Alternatively, the coefficients c_m can be calculated from Formula (23) [49]. The amount of GVD that an FSE can compensate is limited by the total time delay of the N delay elements, which has to be at least $2\Delta t_{max}$ [51]. Digital signal processors usually set the delays Δt equal to the sampling period of the digitized received signal, which is a suitable fraction of the symbol period T_s of the transmitted signal. The number of required delay elements is then determined by the sampling rate, which is sometimes chosen to be twice the symbol rate R_s, so that $\Delta t = T_s/2$.

Compensation of 8 ns/nm GVD in a 32 GBd QPSK signal with a symbol period of 31,25 ps and an optical bandwidth of 50 GHz, for example, requires an FSE with a total delay time of at least 3,2 ns, which is equivalent to about 102,4 symbol periods or at least 205 delay elements with time delay $\Delta t = T_s/2$. Likewise, compensation of 80 ns/nm GVD, which is approximately the dispersion of 4 000 km of dispersion-unshifted fibre at 1,5 µm wavelength, requires at least 2 048 delay elements.

Therefore, compensation of such large amounts of GVD with an FSE would require a fairly large number of multiplications in the DSP. It turns out that it is computationally more efficient to compensate large amounts of chromatic dispersion in the frequency domain [4], [51]. This is achieved by

– first applying a fast Fourier transform (FFT) to $i_c(t)$ to obtain the spectral components of $i_c(t)$ in the frequency domain; then
– multiplying the resulting spectrum with $H^{-1}(\omega)$ of Formula (22); and
– finally applying an inverse FFT (iFFT) to transform the compensated spectrum back into the time domain [51].

This procedure is shown schematically in Figure 16.



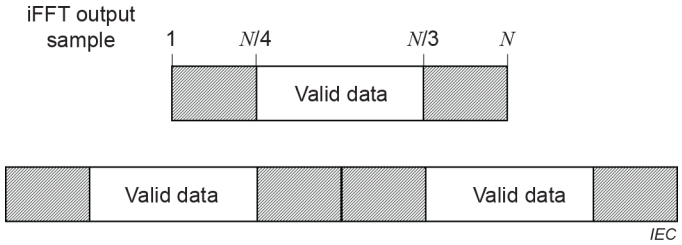
Key

FFT Fast Fourier transform

iFFT Inverse FFT

Figure 16 – Processing steps for CD compensation in the frequency domain

The number of samples used in the FFT should be large enough to cover a time period of at least $4\Delta t_{max}$, which is twice as long as the total delay required in a FSE [4]. The reason for this extended time period is that FFTs assume the signal to be periodic in time with a period equal to the number of samples to which the FFT is applied, so that a relatively large number of samples at the beginning and end of the corrected time samples are calculated under this invalid condition. These samples should be discarded after the procedure is completed.



Key

iFFT Inverse fast Fourier transform

Figure 17 – Data processing for CD compensation in the frequency domain

In the above example of a 32 GBd QPSK signal having experienced 80 ns/nm GVD and being sampled at twice the symbol rate, the FFT should include at least N = 4 096 data samples. After CD compensation and inverse FFT, the first 1 024 and the last 1 024 data samples should be discarded, yielding only 2 048 valid samples. To obtain a continuous stream of compensated signal samples, it is thus necessary to perform the same procedure on an additional set of samples having a 25 % overlap with the previous set of samples, as shown in Figure 17. For best efficiency, the number of samples in each set should be a power of 2, i.e. $N = 2^m$, with m being a suitable integer. If the GVD in the communication system is known to be limited by a certain value, the number m can be chosen to yield the minimal number of samples with which this GVD value can be compensated.

Although computationally more efficient, CD compensation in the frequency-domain introduces additional latency, because of the two Fourier transforms FFT and iFFT [4]. However, the transformation into the frequency domain facilitates compensation of other impairments, such as estimation and removal of potentially large frequency offsets $\Delta\omega$ as well as frequency-dependence in the receiver front end with long correlation times (like those caused by double reflections of the signal).

Excessive frequency offset should be removed from the signal before CD is compensated by multiplication with $H^{-1}(\omega)$, because the mathematical operations of frequency shifting and CD compensation are not commutable [4]. In general, signal impairments should be compensated in the reverse order of how they were introduced, as discussed in Clause 7. Since the optical signal first experienced CD in the fibre link and subsequently the frequency offset in the coherent receiver, the frequency offset should be compensated first and the CD thereafter.

The GVD in a communication link usually is not known accurately enough to determine $H^{-1}(\omega)$ needed for complete CD compensation. Therefore, a CD estimator is often employed in the DSP to determine the amount of GVD that the signal has experienced in the transmission link. This estimated GVD value is then used to calculate $H^{-1}(\omega)$ for the CD compensator. In the case of relatively long transmission links and hence large GVD, the GVD is estimated once the receiver is turned on, and the compensated GVD value remains unchanged during normal operation. A CD compensator with fixed $H^{-1}(\omega)$ is often referred to as a fixed equalizer.

Small amounts of residual GVD, not compensated by the fixed equalizer, are often removed at a later stage in the digital signal processing, where it is possible to estimate the GVD more accurately. Residual GVD is usually compensated in the time domain and in combination with adaptive compensation of polarization-mode dispersion (PMD), as described in more detail in 6.4.3 [49], [51]. Such combination of fixed and adjustable CD compensation allows adaptive compensation of small variations of the GVD in the transmission fibre [51].

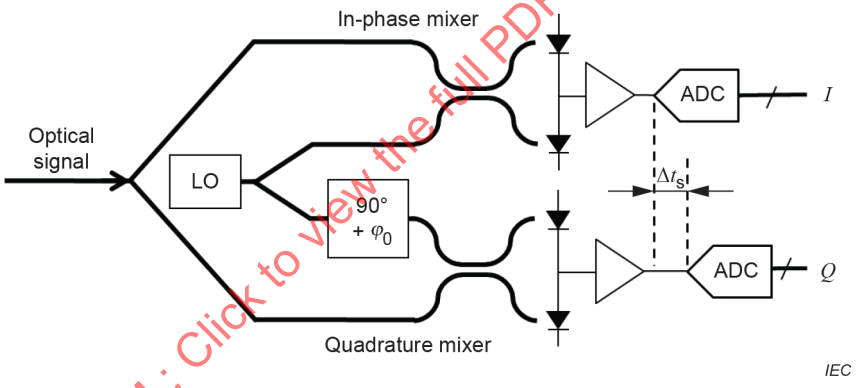
A CD estimator could be located either before or after the fixed CD compensation stage. In the latter case, it is possible to perform a blind search for the GVD value in $H^{-1}(\omega)$ that yields the lowest residual GVD after the CD compensator [54], [55]. For such a blind search, it is not necessary to measure the absolute value of the residual GVD after the compensator, but rather

use a metric for the relative amount of residual GVD as a cost function while the GVD in $H^{-1}(\omega)$ is scanned over a wide range. Such a cost function can be obtained, for example, by calculating spectral correlations in the Fourier transformed signal, or alternatively in the time domain, by characterizing the CD-induced distortion via a constant modulus algorithm [56], [55]. It is also possible to determine residual GVD from the distortion in a special sequence of known symbols (e.g. training symbols) that are intermittently transmitted [57].

NOTE Coherent communication systems often transmit over fibre links that do not employ in-line optical dispersion compensators, so that the transmitted signal experiences substantial amounts of accumulated GVD in the transmission fibre. This mode of operation often minimises signal distortions due to nonlinear effects in the transmission fibre, because the presence of large uncompensated CD limits the undesired accumulation of the nonlinear signal distortions caused by cross-phase modulation and four-wave mixing in the fibre [58]. For this reason, the amount of GVD to be compensated in the coherent receiver can be fairly large, especially in long-distance communication links [6].

6.3.4 Compensation of I-Q skew and phase offset

CD compensation and frequency offset removal both require the electrical phase between the I and Q signals to be very close to 90° at all frequencies within the signal bandwidth. In practice, the I-Q phase often exhibits frequency-dependent as well as frequency-independent offsets from the desired value of 90°. When the optical phase shift in the quadrature mixer deviates from the desired 90° by a phase error of φ_o, as shown schematically in Figure 18, the I-Q phase exhibits a frequency-independent offset. Moreover, when there is a differential time delay (or skew) of Δt_s between the two output signals, as also shown in Figure 18, the I-Q phase exhibits a frequency-dependent phase error.



Key

$90°+\varphi_o$	Optical phase shifter with phase error φ_o
Δt_s	Differential time delay (skew) between I and Q signals
ADC	Analogue-to-digital converter
LO	Local oscillator laser

Figure 18 – Dual coherent mixer with optical phase offset and signal skew

With such phase error in the quadrature mixer and differential time delay between the I and Q signals, the two receiver output signals have the form of Formula (28).

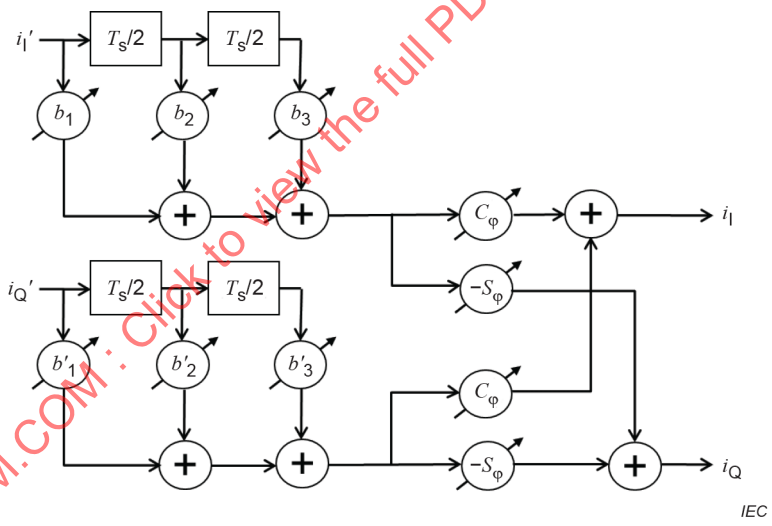
$$\begin{aligned}\tilde{i}_{I'}(t) &= R\sqrt{P_{LO}P_S}\cos(\theta)\,a_s(t)\;\sin\left[\Delta\omega t-\varphi_s(t)+\hat{\varphi}_{tot}(t)\right]+\hat{i}_{tot}(t)\\ \tilde{i}_{Q'}(t) &= R\sqrt{P_{LO}P_S}\cos(\theta)\,a_s(t+\Delta t_s)\;\cos\left[\Delta\omega(t+\Delta t_s)+\varphi_o-\varphi_s(t+\Delta t_s)+\hat{\varphi}_{tot}(t)\right]+\hat{i}_{tot}(t)\end{aligned} \quad (28)$$

For simplicity, it is assumed in Formula (28) that φ_r = 0. The phase errors in Formula (28) can be corrected via digital signal processing, which should be done prior to frequency offset removal and CD compensation [60]. The two signals can be transformed back into the form of Formula (18) by first removing the time delay Δt_s from $\tilde{i}_{Q'}(t)$ and subsequently performing the matrix multiplication of Formula (29), as shown schematically in Figure 19 [41].

$$\begin{bmatrix} \tilde{i}_I(t) \\ \tilde{i}_Q(t) \end{bmatrix} = \begin{bmatrix} 1, & 0 \\ -\tan(\varphi_o), & \cos^{-1}(\varphi_o) \end{bmatrix} \times \begin{bmatrix} \tilde{i}_{I'}(t) \\ \tilde{i}_{Q'}(t - \Delta t_s) \end{bmatrix} \tag{29}$$

The differential time delay between the I and Q signals can be removed by either delaying the I signal by Δt_s to obtain $\tilde{i}_{I'}(t + \Delta t_s)$ or, alternatively, by delaying the I signal and advancing the Q signal simultaneously by $\Delta t_s/2$, so as to obtain $\tilde{i}_{I'}(t + \Delta t_s/2)$ and $\tilde{i}_{Q'}(t - \Delta t_s/2)$, respectively.

In practice, this is accomplished with the help of fractionally spaced equalizers (FSEs; see 6.3.3) with real-valued tap coefficients [48]. Differential time delays smaller than the sampling period of the signal can be removed with relatively short FSEs, having as few as three taps, whereas time delays longer than the sampling period can be removed by first shifting the sampled signals by an appropriate number of samples and then removing any residual delay with an FSE.



Key

T_s	Symbol period
$b_1 \ldots b_3$	Real-valued tap coefficients
C_φ	$\cos(\varphi_o/2)$
S_φ	$\sin(\varphi_o/2)$

Figure 19 – Skew and phase offset removal in a DSP

Alternatively, the differential time delay can be removed in the frequency domain, after Fourier transformation of the signals in Formula (28), by multiplying the spectral components with the complex phasors exp(j $\Delta\omega_s$ $\Delta t_s/2$) and exp(−j $\Delta\omega_s$ $\Delta t_s/2$), respectively. In either case, the phase offset φ_o is subsequently removed via a matrix multiplication of the two signals, as described above, where often a symmetric matrix as shown in Formula (30) is used instead of Formula (29).

$$\begin{bmatrix} \cos(\varphi_o/2), & -\sin(\varphi_o/2) \\ -\sin(\varphi_o/2), & \cos(\varphi_o/2) \end{bmatrix} \tag{30}$$

It is usually not necessary to adaptively adjust the compensation values of Δt_s and φ_o during normal operation, because the skew and phase offset normally do not vary with time. Thus, they can be measured and corrected in an initial calibration of the coherent receiver. This can be accomplished, for example, by mixing the reference signal with an unmodulated signal to generate two sinusoidal beat signals at angular frequency $\Delta\omega$, from which one can determine the relative phase between the I and Q signals. When these beat signals are recorded inside the DSP, they include all differential time delays introduced in the optical and electrical components of the receiver.

As an example, the inset of Figure 20 displays a Lissajous trace of the quadrature signal i_Q versus the in-phase signal i_I, whose slightly elliptical shape reveals a phase error of 3,7° between the two signals.

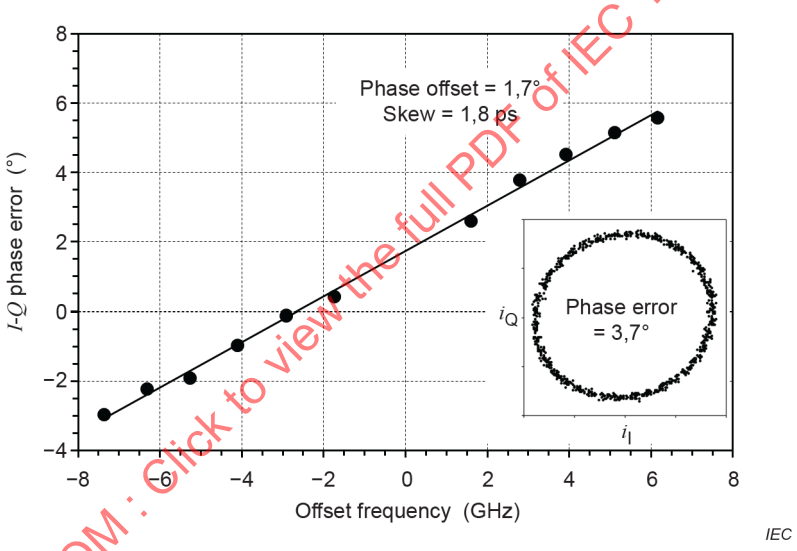


Figure 20 – Example of an I-Q skew and phase error measurement

A set of measurements at different frequency offsets $\Delta\omega$ then yields the phase error as a function of frequency. An example of such measurement is shown in Figure 20, where the I-Q phase errors measured at different frequency offsets form a straight line having a finite slope. The skew Δt_s can then be readily calculated from the slope of the line, and the phase offset φ_o is given directly by its value at $\Delta\omega = 0$.

6.3.5 Spectral shaping and frequency equalization

Digital signal processing can also be employed to reshape the spectrum of the received signal, in order to compensate for undesired frequency dependence in the PDs, TIAs, and ADCs or to remove undesired spectral components from the signals [1]. Such frequency equalization or spectral shaping can be accomplished, for example, with the help of fractionally spaced equalizers (FSEs) having N concatenated delay elements with real-valued tap coefficients b_m, m = 1, …, N + 1, similar to the ones described in 6.3.4 for skew and phase offset removal. The spectral response of an FSE, denoted $T(\omega)$, can be described by the truncated Fourier series shown in Formula (31) [47].

$$T(\omega) = \left[\sum_{n=-N/2}^{N/2} b_{n+N+1} \, e^{j n \omega \Delta t} \right] e^{j (N/2+1) \omega \Delta t} \tag{31}$$

The function $T(\omega)$ in Formula (31) is periodic with a period of $\omega_p = 2\pi/\Delta t$, such that $T(\pi/\Delta t + \omega) = T(-\pi/\Delta t + \omega)$ and $T(\pi/\Delta t + \omega_p) = T(\pi/\Delta t)$. This periodicity does not affect signals that are digitally sampled at a rate of $1/\Delta t$, because these signals do not carry information beyond the Nyquist frequency of the sampled signals, which is $\omega = \pi/\Delta t$. Formula (32) can be used to calculate the tap coefficients b_m that correspond to a given filter shape $F(\omega)$, analogous to Formula (27) for complex tap coefficients [47].

$$b_m = \mathrm{Re}\left[\frac{\Delta t}{2\pi} \int_{-\pi/\Delta t}^{\pi/\Delta t} F(\omega) \, e^{-j (m - N/2 - 1) \omega \Delta t} d\omega \right] \tag{32}$$

The integral in Formula (32) is real-valued if $F(\omega)$ is antisymmetric about $\omega = 0$, with $|F(\omega)| = |F(-\omega)|$ and $\arg\{F(\omega)\} = -\arg\{F(\omega)\}$.

The frequency resolution of an FSE is determined by the number of taps and the total time delay in its elements, i.e. by $\Delta T = N\Delta t$ [47]. Thus, the greater N, the steeper the slope in the spectral response can be. For example, the amplitude and phase response of a gently-sloped fifth-order Bessel filter with a 3 dB bandwidth of one half the symbol rate (i.e. $0{,}5/T_s$) can be approximated with a 5-tap FSE having only four $T_s/2$ delay elements, as shown in Figure 21, which also displays the corresponding five tap coefficients b_n of this filter.

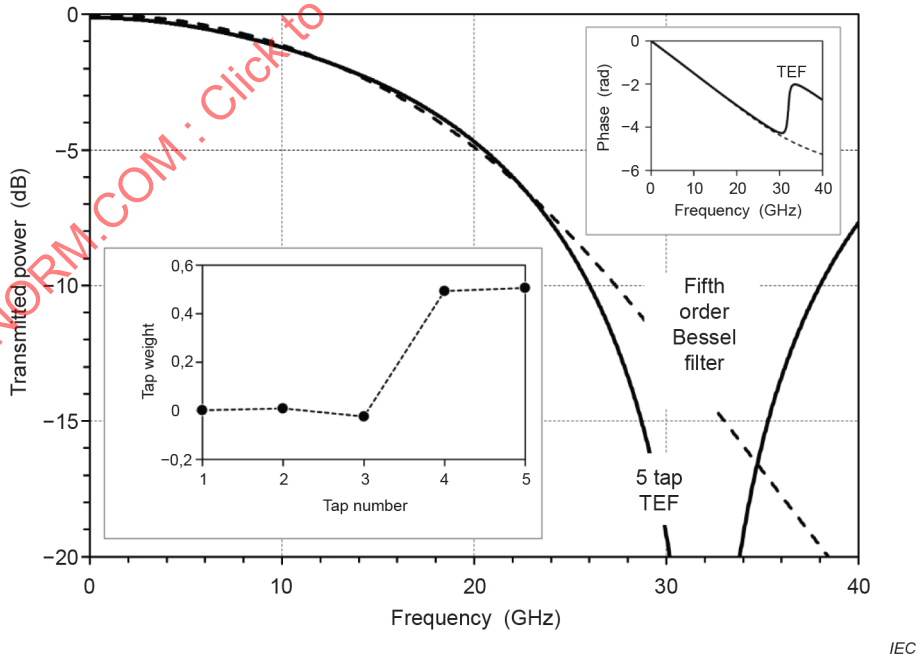


Figure 21 – Fifth-order Bessel filter emulated with a 5-tap FSE

As can be seen from Figure 21, the phase and amplitude of $T(\omega)$ deviate significantly from the desired shape when ω approaches $\pi/\Delta t$ (which corresponds to a frequency of 32 GHz in Figure 21). This discrepancy stems from the aforementioned periodicity of the FSE, i.e. from $T(\pi/\Delta t + \omega) = T(-\pi/\Delta t + \omega)$, which is impossible to realize in the case of Figure 21, because the phase of $F(\omega)$ at $\omega = \pi/\Delta t$ is different from that at $\omega = -\pi/\Delta t$.

FSEs can also be employed for compensation of undesired frequency roll-off caused by non-ideal components in the coherent receiver or by spectral filtering in the communication link (e.g. in ROADMs) [1], [31]. Frequency equalization often requires a filter shape that is similar to that of a high-pass filter, so that the transmission increases with increasing ω and is often highest at $\omega = \pi/\Delta t$. Generation of such filter shapes usually requires FSEs with significantly higher frequency resolution than needed for low-pass filters of the inverted spectral shape. This requirement is a direct consequence of the periodic FSE response, introducing sharp turns in $T(\omega)$ at $\omega = \pi/\Delta t$. If the frequency resolution of the FSE is not high enough to follow these sharp turns, the filter response obtained from Formula (32) often exhibits substantial "ringing" in $T(\omega)$ at lower frequencies.

This periodic meandering about the desired filter shape can be seen in Figure 22, which shows the frequency response of a 9-tap FSE (grey curve) that is employed to equalise the frequency roll-off caused by the 16-GHz Bessel filter shown in Figure 21.

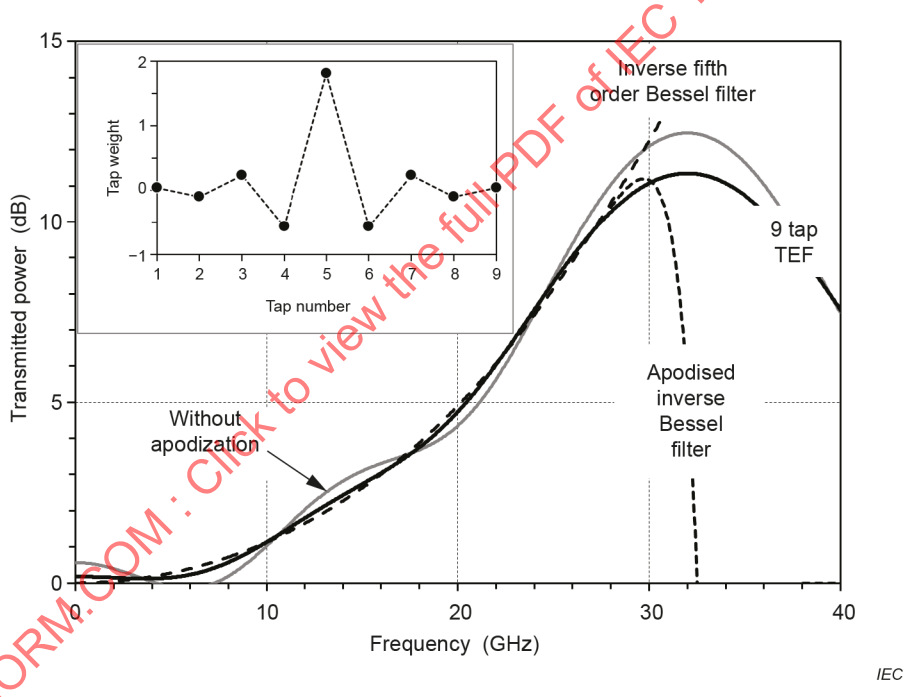


Figure 22 – Shape of inverse Bessel filter generated with a 9-tap FSE

The ringing in the frequency response can be reduced by reshaping the desired equalization function $F(\omega)$ when calculating the tap coefficients b_n, so as to reduce the sharpness of the turns at $\omega = \pi/\Delta t$, which is accomplished by replacing $F(\omega)$ in Formula (32) with an apodised frequency response $W(\omega)\ F(\omega)$, where $W(\omega)$ is a suitable spectral window function with flat frequency response $W(\omega) = 1$ up to some frequency $|\omega_\alpha| < \pi/\Delta t$ and a gentle roll-off towards $W(\omega) \approx 0$ at $|\omega| = \pi/\Delta t$.

Figure 22 displays the effects of such apodization in the case of the 9-tap FSE equalizer used to equalise a 16-GHz Bessel filter. The grey curve shows the FSE response without apodization, where the frequency resolution of the 9-tap FSE is not high enough to follow the sharp drop-off at $\omega = \pi/\Delta t$ and, as a result, introduces undesired ringing in the filter transmission at lower frequencies. This ringing is substantially reduced when apodising the filter shape in such a way that $F(\omega)W(\omega) \approx 0$ at $\omega = \pi/\Delta t$ (corresponding to 32 GHz), as represented by the black curves

in Figure 22. Although the apodization introduces a substantial deviation of the filter response from the desired shape at frequencies above 28 GHz, this drawback is outweighed by better equalization at lower frequencies.

It should be noted that the 9-tap FSE in Figure 22 generates only the amplitude response of the inverse Bessel filter, but not its phase response, which approximately equals a time delay of about 24 ps (at frequencies below 20 GHz). To generate such a large phase delay in addition to the inverse Bessel filter shape requires an FSE with a much larger number of taps as well as stronger apodization. This is shown in Figure 23, where a 19-tap FSE is used to generate amplitude and phase response of the inverse Bessel filter. In a DSP with digitally sampled signals, the FSE does not need to generate the 24 ps time delay, but rather only 8,4 ps, because shifting the signals by one sample period could introduce the remaining 15,6 ps of delay. Such filter function could be generated with the FSE comprising just 17 taps.

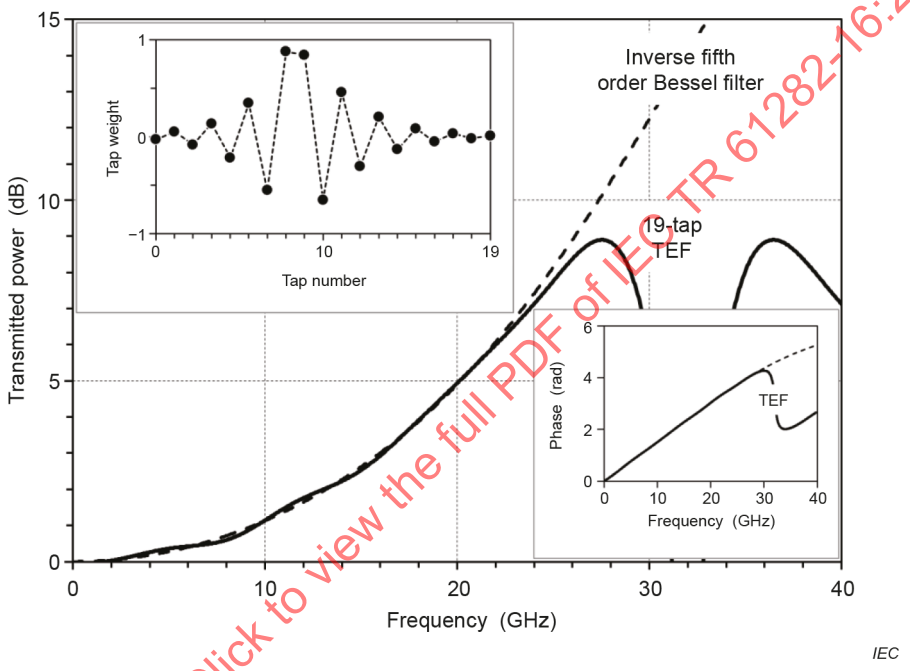


Figure 23 – Amplitude and phase of FSE-generated inverse Bessel filter

Fairly long FSEs are also required to generate so-called “Nyquist WDM” signals, where the transmitted and received optical signals are filtered to exhibit a nearly flat spectrum in the centre and a fairly sharp roll-off at the edges [61]. Such tight signal filtering is difficult to realize with optical filters but can be readily accomplished in the electrical domain using an FSE, as shown in the example of Figure 24 [1], [12]. Because of the sharp filter edges, the FSE needs to have fairly high frequency resolution and, hence, a large number of taps and delay elements. In the example of Figure 24, a 33-tap FSE was used to generate a root-raised-cosine (RRC) shaped filter with a roll-off factor of 0,1. The small ripple in the amplitude response of the FSE indicates that the frequency resolution of the FSE is barely sufficient to generate this almost rectangular filter shape.

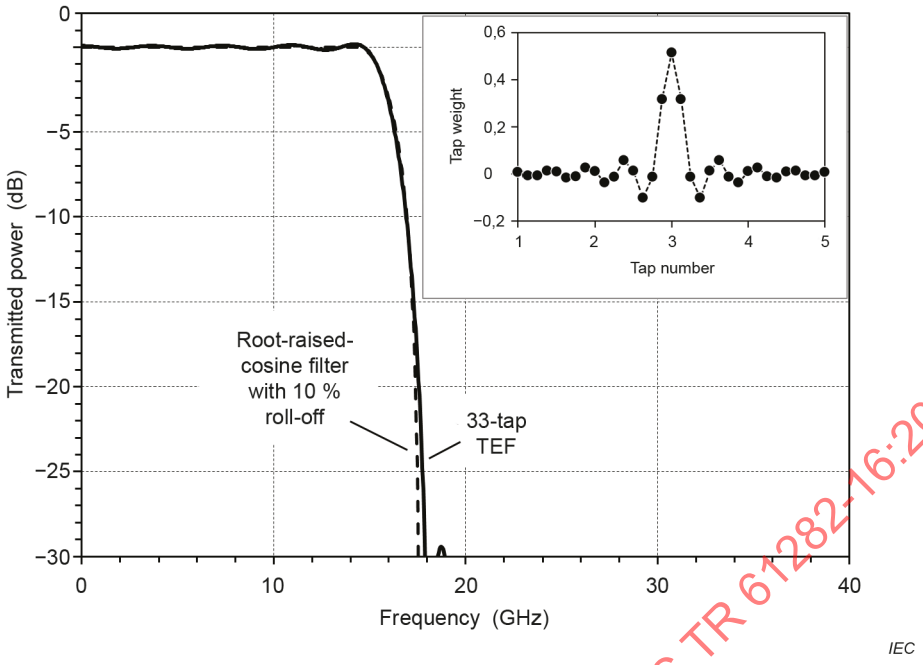


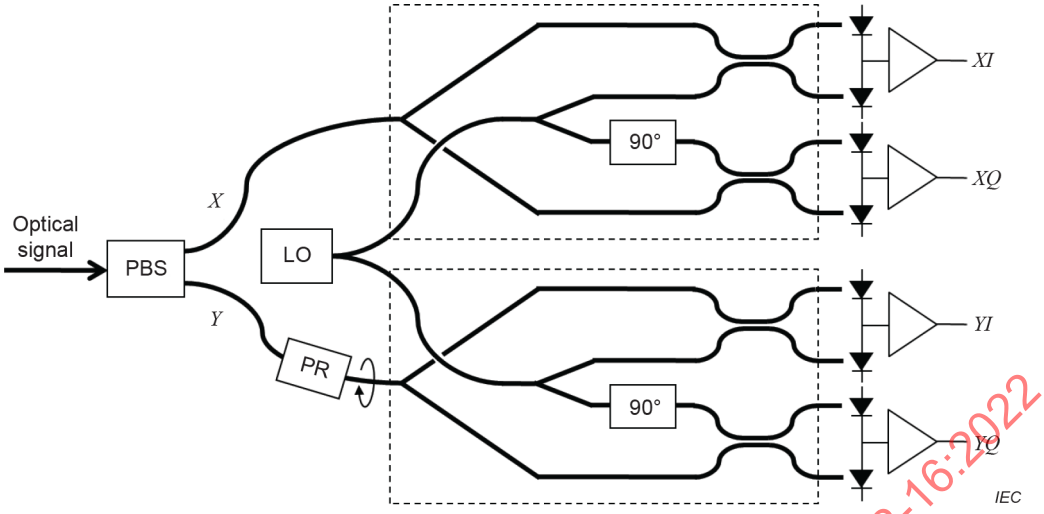
Figure 24 – Root-raised-cosine filter emulated with a 33-tap FSE

6.4 Quadruple mixer with phase and polarization diversity

6.4.1 Principle of operation

Coherent mixers are inherently polarization dependent and, therefore, can detect only single-polarized signals [28]. This drawback can be mended by adding another hybrid mixer to detect the signal component that is orthogonally polarized to the one analysed by the first hybrid mixer, as shown in Figure 25, thus providing phase and polarization diversity [29]. This arrangement not only allows reception of arbitrarily polarized signals but also reception and decoding of polarization-multiplexed signals [1], [9] .

In Figure 25, a polarization beam splitter (PBS) is used to split the received optical signal into two orthogonally polarized components, which are then fed into two identical hybrid mixers of the same design as shown in Figure 10. A 90° polarization rotator (PR) is usually inserted right before the signal input to the second mixer, so as to match the signal's polarization state with that of the LO signal [16]. In this case, the polarization states of the received signals and the reference signals are identical in both hybrid mixers, so that they do not have to be polarization independent. The reference signal from the local oscillator is split between the two hybrid mixers using a simple 3 dB coupler.



Key

90° Optical phase shifter

LO Local oscillator laser

PBS Polarization beam splitter

PR Polarization rotator

Figure 25 – Quadruple coherent mixer with phase and polarization diversity

Alternatively, polarization matching can be achieved by rotating the linear polarization state of the reference signal by 45° and using a polarization splitter to divide the reference signal between the two hybrid mixers. However, this arrangement requires mixers that operate independently of the state of polarization of the signals, because the reference and received signals are polarized orthogonally in the two mixers.

The receiver of Figure 25 generates four different coherent beat signals, which are labelled XI, XQ, YI, and YQ. In the case of a single-polarization signal arriving at the mixer in an arbitrary polarization state, as described by Formula (9), the four beat signals are given by Formula (33), where – for simplicity – all four mixers are assumed to have equal responsivities [1].

$$\begin{aligned} i_{\mathrm{XI}}(t) &= R\,\sqrt{P_{\mathrm{LO}}\,P_{\mathrm{S}}}\,\cos(\theta)\,a_{\mathrm{s}}(t)\,\sin\left[\Phi_{\mathrm{r}}(t)-\Phi_{\mathrm{s}}(t)\right]+\hat{i}_{\mathrm{tot}}(t) \\ i_{\mathrm{XQ}}(t) &= R\,\sqrt{P_{\mathrm{LO}}\,P_{\mathrm{S}}}\,\cos(\theta)\,a_{\mathrm{s}}(t)\,\cos\left[\Phi_{\mathrm{r}}(t)-\Phi_{\mathrm{s}}(t)\right]+\hat{i}_{\mathrm{tot}}(t) \\ i_{\mathrm{YI}}(t) &= R\,\sqrt{P_{\mathrm{LO}}\,P_{\mathrm{S}}}\,\cos(\theta)\,a_{\mathrm{s}}(t)\,\sin\left[\Phi_{\mathrm{r}}(t)-\Phi_{\mathrm{s}}(t)-\psi\right]+\hat{i}_{\mathrm{tot}}(t) \\ i_{\mathrm{YQ}}(t) &= R\,\sqrt{P_{\mathrm{LO}}\,P_{\mathrm{S}}}\,\cos(\theta)\,a_{\mathrm{s}}(t)\,\cos\left[\Phi_{\mathrm{r}}(t)-\Phi_{\mathrm{s}}(t)-\psi\right]+\hat{i}_{\mathrm{tot}}(t) \end{aligned} \tag{33}$$

Note that these four signals contain information on all optical properties of the received optical signal, which are

– the signal amplitude $a_{\mathrm{s}}(t)$,

– the optical phase $\Phi_{\mathrm{s}}(t)$ of the signal, and

– the polarization state of the signal, characterized by the phase angles θ and ψ.

In practical implementations, the reference signals in the two hybrid mixers can have different optical phases when combined with the received signal, so that the phase ψ in Formula (33) does not represent the polarization phase at the input of the coherent receiver. This difference can usually be ignored for the reception of communication signals, unless one is interested in

measuring the polarization state of the transmitted signal at the input of the receiver. In any case, the polarization-diversity mixer avoids the problem of polarization fading that can occur in a single mixer when the polarization state of the received signal fluctuates randomly. For example, in the case of θ = 90° one has $i_{\mathrm{XI}}(t) = i_{\mathrm{XQ}}(t) = 0$, but $i_{\mathrm{YI}}(t)$ and $i_{\mathrm{YQ}}(t)$ have maximal amplitude and carry all information of the transmitted signal. In fact, a DSP can transform the four signals of Formula (27) into just two, as discussed in 6.4.2, and even properly separate polarization-multiplexed signals. In addition, it is possible to remove undesired signal distortions caused by polarization-mode dispersion (PMD) and polarization-dependent loss (PDL) from the signals.

6.4.2 Polarization demultiplexing

The quadruple mixer of Figure 25 does not need an optical polarization controller to ensure continuous reception of the transmitted signal, since it is possible to reconstruct the polarization state of the transmitted signal from the four signals of Formula (27). This can be seen by letting $\Delta\omega = 0$ as well as $\varphi_{\mathrm{r}} = 90°$ and using a complex representation of the four beat signals, which yields the two complex beat signals shown in Formula (34) [1].

$$\begin{aligned} i_{\mathrm{X}}(t) &= i_{\mathrm{XI}}(t) + \mathrm{j}\, i_{\mathrm{XQ}}(t) \\ i_{\mathrm{Y}}(t) &= i_{\mathrm{YI}}(t) + \mathrm{j}\, i_{\mathrm{YQ}}(t) \end{aligned} \tag{34}$$

In the case of a single-polarized signal of the form of Formula (8), the originally transmitted signal can be reconstituted by performing the complex matrix multiplication shown in Formula (35), which describes a general polarization rotation.

$$\begin{bmatrix} \cos(\theta) & \sin(\theta)\, \mathrm{e}^{-\mathrm{j}\psi} \\ -\sin(\theta)\, \mathrm{e}^{\mathrm{j}\psi} & \cos(\theta) \end{bmatrix} \begin{bmatrix} i_{\mathrm{X}}(t) \\ i_{\mathrm{Y}}(t) \end{bmatrix} = \left\{ R \sqrt{P_{\mathrm{LO}}\, P_{\mathrm{S}}}\; a_{\mathrm{s}}(t)\, \mathrm{e}^{\mathrm{j}\left[\varphi_{\mathrm{s}}(t) + \hat{\varphi}_{\mathrm{tot}}(t)\right]} + 2\hat{i}_{\mathrm{tot}}(t) \right\} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{35}$$

Since the polarization state of the received signal is usually unknown and can even fluctuate with time, the parameters θ and ψ in Formula (35) are continuously adjusted to obtain the desired polarization transformation. Just like in the case of optical polarization controllers, a two-dimensional blind search algorithm can be employed to vary θ and ψ until the power in the desired output port is maximal. Hence, the quadruple mixer and the DSP act like an optical polarization controller (PC) with unlimited transformation range [62].

It should be noted that polarization-diversity receivers introduce additional noise when used with single-polarized signals [28]. This can be readily seen when the received signal is linearly polarized parallel with the X-polarization state of the receiver, so that only $i_{\mathrm{X}}(t)$ carries the desired beat signal, just like in the dual-mixer of Figure 10, and $i_{\mathrm{Y}}(t)$ carries only receiver noise [28], [29]. However, this disadvantage is eliminated when the polarization-diversity receiver is used to detect polarization-multiplexed signals, in which case $i_{\mathrm{X}}(t)$ and $i_{\mathrm{Y}}(t)$ always carry beat signals of nearly equal magnitude, independently of the polarization state of the received signal.

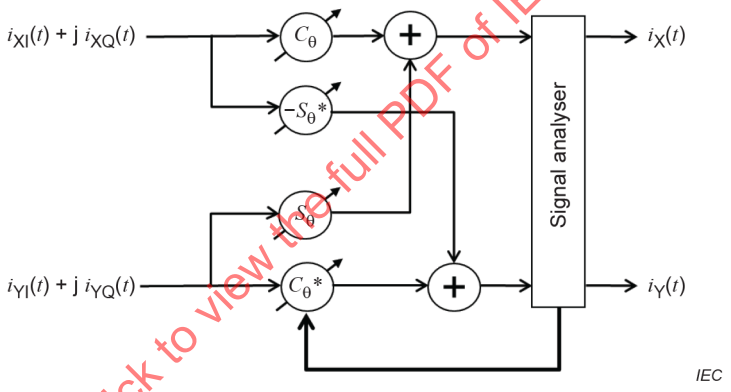
Unlike the single or dual mixers discussed above, the quadruple mixer of Figure 25 is capable of receiving polarization-multiplexed signals, which are composed of two independently modulated signals in orthogonal polarization states, as described by Formula (1). Polarization-multiplexed signals are frequently used to transmit signals at data rates of 100 Gbit/s and above [10], [11].

To decode polarization-multiplexed signals, their two orthogonally polarized signal components need to be properly demultiplexed in the DSP of the coherent receiver. This can be achieved

by using the same matrix multiplication as in Formula (35). However, in the case of polarization-multiplexed signals, the matrix multiplication (or polarization rotation) of Formula (35) produces two output signals, as shown in Formula (36), instead of just one as in Formula (35).

$$\begin{bmatrix} \cos(\theta) & \sin(\theta)\, e^{-j\psi} \\ -\sin(\theta)\, e^{j\psi} & \cos(\theta) \end{bmatrix} \begin{bmatrix} i_X(t) \\ i_Y(t) \end{bmatrix} = R\sqrt{P_{LO}\, P_S} \begin{bmatrix} A_{XC}(t)\, e^{j\hat{\varphi}_{tot}(t)} + \hat{i}_{tot}(t) \\ A_{YC}(t)\, e^{j\hat{\varphi}_{tot}(t)} + \hat{i}_{tot}(t) \end{bmatrix} \tag{36}$$

Since the polarization state of the received optical signal can change with time, because of polarization fluctuations in the fibre link, the parameters θ and ψ in Formula (35) and Formula (36) are usually adjusted adaptively, in such a way that the polarization rotation in Formula (36) always properly separates the two tributaries of the polarization-multiplexed signals [10]. However, adaptive adjustment of θ and ψ requires a different feedback signal for polarization-multiplexed signals than for single-polarization signals, because the two polarization tributaries cannot be distinguished by their optical power (which usually is substantially equal). Typically, the feedback signal for adaptive demultiplexing of polarization-multiplexed signals is derived from an analysis of the modulation content in the two demultiplexed signals, as shown schematically in Figure 26 [1].



Key

C_θ $\cos(\theta)$

C_θ^* Complex conjugate of C_θ

S_θ $\sin(\theta)\exp(-j\psi)$

S_θ^* Complex conjugate of S_θ

Figure 26 – 2 × 2 matrix operation for adaptive polarization demultiplexing

In general, the feedback signals used to demultiplex polarization-multiplexed signals tend to be different for different modulation formats of the transmitted signal. A constant modulus algorithm (CMA) is often used for polarization-multiplexed QPSK signals (PM-QPSK) [62], [64]. Polarization demultiplexing of higher-order QAM signals, such as PM-16QAM or PM-32QAM signals, often employ a radius-directed equalizer (RDE), sometimes also referred to as multi-modulus algorithm (MMA) [4], [60]. Alternatively, the feedback signal can be based on decision-directed least-mean-square algorithms (DD-LMS) [1], [4], [10], [65]. Polarization demultiplexing can also be facilitated by periodically transmitting training symbols to the receiver [60]. A comprehensive review of these techniques and algorithms is beyond the scope of this document. As an example, the principle of the widely used CMA method is described in the following paragraphs.

The CMA method utilizes the fact that the four constellation points of a properly polarization-demultiplexed QPSK signal are confined to a small area in the centre of each quadrant of the

I-Q constellation space, as shown in the rightmost diagram of Figure 27 [9]. However, the demultiplexed signals often carry a small residual frequency offset $\Delta\omega$, which continuously rotates the constellation points of the QPSK signal about their centre [63]. As a result, the four QPSK constellation points are scattered on a narrow circular band of radius r, as shown in the centre diagram of Figure 27. If the polarization states of the two signals are not properly demultiplexed, then the constellation points are spread across a much wider area in the constellation space, as shown in the left diagram of Figure 27 [63]. Hence, the width of the circular band over which the constellation points are scattered can serve as a feedback signal for θ and ψ in Formula (30) [60], [65].

NOTE Polarization multiplexing is usually performed after CD compensation so that potentially large CD-induced distortions are removed from the signal. This is indeed the case in Figure 27.

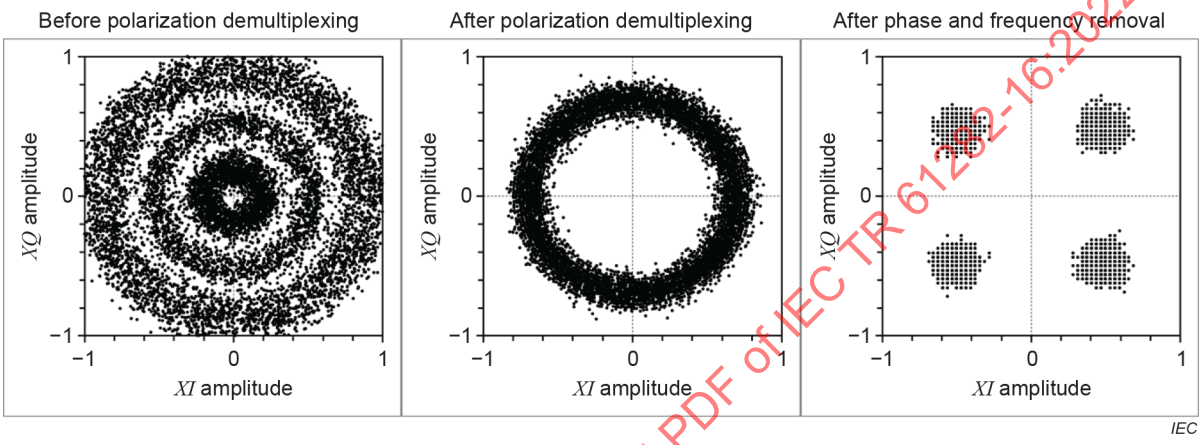


Figure 27 – Constellation points of QPSK signal after polarization demultiplexing

For higher-order QAM signals, like 16QAM, the constellation points are distributed among various circles of different radii, so that a modified feedback signal is needed for polarization control. For example, the constellation points of a polarization-demultiplexed “square” 16QAM signal are located on three different circles, as shown in Figure 28 [64]. In this case, a so-called radius-directed equalizer (RDE) algorithm can be employed, which successively adjusts θ and ψ until the various constellation points are confined to three separate circular bands [60], [65].

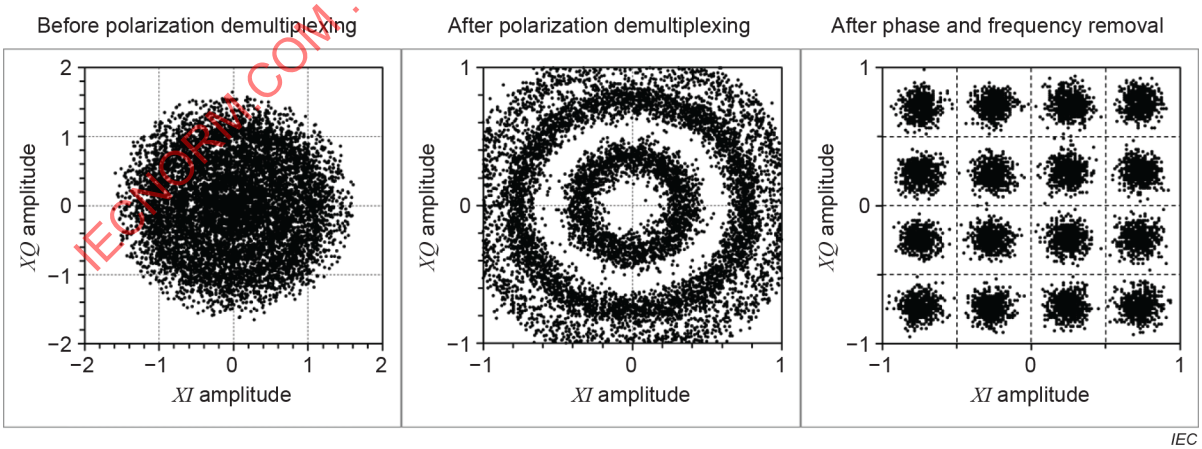


Figure 28 – 16QAM signal before and after polarization demultiplexing

The simple polarization rotation of Formula (30) does not correct for polarization-dependent loss (PDL), which the signal could have experienced in the communication link. In the case of significant PDL, where one (arbitrarily oriented) polarization component of the signal is more attenuated than the orthogonally polarized component, the constellation points of the transformed signal usually deviate from their circular confinement and, hence, impair the RDE-

based polarization control [64]. However, PDL and polarization-mode dispersion (PMD) can be compensated for by using a more complex transformation matrix, as discussed in 6.4.3 and 6.4.4.

6.4.3 Compensation of polarization-mode dispersion

It is well known that polarization-mode dispersion (PMD) in the communication link can lead to severe waveform distortions in the transmitted optical signals, especially at symbol rates of 10 GBd and above [5]. Compensation of PMD at the receiver is usually difficult, because the magnitude and orientation of PMD in the transmission fibre generally vary with time and optical frequency. Thus, a PMD compensator needs to be adaptively adjustable in order to compensate for these time-dependent variations. It is possible to build optical PMD compensators that remove at least some of the effects of PMD, but these are fairly complex devices. For example, an optical compensator for first-order PMD is typically composed of three elements:

– an (endlessly) adjustable optical polarization transformer,

– an optical delay element that generates continuously adjustable differential group delay (DGD) between two orthogonal polarization states, and

– a monitor detector that measures the residual DGD after the compensator and generates suitable error signals for the adjustable polarization transformer and the DGD element [66].

It is thus not surprising that optical PMD compensators have not been widely deployed in fibre optic communication systems. However, with the polarization-diversity receiver of Figure 25, it is possible to implement the functions of an optical PMD compensator in the DSP, so that the adverse effects of PMD can be compensated in the digitized electrical beat signals, similar to the compensation of CD discussed in 6.3.3. PMD compensation can be accomplished by replacing the simple polarization rotation in Formula (36) with a frequency-dependent polarization transformation, which can be described by the unitary matrix shown in Formula (37) [1], [60].

$$\begin{bmatrix} U(\Delta\omega_s)\, e^{j\Theta(\Delta\omega_s)} & V(\Delta\omega_s)\, e^{j\Psi(\Delta\omega_s)} \\ -V(\Delta\omega_s)\, e^{-j\Psi(\Delta\omega_s)} & U(\Delta\omega_s)\, e^{-j\Theta(\Delta\omega_s)} \end{bmatrix} \tag{37}$$

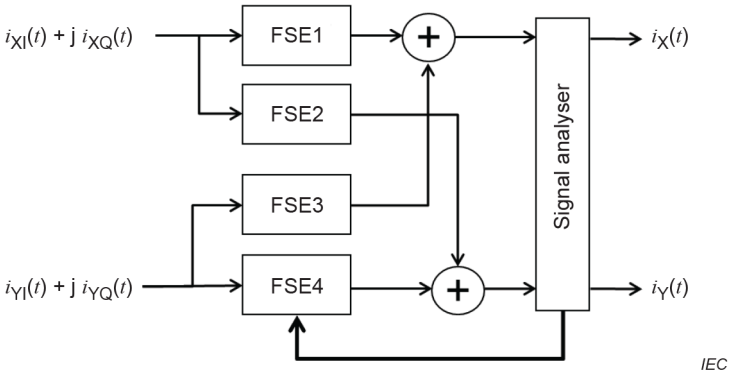
where

$\Delta\omega_s$ is the angular frequency of a spectral component relative to the carrier frequency ω_s,

$U(\Delta\omega_s)$ and $V(\Delta\omega_s)$ denote frequency-dependent attenuation with $U^2(\Delta\omega_s) + V^2(\Delta\omega_s) = 1$,

$\Theta(\Delta\omega_s)$ and $\Psi(\Delta\omega_s)$ are frequency-dependent time delays.

Although multiplication of the received signals with the matrix of Formula (37) can be performed in the frequency domain, it is often computationally more efficient to compensate PMD in the time domain, using relatively short complex-valued FSEs with adjustable tap weights, as shown schematically in Figure 29 [1]. In any case, polarization demultiplexing and PMD compensation can be accomplished simultaneously in a single matrix multiplication [10].



Key

FSE Fractionally spaced equalizer with complex-valued coefficients

Figure 29 – FSE-based compensator for polarization-mode dispersion

The arrangement shown in Figure 29 is a multiple-input, multiple-output (MIMO) dynamic equalizer, which is sometimes referred to as a butterfly equalizer [51]. It may be employed not only for compensation of PMD, as shown in Figure 29, but also for compensation of polarization-dependent loss (PDL) and residual CD. Each FSE introduces variable frequency-dependent attenuation as well as variable time delay. The total time delay in each FSE should be large enough to accommodate the largest DGD that is expected to occur in the communication link [51]. For example, in a 32 GBd QPSK signal that is sampled every 15 625 ps, one can compensate up to 120 ps of DGD using four 9-tap FSEs and adjusting a total of 36 complex tap coefficients. It should be noted that the transformation matrix (37) not only allows compensation of first-order PMD (i.e. DGD) but also of second- and higher-order PMD [67].

Since the PMD in a fibre optic communication link usually fluctuates with time, the various tap weights of the four FSEs need to be adaptively adjusted, so that the PMD is sufficiently compensated and the polarization tributaries are properly demultiplexed, using a blind search algorithm. The presence of PMD in the received signal (after polarization demultiplexing) causes radial spreading of the constellation points, as shown in Figure 30. This effect is similar to the radial spreading shown in Figure 27 and Figure 28, so that a CMA- or RDE-based feedback control loop can be employed to automatically set the various tap weights of the four FSEs [1]. Once the tap coefficients are properly adjusted, the amount of compensated DGD and second-order PMD can be determined from the settings of the coefficients, where the second-order PMD is usually specified in units of ps^2 [67].

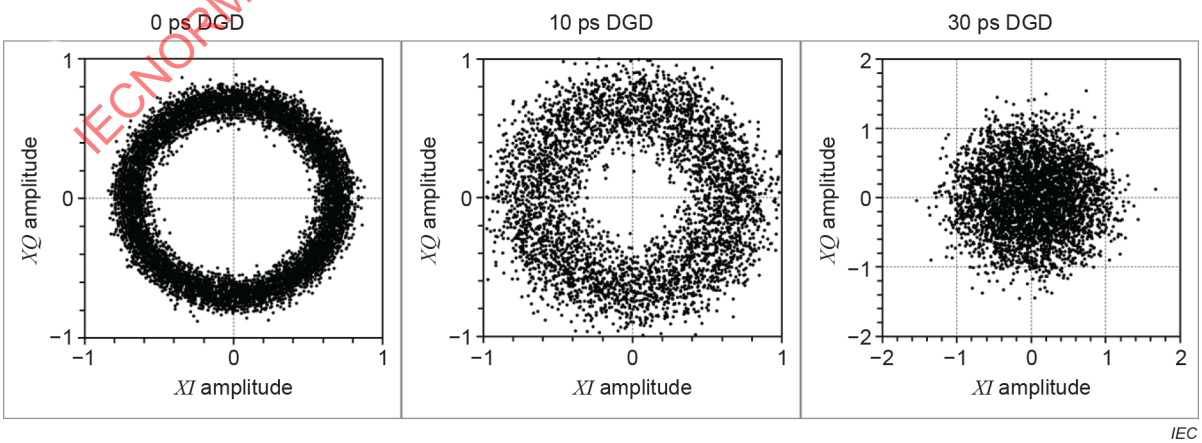


Figure 30 – QPSK signal constellations for various amounts of PMD

However, the control loop has to be fast enough to accommodate rapid polarization rotations, which can occur especially in long fibre optic transmission links. The speed of these rotations can be high, up to several 10 krad/s [10].

6.4.4 Compensation of polarization-dependent loss and residual CD

The matrix compensator of Figure 29 may also be used to simultaneously compensate for polarization-dependent loss (PDL) and residual chromatic dispersion [67], [68]. This can be accomplished by generalizing the matrix in Formula (37) to include the effects of arbitrarily oriented PDL as well as residual GVD, so that it becomes Formula (38) .

$$\begin{bmatrix} U_x(\Delta\omega_s)\, e^{j\Theta_x(\Delta\omega_s)} & V_x(\Delta\omega_s)\, e^{j\Psi_x(\Delta\omega_s)} \\ V_y(\Delta\omega_s)\, e^{j\Psi_y(\Delta\omega_s)} & U_y(\Delta\omega_s)\, e^{j\Theta_y(\Delta\omega_s)} \end{bmatrix} \tag{38}$$

However, compensation of relatively large amounts of residual GVD requires FSEs with relatively long total time delays and, hence, a large number of taps.

As can be seen from Figure 31 and Figure 32, PDL and residual GVD also cause radial spreading of the constellation points, so that again a CMA- or RDE-based feedback control loop can be employed to automatically adjust the eight independent parameters in Formula (38).

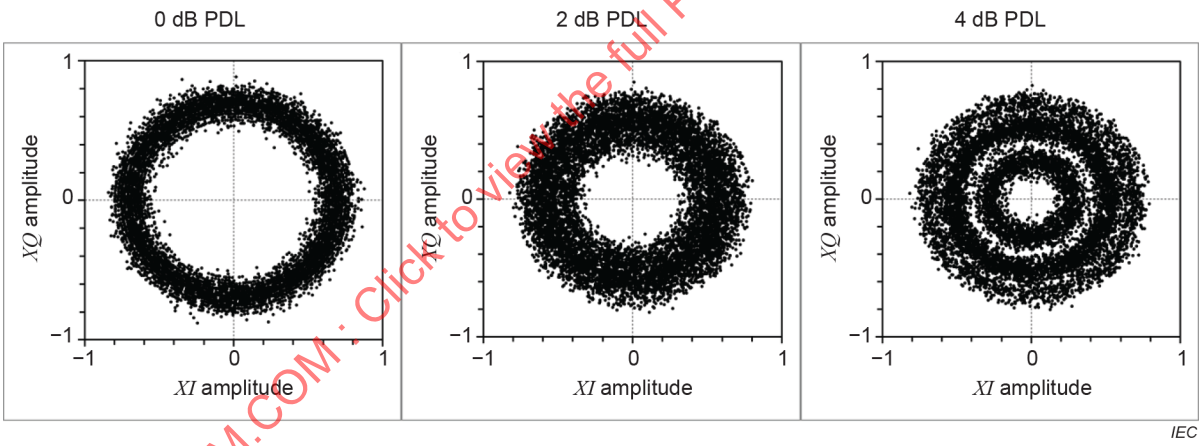


Figure 31 – QPSK signal constellations for various amounts of PDL

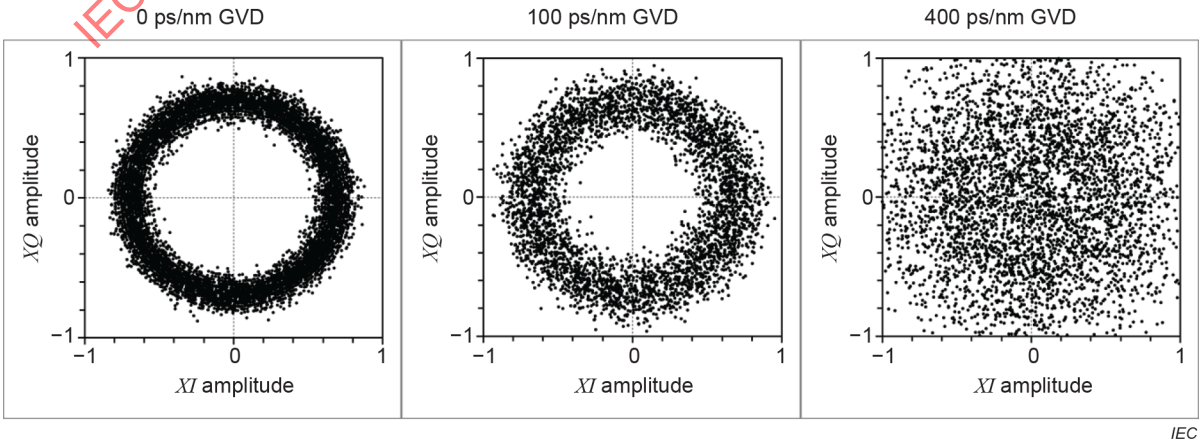


Figure 32 – QPSK signal constellations for various amounts of GVD

However, when using a generalized transformation matrix of the form of Formula (29), one needs to carefully monitor the tap coefficients in the FSEs to make sure that the two polarization-multiplexed signals are properly demultiplexed, because it is possible that Formula (32) generates identical output signals in the two output ports of the demultiplexer, e.g. $i_x(t) \equiv i_y(t)$. This is known as the singularity problem [60].

6.4.5 Carrier phase recovery

Before the transmitted information can be extracted from the received signal, it is necessary to remove undesired phase offsets from the polarization-demultiplexed signals. This processing step is known as carrier phase recovery and requires very fast determination and subtraction of the time-varying phase offsets, because the phase offsets comprise not only slow phase variations $\Delta\varphi_r = \Delta\omega t$ due to residual frequency offsets between the reference and transmitted signals but also rapid optical phase fluctuations $\hat{\varphi}_{tot}(t)$ in the transmitter and LO lasers as well as rapid phase variations due to nonlinear signal interactions in the transmission fibre.

Because of these speed requirements, a feed-forward algorithm is often used for carrier phase recovery rather than a feedback loop. In this case, the phase offset in the signals is determined from an analysis of the received signals before it is subtracted from the signals. Feedforward algorithms include the Viterbi and Viterbi algorithm, Wiener filters, and maximum a posteriori (MAP) estimators [49], [70], [71]. In addition, there are various soft or hard decision-directed phase estimation methods, which of course strongly depend on the modulation format of transmitted signal [70], [71], [72]. A comprehensive review of these methods is beyond the scope of this document. As an example, the following paragraphs describe the widely used Viterbi and Viterbi algorithm, which is also known as the M-th power algorithm [55].

In QPSK signals, the phase offset can be determined with the help of the Viterbi and Viterbi algorithm, which utilises the fact that the fourth power of all four QPSK symbols always equals 1 when there are no phase offsets in the signal, and $\exp(j4\Delta\varphi)$ when a phase offset $\Delta\varphi$ is present. The four QPSK symbols without phase offsets can be described by the complex phasors shown in Formula (39) [49].

$$\begin{aligned} z_1 &= e^{j0} = 1 \\ z_2 &= e^{j\pi/2} = j \\ z_3 &= e^{j\pi} = -1 \\ z_4 &= e^{-j\pi/2} = -j \end{aligned} \tag{39}$$

where

z_n is a complex number in the I-Q space describing the four QPSK symbols n = 1, …, 4.

It is easily verified that Formula (39) yields ${z_n}^4 = 1$ for n = 1, …, 4. However, when an undesired phase offset $\Delta\varphi$ is present in the beat signals, Formula (39) becomes Formula (40).

$$\begin{aligned} z_1 &= e^{j\Delta\varphi} \\ z_2 &= j\,e^{j\Delta\varphi} \\ z_3 &= -e^{j\Delta\varphi} \\ z_4 &= -j\,e^{j\Delta\varphi} \end{aligned} \tag{40}$$

It is easily verified that $z_n^4 = \exp(j4\Delta\varphi)$ for all four symbols in Formula (40), so that phase offsets of up to ±π/4 can be calculated from the argument of the fourth root of z_n^4 and subsequently removed from the received signals [1], [55]. It is important to note that the phase offset can only be determined up to modulo π/2. Thus, if the modulus of the phase offset, $|\Delta\varphi|$, is larger than π/4, then the received symbols cannot be unambiguously mapped to the transmitted symbols. The ambiguity can be resolved by periodically sending a short sequence of known symbols to the receiver, with which it is possible to determine phase offsets up to modulo 2π. These symbols can be frame overhead symbols, special training symbols and/or periodically inserted pilot tones [11].

The carrier phase recovery algorithm thus needs to track and remove small phase offsets that can occur between each set of training symbols. However, the phase offset $\Delta\varphi$ is usually averaged over several consecutive symbol periods before it is subtracted from the received signal, because the calculation of $\Delta\varphi$ is affected by noise in the signal. Furthermore, the average phase offset $<\Delta\varphi>$ is usually determined separately for the X- and Y-polarized signals, as described by Formula (41) [55].

$$\left\langle \Delta\varphi_{\mathrm{x}}(t) \right\rangle = \frac{1}{4(2M+1)} \sum_{m=-N}^{N} \arg\left[i_{\mathrm{x}}{}^4 \left(t + mT_{\mathrm{s}} \right) \right]$$
$$\left\langle \Delta\varphi_{\mathrm{y}}(t) \right\rangle = \frac{1}{4(2M+1)} \sum_{m=-N}^{N} \arg\left[i_{\mathrm{y}}{}^4 \left(t + mT_{\mathrm{s}} \right) \right] \quad (41)$$

The number of symbols over which the phase offset is averaged is typically of the order of 10 and depends on the amount of noise in the signal, in particular on phase noise in the transmitter and receiver lasers [10]. If the laser phase noise is large, then $|\Delta\varphi|$ can temporarily exceed π/4, which leads to an error in the determination of the average carrier phase and possibly to incorrect mapping of the transmitted symbols until the next set of training symbols is received. This error in the carrier phase is called a cycle slip because it causes a cyclic permutation of the decoded QPSK symbols [71].

The occurrence of cycle slips can be reduced by employing transmitter and LO lasers with low phase noise, but they cannot be completely avoided because of the random statistics of laser phase noise and nonlinear phase noise [71]. Thus, even though cycle slips occur relatively infrequently, they can cause a long string of symbol errors. The number of errors can be substantially reduced by encoding the signals differentially, which requires signals with higher SNR, or alternatively by employing sophisticated methods for detecting and correcting the phase errors [71]. Cycle slips can be detected, for example, by periodically transmitting a relatively short sequence of predetermined (and hence known) symbols, frequently referred to as training symbols or pilot signals. These training symbols can greatly improve the performance of carrier phase recovery and are also helpful in polarization demultiplexing and CD compensation.

The characterization of laser phase noise and its impact on coherent communication systems is discussed in more detail in 6.4.6.

6.4.6 Impact of laser phase noise

6.4.6.1 Characterization of laser phase noise

Aside from nonlinear phase noise, which is introduced by nonlinear effects in the communication link, the phase noise in the coherent beat signal is usually the result of the combined phase noise generated in the transmitter and LO lasers [36], [73]. As discussed in 6.4.5, proper recovery of the transmitted carrier phase requires transmitter and LO lasers with sufficiently low phase noise. However, when large amounts of GVD have to be compensated in the receiver, there are further restrictions on the allowable phase noise in the LO laser, as indicated in 6.3.3.

In the following, the impact of laser phase noise in coherent transmission systems is discussed in more detail.

Laser phase noise is often characterised by a single quantity, the instantaneous laser linewidth [73]. Although this parameter is a good measure of the laser phase noise occurring at relatively high frequencies (typically in the GHz range), it does not capture the laser phase noise occurring at lower frequencies (e.g. in the MHz range), which is particularly important for electronic compensation of large amounts of GVD [11], [69].

In some cases, one can deduce the low-frequency laser phase noise from the laser linewidth, but this assumes that the random laser frequency variations exhibit a flat spectrum in the frequency range of interest. It is important to note that this assumption usually is not valid for semiconductor lasers, which are predominantly used in coherent communication systems, because these lasers typically exhibit additional low-frequency laser phase noise that is introduced by electronic noise in the injection current and other frequency control elements of the laser. This additional noise is known as "1/f noise", because it adds more phase noise at lower frequencies than at higher frequencies [74]. Hence, the low-frequency phase noise in the transmitter and LO lasers can be significantly higher than what is estimated from the laser linewidth. For this reason, it is often necessary to specify the frequency dependence of the laser phase noise within a relatively large frequency range [11].

For semiconductor-based lasers, therefore, the magnitude of low-frequency phase variations, occurring over a period of a few nanoseconds or longer, is substantially larger than that of high-frequency noise, occurring over a period of a few picoseconds). This can be clearly seen from Figure 33, which displays the measured time evolution of the optical phase in the output of two unmodulated narrow-linewidth semiconductor lasers.

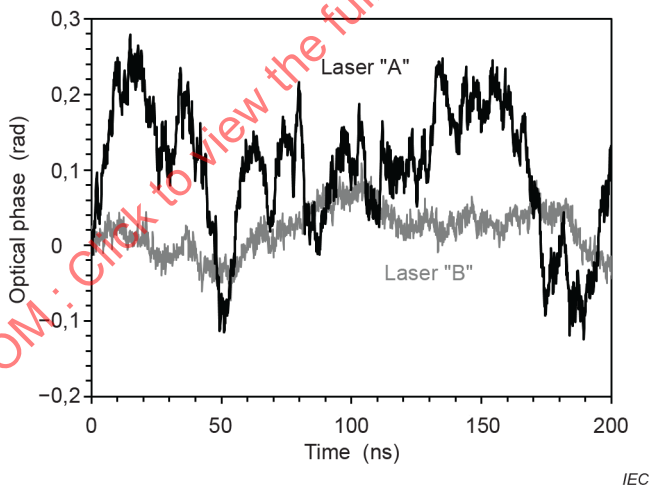


Figure 33 – Optical phase noise of two narrow-linewidth lasers

Low-frequency phase noise in LO lasers becomes especially important when relatively large amounts of CD are to be compensated in the receiver, as discussed briefly in 6.3.3 and in more detail in 6.4.6.2. This low-frequency phase noise can be characterized by a measurement of the laser phase noise spectrum or, alternatively, by a measurement of the laser frequency noise spectrum, as described in the following paragraphs of 6.4.6.1.

There are several different ways to measure the spectrum of the laser phase noise (or laser frequency noise). One technique involves coherent mixing of the (unmodulated) laser under test with a highly stable reference laser in a balanced mixer. The frequency of the reference laser is offset in frequency by a few GHz from that of the laser under test, so that mixing of the two signals produces the sinusoidal beat signal shown in Formula (42).

$$i_{\text{diff}}(t) = R\, a_{\text{r}}\, a_{\text{s}}\, \sin\left[2\pi\langle \Delta f_{\text{o}}\rangle t + \hat{\varphi}_{\text{s}}(t) + \hat{\varphi}_{\text{r}}(t)\right] + \hat{i}_{\text{tot}}(t) \tag{42}$$

where

a_{r} is the output amplitude of the unmodulated reference laser,

a_{s} is the output amplitude of the unmodulated laser under test,

$<\Delta f_{\text{o}}>$ is the average frequency difference between the two lasers,

$\hat{\varphi}_{\text{s}}(t)$ is the phase noise of the laser under test, and

$\hat{\varphi}_{\text{r}}(t)$ is the phase noise of the reference laser, assumed to be substantially smaller than $\hat{\varphi}_{\text{s}}(t)$.

The laser phase noise spectrum is obtained by first extracting the instantaneous beat frequency $\Delta f(t)$ from $i_{\text{diff}}(t)$, which is given by Formula (43).

$$\Delta f(t) = \langle \Delta f_{\text{o}}\rangle + \frac{1}{2\pi}\frac{d}{dt}\left[\hat{\varphi}_{\text{s}}(t) + \hat{\varphi}_{\text{r}}(t)\right] \approx \langle \Delta f_{\text{o}}\rangle + \frac{1}{2\pi}\frac{d}{dt}\hat{\varphi}_{\text{s}}(t) \tag{43}$$

The instantaneous frequency can be determined, for example, by measuring the time period Δt_0 between succeeding zero-crossings in the coherent beat signal, yielding $\Delta f(t) = 2/\Delta t_0(t)$. After subtraction of the frequency offset Δf_{o}, one obtains the instantaneous frequency noise $\Delta \hat{f}_{\text{s}}(t) = (2\pi)^{-1}\, d\hat{\varphi}_{\text{s}}(t)/dt$ of the laser under test, from which the phase noise $\hat{\varphi}_{\text{s}}(t)$ can be calculated by integration, as shown in the example of Figure 33. From these recordings, one can now calculate the frequency noise spectrum and/or the phase noise spectrum of the laser. Both spectra are single-sided power spectra, with spectral densities in units of Hz^2/Hz and rad^2/Hz, respectively. Fourier transformation of $\Delta \hat{f}_{\text{s}}(t)$ and $\hat{\varphi}_{\text{s}}(t)$ yields double-sided amplitude spectra with complex-valued densities $\hat{F}_{\text{ds}}(f)$ and $\hat{\Phi}_{\text{ds}}(f)$, respectively, from which the power spectral densities are calculated by folding of the squared densities $\left|\hat{F}_{\text{ds}}\right|^2(f)$ and $\left|\hat{\Phi}_{\text{ds}}\right|^2(f)$ into the single-sided spectra with densities described by Formula (44), where $f \geq 0$.

$$\begin{aligned}\left|\hat{F}_{\text{ss}}\right|^2(f) &= \left|\hat{F}_{\text{ds}}\right|^2(f) + \left|\hat{F}_{\text{ds}}\right|^2(-f)\\ \left|\hat{\Phi}_{\text{ss}}\right|^2(f) &= \left|\hat{\Phi}_{\text{ds}}\right|^2(f) + \left|\hat{\Phi}_{\text{ds}}\right|^2(-f)\end{aligned} \tag{44}$$

At frequencies $f > 1$ GHz, the frequency noise density is directly proportional to the laser linewidth f_{lw} as described by Formula (45).

$$f_{\text{lw}} = 2\pi\left|\hat{F}_{\text{ss}}\right|^2(f) \tag{45}$$

The original beat signal of Formula (42) is often recorded digitally, using a high-speed digitizing oscilloscope, so that the above-described signal processing can be performed off-line on a computer.

Figure 34 displays the frequency noise spectra of the two narrow-linewidth lasers whose phase noise is shown in Figure 33. The two spectra display the laser frequency noise for Fourier frequencies between 100 kHz and a few GHz. The power spectral density is multiplied here by 2π, so that the laser linewidth can be obtained directly from the high-frequency components of the noise spectra. Over the frequency range displayed in Figure 34, laser "A" exhibits substantially higher frequency noise than laser "B", which is in agreement with the linewidth specifications for these lasers (approximately 300 kHz for laser "A" and 10 kHz for laser "B"; the increase in noise above 1 GHz is an artefact, which is introduced by the measurement equipment). For both lasers, the frequency noise spectrum between 10 MHz and 1 GHz is relatively flat and smooth, which indicates that there are no significant contributions from electronic noise. This frequency range is important for compensation of large amounts of GVD, as discussed below. However, the frequency noise increases rapidly at frequencies below 1 MHz, especially for laser "A", which is caused by electronic noise in the laser control signals ($1/f$ noise).

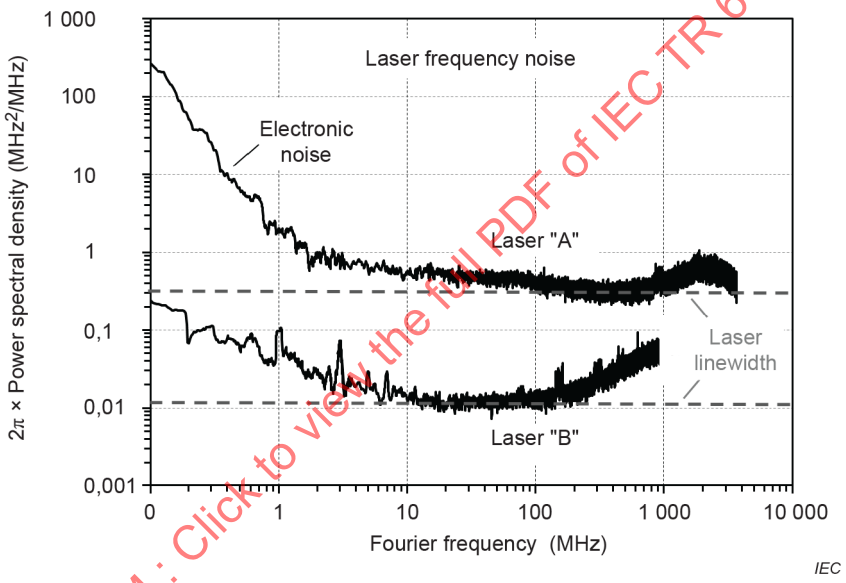


Figure 34 – Optical frequency noise spectra of two lasers

The laser frequency noise spectra shown in Figure 34 hence provide information on the laser linewidth and the presence of excessive frequency noise. For coherent communication systems, it is more important to assess the stability of the laser phase over time, which can be done more easily by analysing the closely related laser phase noise spectrum.

Figure 35 displays the phase noise spectra corresponding to the frequency noise spectra of Figure 34. For frequencies above 1 MHz, the phase noise spectral densities of both lasers decrease exponentially with increasing frequency, at a rate of about 20 dB per decade, which indicates that $\left|\hat{\Phi}_{\mathrm{ss}}\right|^2(f)$ is proportional to $1/f^2$. This dependency is consistent with the fact that the frequency noise density $\left|\hat{F}_{\mathrm{ss}}\right|^2(f)$ is approximately constant above 1 MHz, as seen in Figure 34.

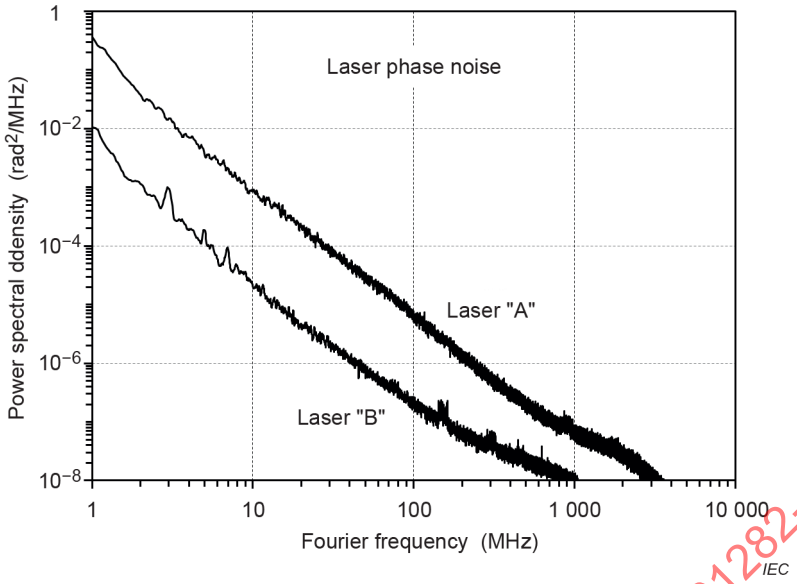


Figure 35 – Optical phase noise spectra of two lasers

From the phase noise power spectrum one can readily calculate the standard deviation σ_φ of the total random phase noise that occurs within a given frequency interval. With f_{min} denoting the lower and f_{max} the upper limit of this interval, the variance of the random phase noise in this frequency interval $[f_{min}, f_{max}]$ is given by Formula (46).

$$\sigma_\varphi{}^2 = \int_{f_{min}}^{f_{max}} \left|\hat{\Phi}_{ss}\right|^2 (f)\, df \tag{46}$$

When $\left|\hat{\Phi}_{ss}\right|^2 (f)$ is proportional to $1/f^2$, which is approximately the case in Figure 35, Formula (46) becomes Formula (47).

$$\sigma_\varphi{}^2 = f_{min}\left(1 - \frac{f_{min}}{f_{max}}\right) \left|\hat{\Phi}_{ss}\right|^2 (f_{min}) \tag{47}$$

In the case of $f_{max} >> f_{min}$, Formula (47) can be approximated by $\sigma_\varphi{}^2 \approx f_{min} \left|\hat{\Phi}_{ss}\right|^2 (f_{min})$, which shows that the total phase noise in a sufficiently wide interval is dominated by the noise around f_{min}. With f_{min} = 300 MHz, for example, from Figure 35 $\sigma_\varphi \approx$ 0,01 rad for laser "A" and $\sigma_\varphi \approx$ 0,001 rad for laser "B" are obtained.

The laser frequency or phase noise spectra can also be measured non-coherently, using a narrow-band optical high-pass or low-pass filter with sharp roll-offs at the shoulders. This method is often used when a reference laser with sufficiently low phase noise is not available for coherent mixing. In the non-coherent phase noise measurement, the laser under test is set to a frequency where the roll-off at the filter shoulder is approximately linear, as shown schematically in Figure 36, so that the laser frequency variations $\Delta\hat{f}_s(t) = d\hat{\varphi}_s(t)/dt$ are converted into optical intensity variations $\Delta P_s(t)$, as described by Formula (48).

$$\Delta P_s(t) = \Delta \hat{f}_s(t)\, P_o\, dT(f)/df \quad (48)$$

where

$T(f)$ is the normalized filter transmission, and

P_o is the average optical power.

The resulting intensity variations are then detected by a high-speed photodetector, whose output can either be recorded digitally and processed as described above or, alternatively, fed directly into an electrical RF spectrum analyser, which calculates the power spectral density of the optical intensity variations. Obviously, this method requires careful calibration of the filter slope and the received optical power in order to deduce the power spectral density of the laser frequency variations from the power spectral density of the optical intensity variations.

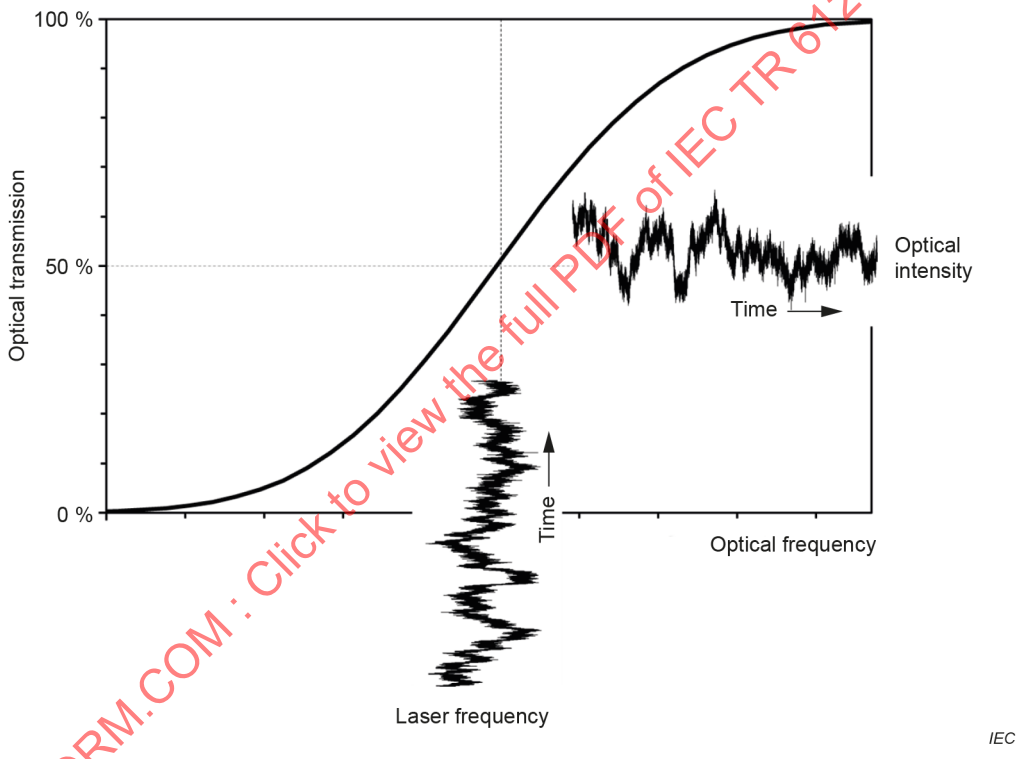


Figure 36 – Laser phase noise measurement with optical bandpass filter

6.4.6.2 Impact of phase noise in transmitter and LO lasers

As discussed in 6.4.5, recovery of the carrier phase in the coherently received signal requires the optical phases of the transmitter and LO lasers to be constant over the time period used to determine the phase offset in the received signals [73]. This time period is proportional to the number of symbols over which the phase offset is averaged, which is typically between 10 and 100. Thus, for a 32 GBd signal, the transmitter and LO laser phases should not significantly fluctuate over a time period of about 3 ns. This means that the total laser phase noise at frequencies above f_{min} ≈ 30 MHz should be sufficiently small. In Figure 35, the standard deviation of the total phase noise above 30 MHz is σ_φ ≈ 0,03 rad for laser “A” and σ_φ ≈ 0,003 rad for laser “B”. Both values are small enough for robust carrier phase recovery [4], [75].

For the purpose of carrier phase recovery, the requirements for laser phase stability are usually identical for the transmitter and LO lasers. In general, the phase stability requirements decrease with increasing symbol rate, because the averaging time for carrier phase recovery is inversely

proportional to the symbol rate. However, the tolerable laser phase noise depends strongly on the cardinality of the modulation format of the transmitted signal. Binary ASK and PSK signals, for example, can tolerate more laser phase noise than QPSK signals, and QPSK signals more than 16QAM signals [75]. Thus, transmission of 64QAM or 256QAM signals requires transmitter and receiver lasers with very low phase noise [25].

However, there are additional phase stability requirements for the LO laser when the receiver needs to compensate relatively large amounts of GVD [4]. It is important to note that in this case, the phase stability requirements for the LO laser increase with increasing symbol rate, in sharp contrast to the above-described phase stability requirements for carrier phase recovery. Consider the example of a single-polarized phase-shift-keyed (PSK) signal which has experienced a group velocity dispersion of D_L in the communication link. Neglecting shot noise and amplifier noise, this signal would produce the coherent beat signal of Formula (49).

$$i_c(t) = i_I(t) + \mathrm{j}\, i_Q(t) = R\, a_r\, \mathrm{e}^{\mathrm{j}\,\hat{\varphi}_r(t)} \int_{-\Delta t_{max}}^{\Delta t_{max}} h(\tau)\, a_s\, \mathrm{e}^{\mathrm{j}\left[\varphi_s(t-\tau)+\hat{\varphi}_s(t-\tau)\right]}\, d\tau \tag{49}$$

where

$h(t)$ describes the effects of GVD in the time domain, as defined in Formula (23),

$\hat{\varphi}_s(t)$ is the phase noise in the transmitter laser,

$\hat{\varphi}_r(t)$ is the phase noise in the LO laser,

$\Delta t_{max} = D_L\, \Delta\lambda_s/2$, with $\Delta\lambda_s$ being the width of the transmitted signal spectrum.

After electronic CD compensation, which corresponds to a convolution of $i_c(t)$ with $h^{-1}(t)$ in the time domain, as described by Formula (24), the received signal takes the form of Formula (50).

$$\left[h^{-1} * i_c\right](t) = R\, a_r\, a_s \int_{-\Delta t_{max}}^{\Delta t_{max}} h^{-1}(\tau')\, \mathrm{e}^{\mathrm{j}\,\hat{\varphi}_r(t-\tau')} \int_{-\Delta t_{max}}^{\Delta t_{max}} h(\tau)\, \mathrm{e}^{\mathrm{j}\left[\varphi_s(t-\tau-\tau')+\hat{\varphi}_s(t-\tau-\tau')\right]}\, d\tau\, d\tau' \tag{50}$$

For sufficiently small LO phase noise $\hat{\varphi}_r(t)$, Formula (50) can be approximated by Formula (51).

$$\left[h^{-1} * i_c\right](t) \approx R\, a_r\, a_s\, \mathrm{e}^{\mathrm{j}\left[\varphi_s(t)+\hat{\varphi}_s(t)\right]} \int_{-\Delta t_{max}}^{\Delta t_{max}} \mathrm{e}^{\mathrm{j}\,\hat{\varphi}_r(t-\tau')}\, d\tau' \tag{51}$$

It can be seen from Formula (51) that electronic CD compensation is not affected by phase noise in the transmitted signal, which is represented by $\hat{\varphi}_s(t)$. However, laser phase noise in the reference signal, represented by $\hat{\varphi}_r(t)$, interferes with the CD compensation process and thus introduces undesired amplitude and phase variations in the signal. This effect is sometimes called equalization-enhanced phase noise [69], or alternatively, phase noise to amplitude noise conversion [38], because the underlying phenomenon is very similar to the laser phase noise

to intensity noise conversion introduced by chromatic dispersion in direct-detection optical communication systems [76].

These undesired waveform distortions are described by the right-most term in Formula (51). Hence, the transmitted signal is recovered with high fidelity only when $\hat{\varphi}_r(t) \approx 0$ over the total integration time of the convolution in Formula (50), which is $2\Delta t_{max} = D_L \Delta\lambda_s$. Compensation of large amounts of GVD, therefore, is only possible when the phase variations in the reference signal are sufficiently small over a relatively long time period $2\Delta t_{max}$. Consequently, the phase noise of the LO laser should be small at frequencies above $f_{min} \approx 0{,}1/(D_L \Delta\lambda_s)$.

For the example of a 32 GBd signal with an optical bandwidth of $\Delta\lambda_s$ = 0,4 nm, the integration time $2\Delta t_{max}$ is around 3 ns when D_L = 8 ns/nm and 32 ns when D_L = 80 ns/nm, where the latter corresponds to about 4 000 km of dispersion-unshifted B-652 fibre at 1,5 µm wavelength. In the first case, f_{min} ≈ 30 MHz, and in the second case f_{min} ≈ 3 MHz. Since the total phase noise impacting electronic CD compensation is dominated by the phase noise around f_{min} , the phase stability requirements for the LO laser increase steeply with the amount of GVD to be compensated. At f_{min} = 3 MHz, for example, from Figure 35 $\sigma_\varphi \approx 0{,}2$ rad for laser "A" and $\sigma_\varphi \approx 0{,}02$ rad for laser "B" are obtained. Thus, the phase noise in laser "B" is small enough to allow electronic compensation of up to 8 000 ps/nm GVD in a 32 GBd signal, whereas the phase noise in laser "A" could introduce significant signal distortions when used as a local oscillator. However, laser "A" could be employed as a transmitter laser, because the phase noise of the transmitter laser does not adversely affect the process of CD compensation.

As mentioned above, the phase stability requirements for the LO laser become more stringent with increasing symbol rate, because $\Delta\lambda_s$ and Δt_{max} increase proportionally with the symbol rate of the transmitted signal, so that f_{min} decreases with increasing symbol rate. This trend is opposite to the laser phase noise requirements for carrier phase recovery (see 6.4.5), where f_{min} increases with increasing symbol rate.

6.5 High-resolution spectral analysis with coherent receivers

6.5.1 Measurement methods

Coherent mixers are inherently frequency selective and, hence, can be used as optical spectrum analysers (OSAs). It turns out that the spectral resolution of coherent OSAs can be much higher than that of conventional, grating-based optical spectrum analysers or high-finesse interferometers [77]. This feature is of great interest for test and measurement applications, since it allows detection and analysis of relatively fine features in the spectra of optical signals, such as narrow-band peaks or discrete lines, or even complex signal analysis [77].

There are two different approaches to analyse an optical spectrum with a coherent receiver.

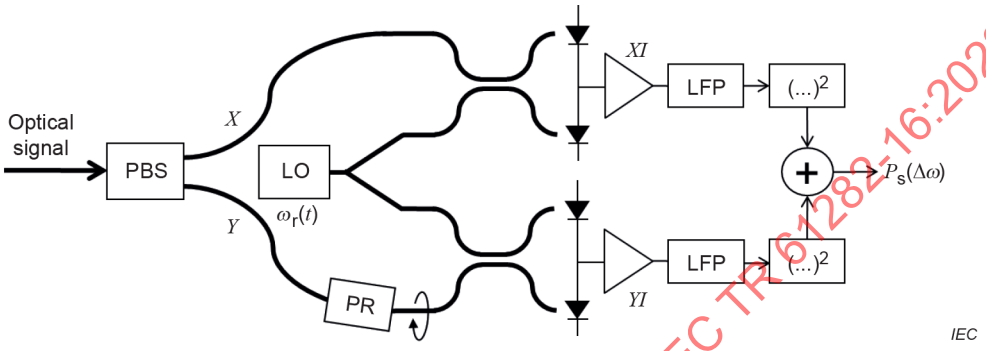
The first method is based on a high-speed coherent receiver with phase and polarization-diversity, as displayed in Figure 25, and digital recording of the time-varying optical signal over a certain measurement time T_m. The optical spectrum is then obtained from a Fourier transform of the recorded samples, as discussed in 6.3.2, which yields the complete complex amplitude spectrum of the signal to be analysed, i.e. the amplitude and phase spectral densities for the two orthogonal polarization components. This measurement technique is frequently used in optical modulation analysers (OMAs). The power spectral density of the optical spectrum can be directly calculated from the recorded samples without prior compensation of CD and PMD. The frequency resolution of the resulting spectrum is inversely proportional to the measurement time T_m, whereas the total frequency range of the spectral measurement is limited to twice the electrical bandwidth of the coherent receiver.

Another and less complex method for measuring the optical power spectrum with a coherent receiver is to use a low speed receiver in conjunction with a slowly scanning LO laser, so that

the power spectrum can be measured directly in the frequency domain [78]. These measurements do not require a phase-diversity receiver and, hence, can be performed with a coherent dual-mixer receiver having only polarization diversity, as described in 6.5.2.

6.5.2 Dual mixer with polarization diversity

Figure 37 displays a schematic diagram of a dual coherent mixer with polarization diversity, which is often used in high-resolution optical spectrum analysers (HR-OSAs). The optical spectrum is measured by recording the combined signal power detected in the balanced XI and YI mixers while the LO laser is slowly scanned over a predetermined frequency range.



Key

LO Local oscillator laser
LPF Low-pass filter
PBS Polarization beam splitter
PR Polarization rotator

Figure 37 – Dual coherent mixer with polarization diversity

The frequency resolution of this spectrum analyser is twice the electrical bandwidth B_m of the coherent receiver, which can be adjusted to the desired value by low-pass filtering the electrical beat signals $i_{XI}(t)$ and $i_{YI}(t)$. In the case of an ideal rectangular-shaped filter with bandwidth $B_m = 1/T_m$, the filtered beat signals can be described by Formula (52).

$$\bar{V}_{XI}(t) = Z \int_{-T_m/2}^{T_m/2} i_{XI}(t)\, dt = Z\, R \cos(\theta)\; a_r \int_{-T_m/2}^{T_m/2} a_s(t) \sin\left[\omega_r t + \hat{\varphi}_r(t) - \Phi_s(t)\right] dt + \hat{V}_{noise}(t) \approx$$
$$Z\, R \cos(\theta)\; a_r \left\{ \left\langle \cos\left[\hat{\varphi}_{tot}(t)\right] \right\rangle \mathrm{Im}\left[A_s(\Delta\omega)\right] + \left\langle \sin\left[\hat{\varphi}_{tot}(t)\right] \right\rangle \mathrm{Re}\left[A_s(\Delta\omega)\right] \right\} + \hat{V}_{noise}(t) \quad (52)$$

$$\bar{V}_{YI}(t) = Z \int_{-T_m/2}^{T_m/2} i_{YI}(t)\, dt = Z\, R \sin(\theta)\; a_r \int_{-T_m/2}^{T_m/2} a_s(t) \sin\left[\omega_r t + \hat{\varphi}_r(t) - \Phi_s(t) - \psi\right] dt + \hat{V}_{noise}(t) \approx$$
$$Z\, R \sin(\theta)\; a_r \left\{ \left\langle \cos\left[\hat{\varphi}_{tot}(t) + \psi\right] \right\rangle \mathrm{Im}\left[A_s(\Delta\omega)\right] + \left\langle \sin\left[\hat{\varphi}_{tot}(t) + \psi\right] \right\rangle \mathrm{Re}\left[A_s(\Delta\omega)\right] \right\} + \hat{V}_{noise}(t)$$

Formula (52) is calculated for a transmitted signal of the form of Formula (9), where $A_s(\omega)$ denotes the Fourier transform of the transmitted signal amplitude $a_s(t)$, which is given by Formula (53).

$$A_s(\omega) = \int_{-\infty}^{\infty} a_s(t)\, e^{j\varphi_s(t)} e^{-j\omega t}\, dt \tag{53}$$

The optical phase of the LO laser fluctuates during the slow tuning of its frequency ω_r, so that the two filtered beat signals of Formula (52) carry a combination of the in-phase and quadrature components of $A_s(\omega)$. After squaring and summing of the two beat signals in Formula (52), the resulting electrical signal is described by Formula (54), which is proportional to the power spectral density of the received optical signal.

$$P_s(\Delta\omega) = \left[\bar{V}_{XI}(t)\right]^2 + \left[\bar{V}_{YI}(t)\right]^2 \approx \frac{1}{2}\left(Z\,R\;a_r\right)^2 \left|A_s(\Delta\omega)\right|^2 + \hat{P}_{noise}(t) \tag{54}$$

In practice, the maximal achievable frequency resolution is limited by optical phase noise in the LO laser, which causes random frequency fluctuations. It is possible to achieve a frequency resolution of 5 MHz with such a coherent OSA. The dynamic range of these instruments is limited by electrical noise in the coherent receiver, which is usually dominated by shot noise from the LO laser. Non-coherent beat noise can be minimized by using balanced mixers with high common-mode-rejection ratio (CMRR), as discussed in 6.2.2.

6.5.3 Examples of high-resolution spectral analysis

Figure 38 displays an example of a high-resolution optical spectrum obtained with a high-speed quadruple coherent mixer, as described in 6.4. The spectrum in Figure 38 was calculated from a Fourier transform of 4 096 digitally recorded samples taken at a rate of about 64 GHz, which corresponds to a total measurement time of about T_m = 64 ns. The frequency resolution of this spectrum is about 15 MHz and thus substantially higher than that of conventional grating-based OSAs, which are typically of the order of a few GHz. Higher resolutions can be achieved by collecting a larger number of samples for the Fourier transform. In either case, a single measurement, as the one shown in Figure 38, usually yields a relatively noisy spectrum. Smoother spectra can be obtained by repeating the measurement several times and averaging the power spectral densities of the individual measurements. The maximal frequency range of the spectrum is determined by the sampling rate of the high-speed ADCs that record the signals, which is ±32 GHz in the example displayed in Figure 38.

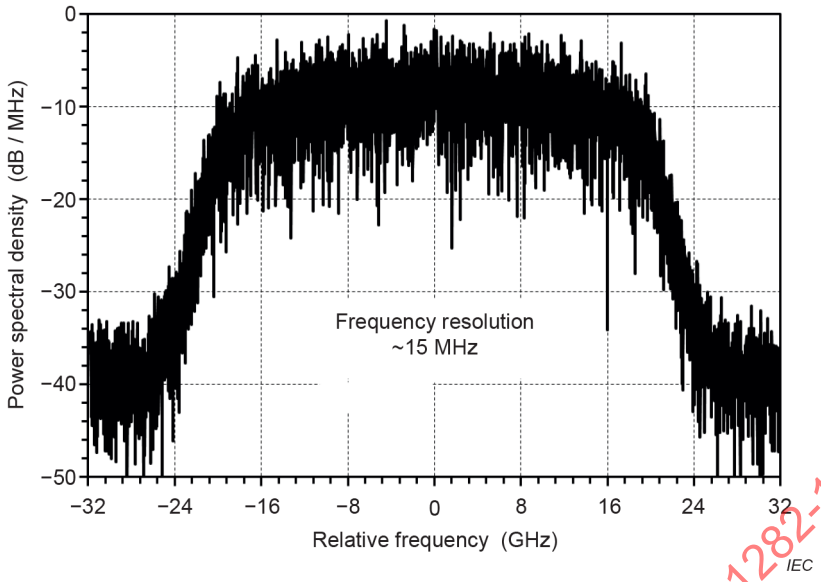


Figure 38 – High-resolution optical spectrum of a 32 GBd QPSK signal

Spectral measurements as shown in Figure 38 are often performed in optical modulation analysers (OMAs), also known as optical constellation analysers. Precise measurement of the optical power spectrum requires a well calibrated and equalized coherent receiver. The dynamic range of the measurement depends on the effective number of bits (ENOB) available in the high-speed ADCs that are used to record the time-varying signals. For an 8-bit ADC with a typical ENOB between 5 and 6, the dynamic range of the spectral measurement is of the order of 30 dB to 40 dB.

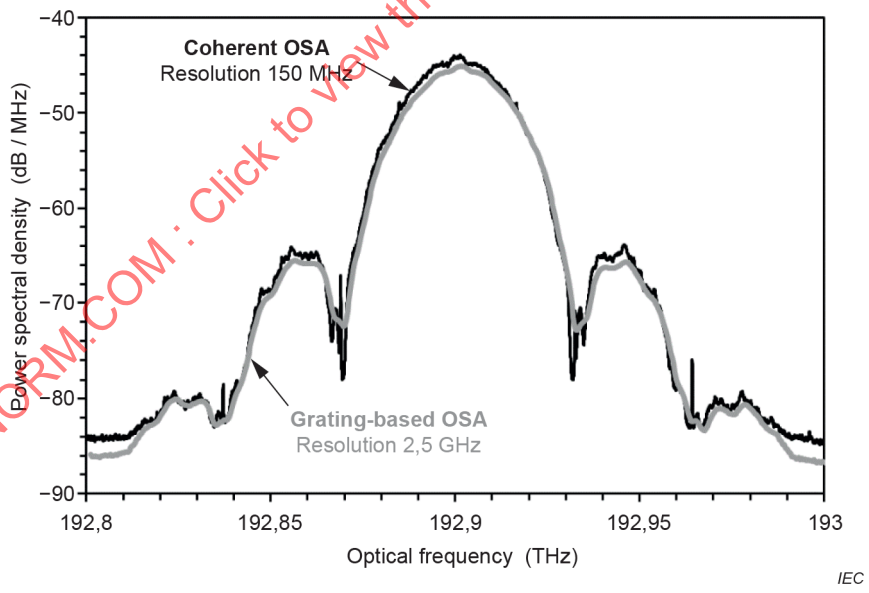


Figure 39 – QPSK signal measured with coherent OSA and with grating-based OSA

Figure 39 displays the spectrum of a 32 GBd QPSK signal measured with a commercial coherent HR-OSA, which is based on a dual coherent mixer with frequency-tuned LO laser. The spectral resolution of this instrument is about 150 MHz, which is fine enough to resolve the discrete clock tones in the modulated signal, located at ±32 GHz and ±64 GHz from the centre of the spectrum. A measurement of the same signal with a conventional grating-based OSA, shown as a grey curve in Figure 39, reveals that the grating-based OSA does not resolve these fine spectral features (which are usually undesired, because they indicate signal distortions). The maximal frequency range of a coherent OSA depends on the tuning range of the LO laser

and can be as wide as the entire C-band or L-band. The dynamic range of a coherent OSA is usually significantly higher than what can be obtained with a high-speed coherent receiver.

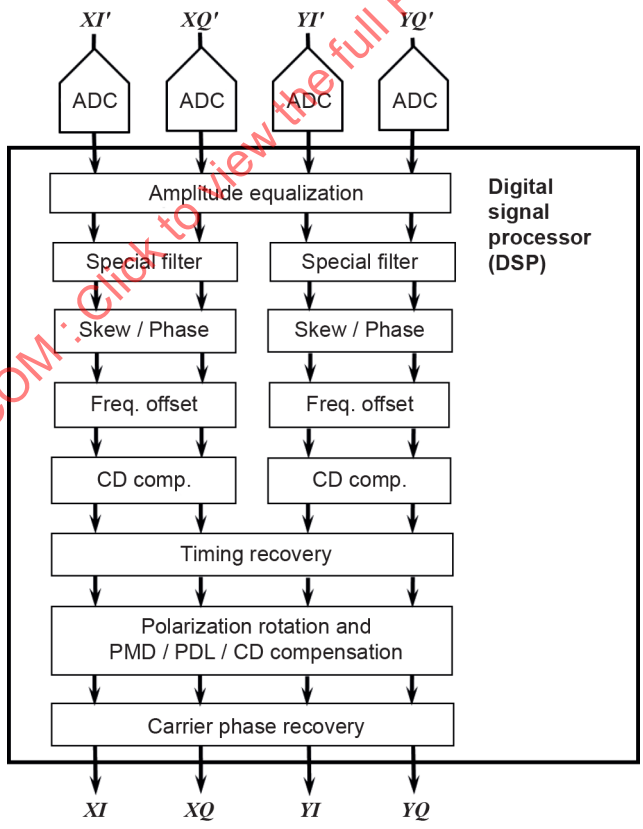
7 Digital signal processing in coherent receivers

7.1 Basic features of digital signal processing

As briefly discussed in Clause 6, the use of coherent receivers usually requires extensive signal processing to retrieve the transmitted data from the received signals. In general, the signal processing steps needed to recover the transmitted symbols are as follows [60]:

– remove undesired distortions from the transmitted signals, such as CD, PMD, PDL, and frequency-dependence of the receiver components;
– recover the timing (i.e. beginning and end) of the received symbols and synchronize the ADC sampling and signal processing with the frequency and phase of the symbol clock;
– transform the polarization state of the received signal so that two orthogonally polarized components of polarization-multiplexed signals can be separated; and
– remove undesired frequency and phase offsets from the received signals so that the in-phase and quadrature components of the transmitted signal can be separated.

These signal processing steps are usually performed after the coherently received analogue signals have been converted into digital signals, using high-speed digital analogue-to-digital converters (ADCs). Figure 40 shows a block diagram of the basic functions that are typically performed in a digital signal processor (DSP) [10].



Key

ADC Analogue-to-digital converter
XI In-phase component of *X*-polarized signal at transmitter
XI' In-phase component of *X*-polarized signal at receiver
XQ Quadrature component of *X*-polarized signal at transmitter

XQ' Quadrature component of *X*-polarized signal at receiver

YI In-phase component of *Y*-polarized signal at transmitter

YI' In-phase component of *Y*-polarized signal at receiver

YQ Quadrature component of *Y*-polarized signal at transmitter

YQ' Quadrature component of *Y*-polarized signal at receiver

Figure 40 – Typical digital signal processing steps in a coherent receiver

As indicated in Figure 40, the various signal processing steps in a coherent receiver are typically performed in the following sequence [60], [80]:

1) Equalization of the signal amplitudes obtained from the four ADCs (this step can also include removal of undesired DC offsets in the signal).
2) Individual spectral filtering of the four signals to
   a) compensate frequency dependent attenuation in the coherent receiver and ADCs, and
   b) remove excessive noise from the signals via spectral shaping.
3) Compensation of skew and phase errors introduced in the coherent receiver.
4) Removal of frequency offset between the LO laser and the transmitted signal.
5) Compensation of chromatic dispersion (CD) that the received signal has experienced in the transmission link (this is typically a fixed compensator).
6) Polarization rotation (for proper polarization de-multiplexing of the signals) and simultaneous compensation of polarization-mode dispersion (PMD), polarization-dependent loss (PDL) and small amounts of residual CD (this is always an adaptive element).
7) Removal of any residual frequency offset as well as phase offset from the signals, so as to recover the carrier phase (this is also an adaptive element).

It is possible to combine some of the processing steps of Figure 40 into a single operation. For example, spectral filtering, skew compensation, frequency offset removal, and CD compensation could be performed simultaneously as one operation in the frequency domain. However, it is important that these operations are executed in the proper sequence, because some of the operations do not commute, as for example frequency offset removal and CD compensation [4]. Therefore, the various compensation steps should be performed in the following order.

a) Signal impairments introduced in the coherent receiver should be compensated first and in the reverse order of how they occur.
b) Signal impairments introduced in the transmission line, such as CD, PMD, and polarization rotations, should be compensated after all impairments introduced in the coherent receiver have been properly compensated. Slowly varying transmission impairments, like CD, should be compensated before more rapidly varying impairments, such as those introduced by PMD, PDL, polarization rotations and optical phase variations, to facilitate fast feedback control.
c) Certain signal impairments introduced in the transmitter, such as *I*-*Q* skew and *I*-*Q* offsets, can be compensated in the receiver, but only after all impairments introduced in the receiver and in the transmission line have been properly compensated.

The most important processing steps are CD compensation, PMD compensation, and polarization tracking as well as carrier phase recovery. CD compensators are typically fixed equalizers, which introduce a predetermined and constant amount of group velocity dispersion (GVD), whereas PMD compensators/polarization rotators are adaptive equalizers, which are continuously adjusted based on feedback from a signal analyser located at the output of the equalizer (see 6.4.3) [60], [80]. Similarly, recovery of the carrier phase requires an adaptive equalizer, which is adjusted based on a feed-forward or feedback signal (see 6.4.5) [80]. These three processing steps are essential for recovering the transmitted data from the coherently received signals, which can be highly distorted by CD, PMD, polarization rotation, and phase and frequency offsets, as shown in Figure 41 for the example of a 100 Gbit/s PM-QPSK signal.

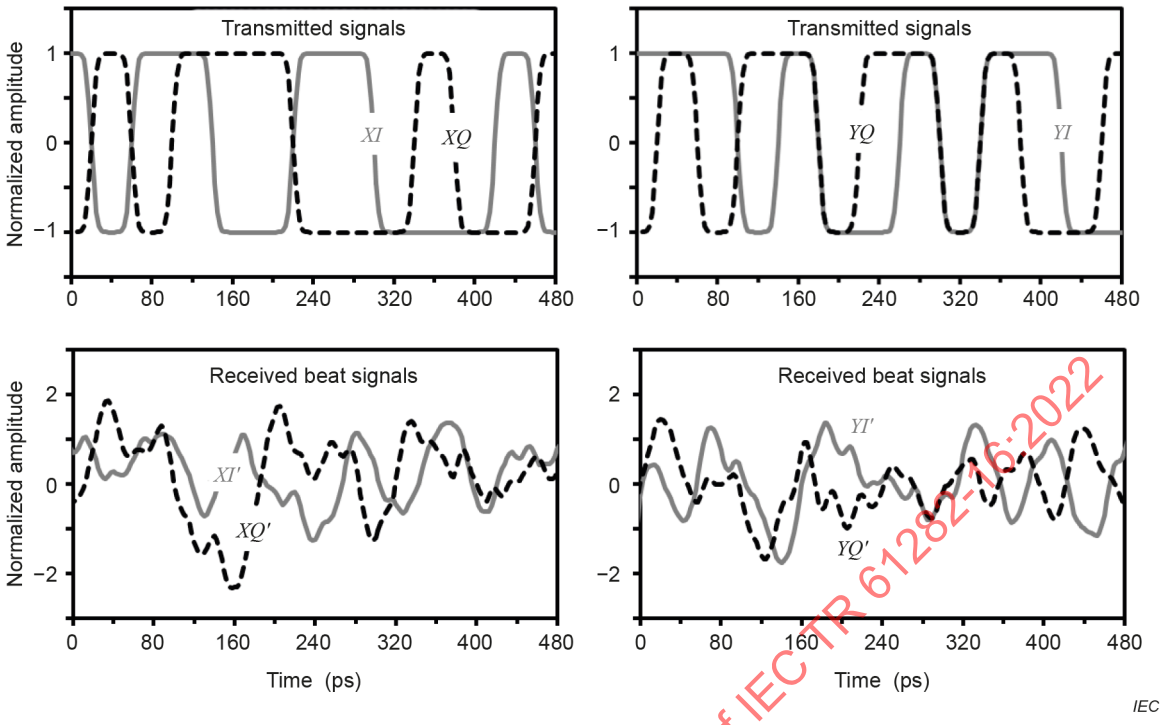


Figure 41 – 100 Gbit/s PM-QPSK signal before and after fibre transmission

The four graphs in the simulation of Figure 41 display the four tributaries of a 100 Gb/s PM-QPSK signal with a symbol rate of 25 GBd before and after transmission over 400 km of B-652 fibre, which is assumed to introduce about 8 000 ps/nm GVD and about 40 ps of differential group delay (DGD). Before transmission, the four tributaries XI, XQ, YI, and YQ to the PM-QPSK signal have clean binary ASK waveforms (see upper row of Figure 41), whereas after transmission through the fibre link, the four coherently received beat signals XI', XQ', YI', and YQ' are highly distorted by CD, PMD, and polarization rotation (see lower row of Figure 41, where an ideal receiver is assumed). It is clearly seen from Figure 41 that the waveforms of the received signal components do not show any resemblance with the waveforms of the transmitted signal components, but rather the characteristics of random noise, even though no optical or electrical noise was added to the transmitted and received signals.

Figure 42 then displays simulations of how the four received beat signals are de-convoluted at various stages in the DSP. The upper row of Figure 42 shows the four received signal components after frequency offset removal and after CD compensation (which is assumed to be perfectly matched to the GVD experienced in the transmission fibre). At this stage, the signals are still heavily distorted and, therefore, do not resemble the transmitted signal components.

The middle row of Figure 42 shows the four signal components after PMD compensation and proper polarization de-multiplexing. At this stage, the signals show signs of periodic modulation (which somewhat resembles duobinary modulated signals [8]) but are still significantly distorted. Finally, the last row of Figure 42 displays the four signal components after recovery of the carrier phase (i.e. after phase offset removal). Comparison with Figure 41 reveals that the binary ASK data have been recovered with high fidelity (although moderate low-pass filtering was introduced in the signal processing steps).

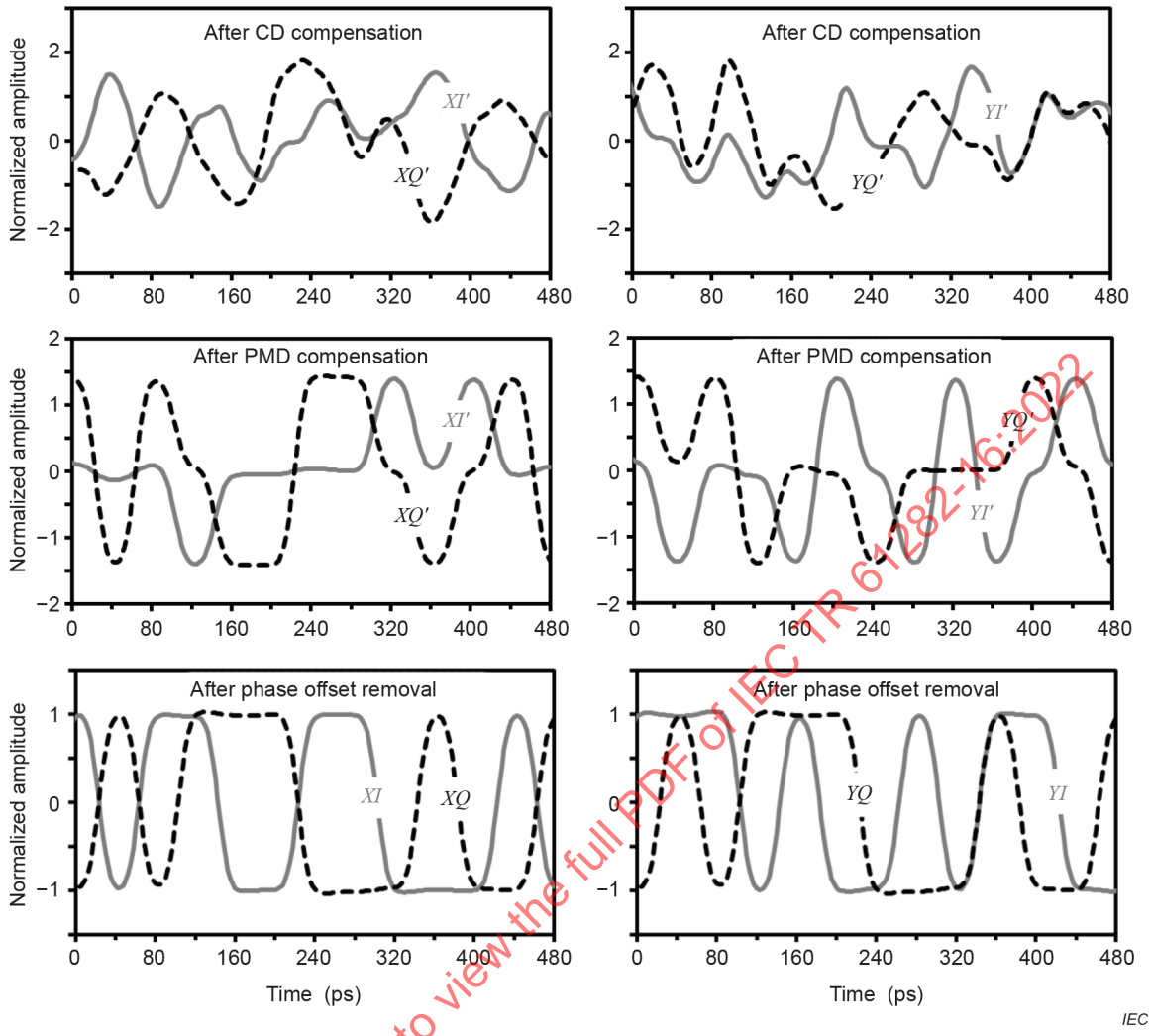


Figure 42 – De-convolution of a 100 Gbit/s PM-QPSK signal at various DSP stages

It should be noted that the simulations of Figure 38 and Figure 39 display highly oversampled signals, which can only be observed in certain software-based DSPs, which are used, for example, in optical modulation analysers (OMAs). The sampling rate of real-time, hardware-based DSPs, which are typically employed in optical communication systems, is much lower, typically ranging from 1 sample per symbol to 2 samples per symbol, and can even be different in the various processing steps [60].

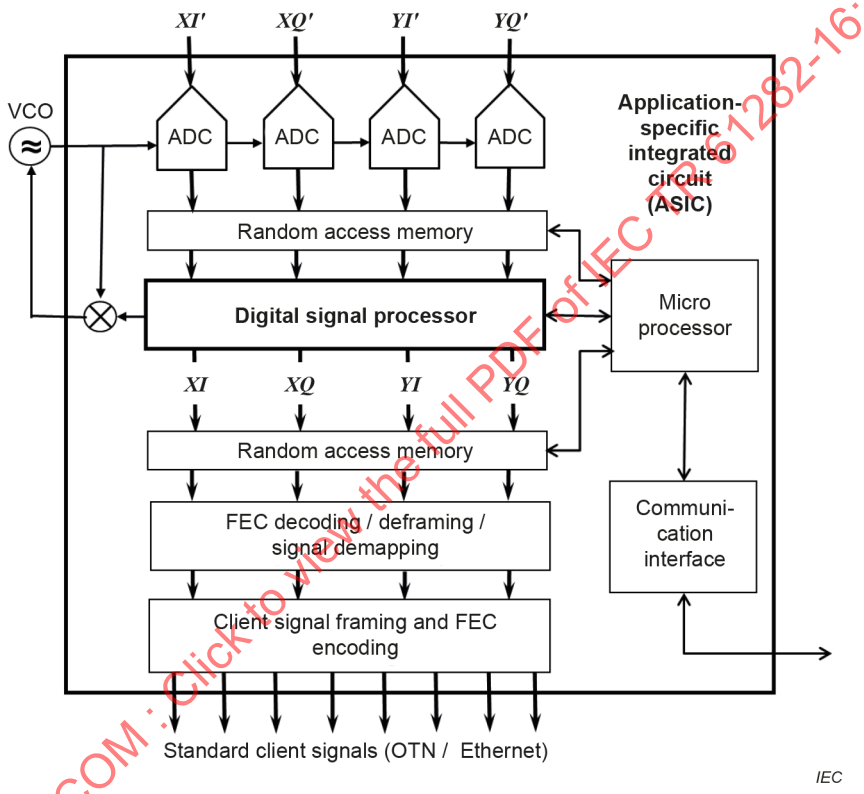
DSPs for fibre optic communication systems usually are implemented in specially designed high-speed electronic integrated circuits (ASICs), because they need to process the digital signals very rapidly at the rate of the incoming signals, so as to ensure continuous and low-latency reception [10]. DSPs for optical modulation analysers (OMAs), on the other hand, are typically implemented in software, which is often executed on a general-purpose computer. As a result, the processing speeds of software-based DSPs are usually several orders of magnitude slower than that of ASIC-based DSPs. However, software-based DSPs are more versatile than ASIC-based DSPs, so that they can be used with a larger variety of received signals, having different modulation formats as well as different symbol rates. Other differences between these DSPs are described in 7.2 and 7.3.

7.2 Real-time DSPs for fibre optic communication systems

7.2.1 Basic functions

As discussed in 7.1, DSPs for optical communication systems have to process the signals at very fast speeds and, therefore, are typically implemented in specially designed ASICs [10]. The ADCs are usually integrated with the DSP in the same ASIC, together with other functions and elements, as shown in the schematic diagram of Figure 43.

The sampling of the signal in the high-speed ADC is often synchronized with the symbol rate of the received signals [80]. This is achieved by clocking the ADC with an external voltage-controlled oscillator (VCO), which is part of a phase-locked loop (PLL) that aligns the ADC sampling time with the symbol slots of the de-convoluted signals, as shown in Figure 43.



Key

ADC Analogue-to-digital converter

FEC Forward error correction

OTN Optical transport network

VCO Voltage-controlled oscillator

Figure 43 – Block diagram of specially designed integrated circuit with DSP

The number of numerical operations performed in a DSP is usually very large, thus requiring considerable amounts of electric power. Power consumption is especially important for applications in compact coherent transceiver modules. To minimize power consumption without losing the desired functionalities, ASIC-based DSPs sample the received signals at relatively low rates, typically between 1 sample per received symbol and 2 samples per received symbol, including fractional sampling (e.g. 1,5 samples per symbol). In addition, the samples are digitized with relatively low resolution, typically not exceeding 8 bits per sample.

It is not uncommon that sampling rate and digital resolution in the DSP vary between the various processing steps that the received signal undergoes. For example, the signals are often down-sampled to just 1 sample per symbol after polarization demultiplexing (see 6.4.2).

ASIC-based DSPs frequently include additional signal processing functions, such as CD estimation, timing recovery and cycle slip detection. The CD estimator is usually located after the CD compensator and often used only during the initial start-up of the DSP to determine the optimal value of GVD to be compensated, as discussed in 6.3.3.

7.2.2 Timing recovery

Timing recovery is needed in digital communication receivers to identify the beginnings and ends of the transmitted symbols, so that they can be properly decoded [81]. Just like in conventional direct-detection receivers, timing recovery is often accomplished by detecting the transitions between symbols, using for example a Gardner timing-error detector [55], [80]. In coherent receivers, the timing-error detector is integrated into the DSP and typically located after the CD compensation stage, so as to minimize the impact of inter-symbol-interference introduced by CD [55]. Timing recovery is often combined with a phase-locked loop that synchronizes the high-speed ADC sampling times with the symbol periods [60].

7.2.3 Cycle slip detection

An increasingly important processing step in modern DSPs is the detection and correction of cycle slips, which can occur during carrier phase recovery as a result of excessive phase noise in the received signal (see 6.4.5) [79]. A cycle slip can either swap or invert the in-phase and quadrature components of the demodulated signal, so that XI becomes XQ, for example, and XQ becomes XI. This undesired swap usually lasts for only a relatively short period of time, as explained below, but can cause a large number of consecutive (and hence not correctable) errors [71].

In digital communication systems, the payload data is usually encapsulated into frames, which carry various overhead data, which include labels that uniquely identify the four transmitted lanes XI, XQ, YI, and YQ [11]. These labels are needed for the DSP to route the four de-convoluted signals to the designated output ports after polarization de-multiplexing. They also allow the DSP to detect the occurrence of cycle slips in the transmitted data, which can be introduced during the carrier phase recovery (as discussed in 6.4.5). Since the lane-identifying labels are typically transmitted at the beginning of each frame, they can only detect in which frame the cycle slip occurred, but not which symbols in the frame were affected by the slip. Thus, when a cycle slip occurs in the middle of a frame, it will only be detected at the beginning of the subsequent frame, which means that data received after the cycle slip and before the beginning of the next frame could be incorrectly decoded. This problem can be mitigated by transmitting differentially encoded signals, in which the number of incorrectly decoded symbols after a cycle slip is limited to only two [71]. However, differentially encoded signals are more sensitive to optical noise than non-differentially encoded signals and, hence, cannot be transmitted over the same long distances.

In non-differentially encoded signals, the number of incorrectly decoded symbols can be minimized by periodically interspersing training symbols or so-called pilot tones with the payload data [11]. Since the interspersed training symbols and their location within the frames are known to the DSP, it can determine whether a training symbol has been incorrectly decoded and, thus, more accurately locate the occurrence of a cycle slip. Such insertion of training symbols or pilot tones of course increases the overhead in the transmitted data (and hence the line rate), but it allows the DSP to mitigate many of the errors resulting from a cycle slip, so that the number of errors remains small enough to be correctable by forward error correction (FEC) techniques [82].

7.2.4 Compensation of nonlinear transmission effects

It is well known that nonlinear optical effects in the transmission fibre can severely impair the transmitted waveform [83]. These nonlinear effects include self-phase modulation (SPM), which

results from nonlinear interaction of the received signal with itself, as well as cross-phase modulation (XPM) and four-wave mixing (FWM), which are caused by nonlinear interactions of the transmitted signal with signals in other DWDM channels [59]. However, XPM and FWM can also result from nonlinear interactions between the various frequency components of the coherently received signal, especially in the case of multi-carrier signals. These interactions are often referred to as intra-channel XPM and intra-channel FWM [84].

If the local properties of the transmission link are sufficiently well known, such as the local optical power levels and local GVD along the fibre, it is possible to retrace at least the intra-channel nonlinear interactions along the transmission link and thereby recover the transmitted waveform. One of these nonlinear retracing methods is known as digital backpropagation (DBP) [85]. It involves repeated forward and backward Fourier transformations of the signal as it propagates through the transmission fibre as well as calculations of the local intra-channel nonlinear interactions (split-step DBP), which can in principle be performed in a DSP. In practice, this is rarely done because of the prohibitively high power consumption required for such complex calculations [4], [84].

A less power-hungry method for calculating intra-channel nonlinear distortions is the use of a Volterra series transfer function of suitable order n [84]. Another technique for mitigating intra-channel nonlinear distortions is to apply a nonlinear Fourier transform (NLFT) to the transmitted signal and, likewise, an inverse NLFT to the received signal, both of which require substantially more calculations and, hence, electrical power than linear Fourier transforms [86]. It should be noted that the aforementioned methods typically only mitigate intra-channel nonlinear distortions but not nonlinear interactions with other DWDM channels, as this would require knowledge of the waveforms of all transmitted DWDM signals. The accuracy of these methods can be significantly limited by time-varying polarization effects, such as PMD and PDL [4], [87].

Inter-channel nonlinear effects, such as XPM and FWM, can also cause significant signal distortions. In a transmission link without CD, where all DWDM signals propagate at the same speed, the XPM-induced distortions would accumulate synchronously and become potentially large. However, in a transmission fibre with large local GVD (i.e. with large CD coefficient) and without in-line CD compensation, the XPM-induced distortions accumulate more randomly, especially in systems with many DWDM channels [58]. In this case, the inter-channel nonlinear distortions exhibit similar characteristics as random optical noise [88]. In addition, the transmitted signal can be distorted by nonlinear interactions with random optical noise from in-line optical amplifiers, which is known as nonlinear phase noise (NLPN) [4].

It has been shown that such random nonlinear signal distortions can be effectively mitigated with the help of fast adaptive linear equalizers [4], [40], [88]. The power consumption of these equalizers is comparable to that of the adaptive equalizers discussed in 6.4.3 and 6.4.4. Therefore, these fast adaptive filters are employed in some high-end DSPs to compensate for inter-channel nonlinear optical interference noise.

7.2.5 FEC decoding and performance monitoring

The DSPs used for coherent communication systems typically include additional functions, such as hard-decision or soft-decision decoders for forward-error correction (FEC) and framers to convert the received signals into the desired format of the receiver client (which could be an Ethernet router or a data/telecommunication switch) [89].

In addition, DSPs often include signal processing stages for analysing the quality of the received signal, which can measure or estimate the electrical or even optical signal-to-noise ratio (SNR) in the received signal [90]. Moreover, the above-described CD and PMD compensation stages allow continuous monitoring of the GVD, DGD, and PDL that the signal has experienced in the transmission link, as well as monitoring of the frequency stability of the received signal.

7.3 Software-based DSPs for optical modulation analysers

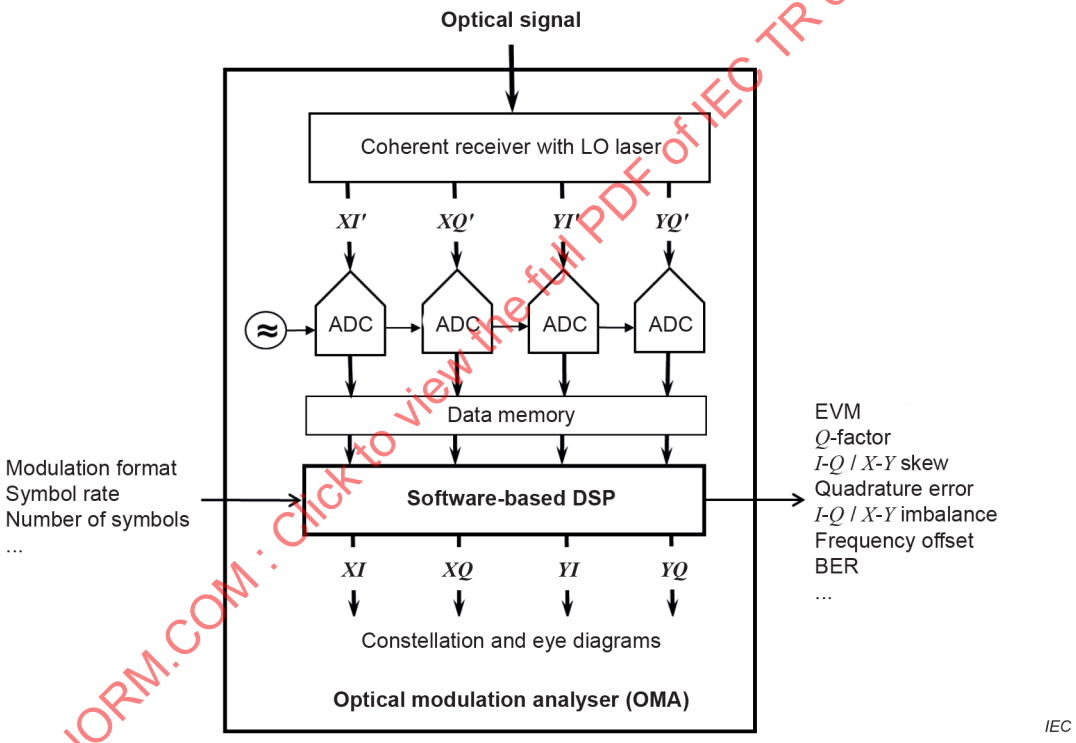
Optical modulation analysers (OMAs) are test instruments that analyse and measure the properties of vector-modulated optical signals. The capabilities of OMAs are similar to electrical modulation analysers used for complex modulated electrical or radio signals. An OMA is typically composed of the following three components:

a) a high-quality and well calibrated coherent receiver (including LO laser);

b) four high-speed ADCs with large data storage capability; and

c) a software-based DSP executed on a dedicated computer.

The coherent receiver usually comprises four balanced coherent mixers with high-speed photodetectors, but often does not include transimpedance amplifiers, so as to obtain the widest possible frequency response and the most linear transfer function. As a result, OMAs typically require substantially higher signal power than the coherent receivers used in optical communication systems. The coherent receiver usually includes a tuneable LO laser, which can be tuned to the centre frequency of the signal to be analysed, and is operated as an intradyne receiver (see 6.3.2).



Key

ADC Analogue-to-digital converter

BER Bit-error ratio

EVM Error-vector magnitude

GVD Group velocity dispersion

Figure 44 – Block diagram of OMA with software-based DSP

The four ADCs are often part of a real-time digitizing oscilloscope, which includes data storage and data management. The sampling rate and the bandwidth of the ADCs should be large enough to capture the entire spectrum of the signal to be analysed. Sampling rate and bandwidth can be fixed or variable over a certain range. It is important to note that ADC sampling is not synchronized with the symbol rate of the signal to be analysed. This means that the received signals typically need to be re-sampled at an integer multiple of the signal's symbol rate when they are analysed in the DSP.

The software-based digital signal processing is usually executed on a dedicated computer, which can be part of the digitizing oscilloscope or otherwise connected to the data storage of the oscilloscope. Such digital signal processing on an external computer is often also referred to as “offline” signal processing. In order to analyse the signal under test, the DSP requires input from the user about the modulation format of the signal and its precise symbol rate. It also requires input on the number of subsequent symbols that the user wants to be analysed, as well as on the amount of GVD that the signal has experienced. The more symbols are to be analysed the longer it takes to obtain the results.

Once triggered by the user, the output signals of the coherent receiver are synchronously sampled by the four ADCs over a time period that corresponds to the desired number of symbols (typically a few thousand symbols). The sampled signals are then recorded in the data storage and subsequently transferred to the computer for analysis by the DSP.

The DSP then re-samples the data to an integer multiple of the signal’s symbol rate (typically around 8 samples per symbol or 10 samples per symbol) and processes the data as described in 7.1, using calibration data for the coherent receiver to remove amplitude imbalance, skew, phase offset, and frequency dependence of the coherent receiver and ADCs. In addition to the processing steps shown in Figure 40, the DSP has to include a re-timing step to determine the beginning and end of a symbol period.

Once the signal analysis is completed, the OMA can display the decoded signals in the time domain as constellation diagrams (see Figure 4), eye diagrams, phase diagrams, or in the frequency domain as amplitude or power spectra. The DSP can also track variations in the polarization state of the received signal that occurred over the measurement time. Usually, the actual measurement time is much shorter than the signal processing time.

It is important to note that an OMA usually cannot determine which of the decoded signals was transmitted as the in-phase component or as the quadrature component. It also cannot distinguish between the X- and Y-polarized components if the same set of symbols is transmitted in both polarizations. It is thus not uncommon that the OMA maps the transmitted lanes incorrectly to the received lanes, as shown in the example of Formula (55).

$$\begin{pmatrix} XI_{\text{transmitted}} \\ XQ_{\text{transmitted}} \\ YI_{\text{transmitted}} \\ YQ_{\text{transmitted}} \end{pmatrix} \Rightarrow \begin{pmatrix} YI_{\text{received}} \\ YQ_{\text{received}} \\ XQ_{\text{received}} \\ XI_{\text{received}} \end{pmatrix} \tag{55}$$

However, the in-phase or quadrature components of the transmitted X-polarized signal are usually mapped into the same received polarization state (i.e. either the X- or Y-polarized state) so that they do not get paired with an in-phase or quadrature component of the transmitted Y-polarized signal. The OMA can choose different mappings between the transmitted and received lanes when another set of symbols from the same transmitted signal is analysed, although most OMAs try to identify the received lanes by unique impairments (see below), so that the same mapping as in the previous run is used.

Another useful feature of OMAs is that they can quantify the quality of vector-modulated signals by measuring some of the following parameters on the decoded signals:

- error-vector magnitude (EVM), Q-factor, and bit-error ratio (BER) estimated from EVM;
- I-Q imbalance, I-Q bias, and I-Q skew;
- quadrature error;
- X-Y imbalance and X-Y skew;

- frequency offset between signal and LO;
- accumulated differential group delays (DGD and GVD).

The EVM measures the deviation of the constellation points of the decoded signal from the ideal position and is usually expressed in units of percent (%). It is equivalent to the Q-factor, which is used for intensity- or amplitude-shift-keyed signals, and in fact uniquely related to the Q-factor, as described in [91]. Under the assumption that the spread in the constellation points has a Gaussian distribution, one can estimate the expected BER from the Q-factor. EVM is a global parameter characterizing the overall signal quality, which can be impacted by noise as well as other impairments introduced in the optical transmitter, such as

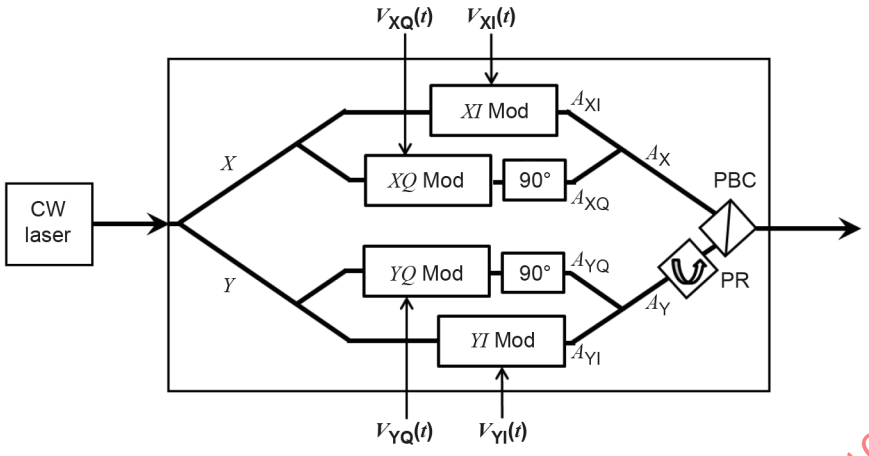
– I-Q imbalance, which characterizes the ratio of the modulation index of the in-phase signal over the modulation index of the quadrature signal and is usually expressed in units of decibel (dB).
– I-Q bias, which characterizes a common offset (or shift) of the constellation points form their ideal positions. It is calculated from the offset in the in-phase component (I-bias) and that in the quadrature component (Q-bias), and usually expressed in units of percent (%).
– I-Q skew, which is the time delay between the symbols of the in-phase component and symbols of the quadrature component and is usually expressed in units of picosecond (ps). Note that OMAs usually cannot reliably determine skews larger than half a symbol period.
– Quadrature error, which characterizes the deviation of the optical phase between the in-phase component and quadrature components of the transmitted signal from 90° and is usually expressed in units of degree (°) or radians (rad).
– X-Y imbalance, which characterizes the ratio of the average modulation index of the X-polarized signal over the average modulation index of the Y-polarized signal and is usually expressed in units of decibel (dB).
– X-Y skew, which is the time delay between the symbols of the X-polarized signal (average of the XI- and XQ-component) and the symbols of the Y-polarized signal (average of the YI- and YQ-component) and is usually expressed in units of picosecond (ps).

It is also possible to measure the BER of a transmitted signal with an OMA, provided that standard test vectors, like unframed pseudo-random binary sequence (PRBS) of a given length, are transmitted in the four lanes. However, it should be noted that such BER measurements can be different than what would be measured in a coherent receiver of an optical communication system, because of the differences in the digital signal processing. Some OMAs allow the addition of user-defined signal processing steps and/or modification of the default processing functions, so as to better emulate the performance of a communication receiver.

8 Transmitters for vector-modulated signals

8.1 Generation of vector-modulated signals

Vector-modulated signals are comprised of four real-valued components $A_{XI}(t)$, $A_{XQ}(t)$, $A_{YI}(t)$, and $A_{YQ}(t)$, each of which can carry independent information data, as described in 5.3. These four signal components are typically generated by four parallel amplitude modulators, which are imbedded in a double-layered structure of Mach-Zehnder interferometers, as shown schematically in Figure 45 [18]. The four amplitude modulators are driven by four time-varying voltages $V_{XI}(t)$ … $V_{YQ}(t)$, which usually carry independent digital information.



Key

CW	Continuous wave
Mod	Modulator
PBC	Polarization beam combiner
PR	Polarization rotator

Figure 45 – Typical arrangement for generation of vector-modulated signals

The two modulators in each of the two inner interferometers generate the in-phase and quadrature components to form the complex amplitudes $A_X(t) = A_{XI}(t) + jA_{XQ}(t)$ and $A_Y(t) = A_{YI}(t) + jA_{YQ}(t)$. The outer interferometer serves as a polarization multiplexer for $A_X(t)$ and $A_Y(t)$ [18]. The combined output signal of the four modulators thus forms a stream of four independent signals, which are simultaneously transmitted over the communication link.

The modulators themselves are often based on Mach-Zehnder interferometers (see 8.2), thus forming a triple-layer of nested Mach-Zehnder interferometers, as shown in Figure 3.

In the arrangement of Figure 45, the linearly polarized output light from a continuous-wave laser is split into four signals having equal power, which are then modulated independently to generate the four real-valued amplitudes $A_{XI}(t)$ to $A_{YQ}(t)$. The quadrature components $A_{XQ}(t)$ and $A_{YQ}(t)$ are then shifted in the optical phase by 90° before they are combined with the in-phase components $A_{XI}(t)$ and $A_{YI}(t)$ to form the complex amplitudes $A_X(t)$ and $A_Y(t)$. Finally, the polarization state of $A_Y(t)$ is rotated by 90° before it is combined with the orthogonally polarized signal $A_X(t)$ (typically in a polarization beam combiner).

The entire modulator structure of Figure 45 is usually implemented in an integrated optical substrate with single-mode optical waveguides, as discussed in 10.2.

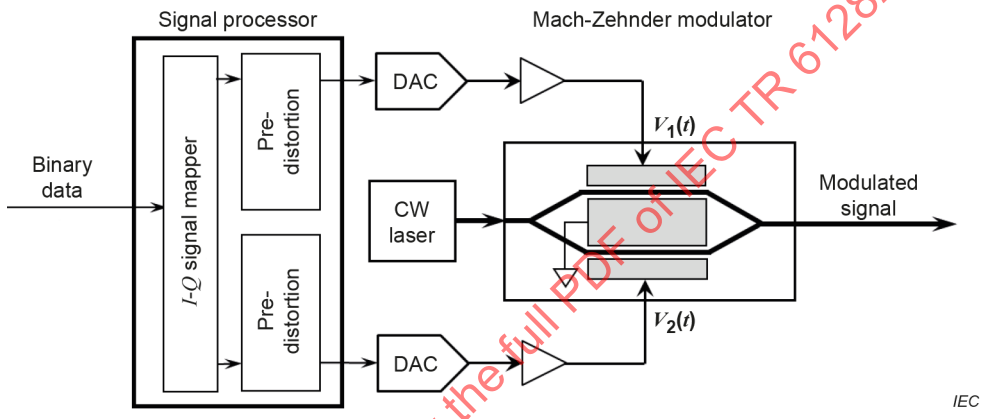
The operation of a single Mach-Zehnder modulator is described in more detail in 8.2, whereas the operation of parallel dual and quadruple Mach-Zehnder modulators is described in 8.3 and 8.4.

8.2 Single Mach-Zehnder modulator

8.2.1 Principle of operation

Mach-Zehnder modulators (MZMs) were initially introduced in fibre optic communication systems to generate chirp-free intensity modulation at symbol rates of 10 GBd and above. These single MZMs are typically implemented with single-mode optical waveguides, which are imbedded in an electro-optic material, such as lithium niobate, silicon, or III-V semiconductors (e.g. indium phosphate or gallium arsenide). Coherent communication systems typically use four MZMs, arranged in parallel, to generate complex vector-modulated signals. Subclauses 8.2, 8.3, and 8.4 describe the principle of operation of these MZMs but do not discuss the details of their specific implementation.

NOTE Modulators implemented in III-V semiconductor materials or in silicon often require additional fixed offset voltages to be added to the modulating signals. These offset voltages are different from the bias voltages discussed below, which are used to set the MZM to the desired operating point. For clarity, the offset voltages are ignored throughout Clause 8.



Key

CW Continuous wave laser

DAC Digital-to-analogue converter

Figure 46 – Differentially driven Mach-Zehnder modulator

MZMs are typically designed to modulate the amplitude of an optical signal in a fixed, predetermined polarization state. Hence, at least two MZMs are required to generate polarization-multiplexed signals (see 8.4). A single MZM typically employs two high-speed electro-optic phase shifters, one in each arm, as shown in Figure 46. These two phase shifters are needed in order to generate chirp-free signals. When these phase shifters are driven by two independent modulation voltages $V_1(t)$ and $V_2(t)$, the output signal of the MZM exhibits a combination of phase and amplitude modulation, which is given by Formula (56).

$$A_{\text{out}}(t) = A_{\text{in}} \exp\left[\mathrm{j}\frac{V_1(t)+V_2(t)}{2V_\pi}\frac{\pi}{2}\right] \cos\left[\frac{V_1(t)-V_2(t)}{2V_\pi}\frac{\pi}{2}\right] \quad (56)$$

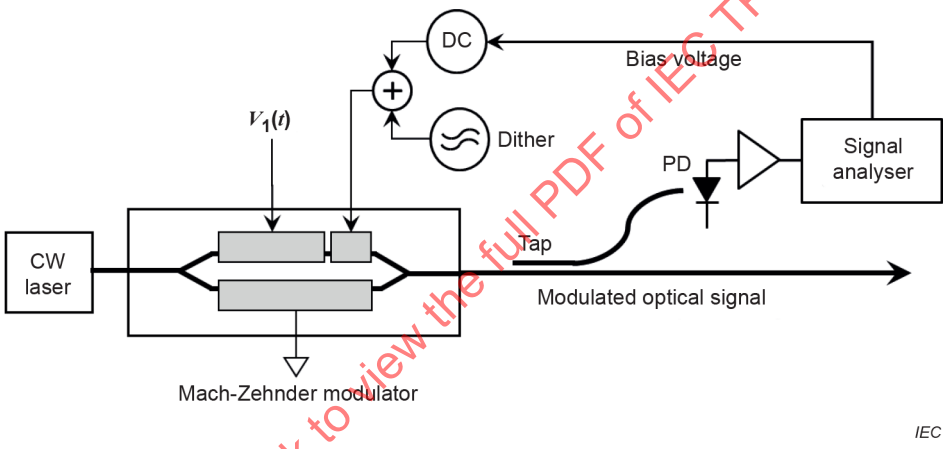
where

A_{in} denotes the unmodulated amplitude of the input signal to the modulator,

V_π is the voltage required to introduce a differential optical phase shift of π between the two arms of the MZM.

It should be noted that Formula (56) assumes perfectly matched optical path lengths in the two arms of the MZM, so that $A_{out} = A_{in}$ when $V_1(t) = V_2(t) \equiv 0$. In practice, the optical path lengths of the two arms are often not perfectly matched, which has the effect that an undesired optical phase shift is introduced between the two signals when they are combined at the output of the MZM. However, these phase shifts can be compensated by either applying an additional bias voltage V_b to the two modulating voltages, so that $V_1(t)$ is replaced by $V_1(t) + V_b/2$ and $V_2(t)$ by $V_2(t) - V_b/2$, or alternatively by employing an additional adjustable optical phase shifter in one or both arms of the interferometer, as shown in Figure 47. These additional phase shifters are typically low-speed electro-optic or thermal heater-based phase shifters, which are adjusted adaptively via automatic feedback control, as shown schematically in Figure 47 and discussed in more detail in 8.2.3. In the following, it is assumed that the undesired phase shifts are completely compensated.

The phase and amplitude excursions in Formula (56) can be independently adjusted via $V_{sum}(t) = V_1(t) + V_2(t)$ and $V_{diff}(t) = V_1(t) - V_2(t)$, respectively. Hence, a single MZM can, in principle, generate a complex modulated signal. However, this mode of operation is rarely used in practice, because it requires precise control of the amplitudes of $V_{sum}(t)$ and $V_{diff}(t)$ to generate the desired combination of phase and amplitude modulation.



Key

CW Continuous wave laser
DC Direct-current bias voltage
PD Photodiode

Figure 47 – Mach-Zehnder modulator with adaptive bias control

Instead, complex modulated signals are usually generated with the help of two parallel MZMs, as described in 8.3.1. In this case, each MZM is operated in such a way that the high-speed phase modulations in the two arms of the MZM are complementary, having phase excursions that are equal in magnitude but opposite in sign. This can be accomplished by driving the MZM either differentially with complementary voltages $V_2(t) = -V_1(t)$ or in a single-ended configuration with $V_2(t) = 0$, as shown in Figure 47. Either arrangement is capable of generating chirp-free amplitude modulation whose amplitude variations are described by Formula (57).

$$A_{out}(t) = A_{in} \cos\left[\frac{V_1(t)}{V_\pi}\frac{\pi}{2}\right] \tag{57}$$

Intensity modulation is typically used in non-coherent communication systems and can be generated by adding a constant bias voltage of $V_b = -V_\pi/2$ to the modulating voltage $V_{mod}(t)$,

so that $V_1(t) = V_{mod}(t) - V_\pi/2$. The optical output amplitude $A_{out}(t)$ and output power $P_{out}(t)$ are given by Formula (58).

$$A_{out}(t) = A_{in} \cos\left[\frac{V_{mod}(t) - V_\pi/2}{V_\pi}\frac{\pi}{2}\right]$$
$$P_{out}(t) = \left|A_{out}(t)\right|^2 = \frac{\left|A_{in}\right|^2}{2}\left\{1 + \sin\left[\frac{V_{mod}(t)}{V_\pi}\pi\right]\right\} \quad (58)$$

In this mode of operation, the total voltage swing of $V_{mod}(t)$ is typically set to V_π, so that $-V_\pi / 2 \le V_{mod}(t) \le V_\pi$ $V_\pi / 2$, as shown in Figure 48, which has the effect that the optical intensity variations are fairly insensitive to small variations in the voltage swing of $V_{mod}(t)$. However, as can be seen from Formula (58), half of the output power of the MZM remains unmodulated, so that a strong optical carrier component is present in the transmitted signal, which in certain cases can give rise to undesired nonlinear effects in the transmission fibre.

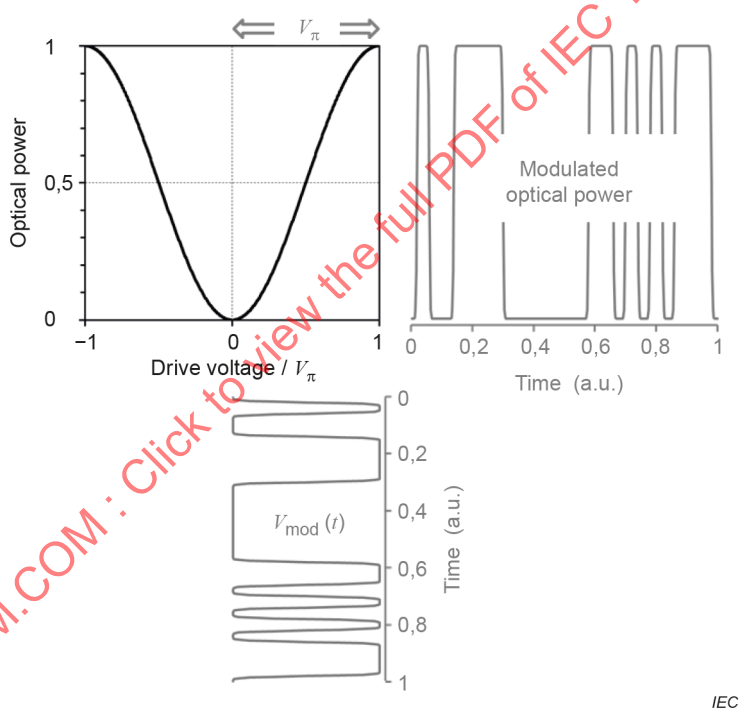


Figure 48 – MZM operation for intensity modulation

Carrier-suppressed amplitude modulation can be generated by adding a bias voltage of $V_b = -V_\pi$ to the modulating voltage $V_{mod}(t)$, so that the optical output amplitude of the modulator is described by Formula (59).

$$A_{out}(t) = A_{in} \cos\left[\frac{V_{mod}(t) - V_\pi}{V_\pi}\frac{\pi}{2}\right] = A_{in} \sin\left[\frac{V_{mod}(t)}{V_\pi}\frac{\pi}{2}\right] \quad (59)$$

In this mode of operation, the voltage swing of $V_{mod}(t)$ can be as large as $2V_\pi$, but the optimal value depends on the modulation format. For binary phase-shift keying (BPSK), for example,

the voltage swing can be set to 2 V_{π}, so that – V_{π} ≤ $V_{\text{mod}}(t)$ ≤ V_{π}, as shown in Figure 49. This has the effect that the average optical output power of the MZM is maximal, but that the amplitude response of the MZM is very nonlinear. This nonlinear response is not relevant when conventional BPSK signals are generated, but it would severely distort amplitude-shift-keyed or pulse-amplitude-modulated signals with more than two amplitude levels. Moreover, the nonlinear response limits the capabilities for signal pre-distortion in the transmitter, as will be discussed in 9.1.5.

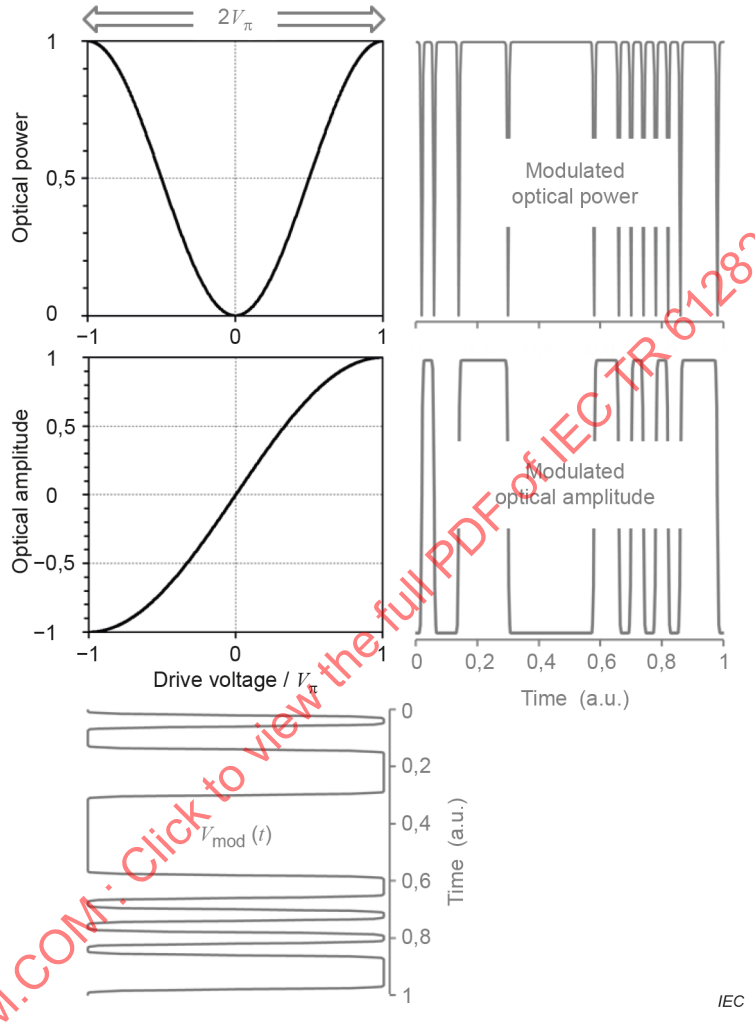


Figure 49 – Nonlinear MZM operation for binary PSK/ASK

Hence, for modulation formats with more than two levels (e.g. for quaternary amplitude shift keying (QASK)), the voltage swing is usually reduced to a much smaller range, where the optical output amplitude of the modulator becomes a linear function of the modulating voltage $V_{\text{mod}}(t)$, as described by Formula (60).

$$A_{\text{out}}(t) \approx A_{\text{in}} \frac{\pi}{2} \frac{V_{\text{mod}}(t)}{V_{\pi}} \tag{60}$$

As an example of such linear modulation, Figure 50 depicts the generation of a quaternary ASK (QASK) signal. The MZM in Figure 50 is operated in a quasi-linear regime, where the total voltage swing is limited to about 0,6 V_{π}. The upper two diagrams show the amplitude and intensity response of the MZM to the four-level QASK drive signal. For comparison, the dashed

curves show the amplitude response of an ideal linear modulator. QASK signals are frequently generated to form the in-phase and quadrature components of 16QAM signals, as described in 8.3.1.

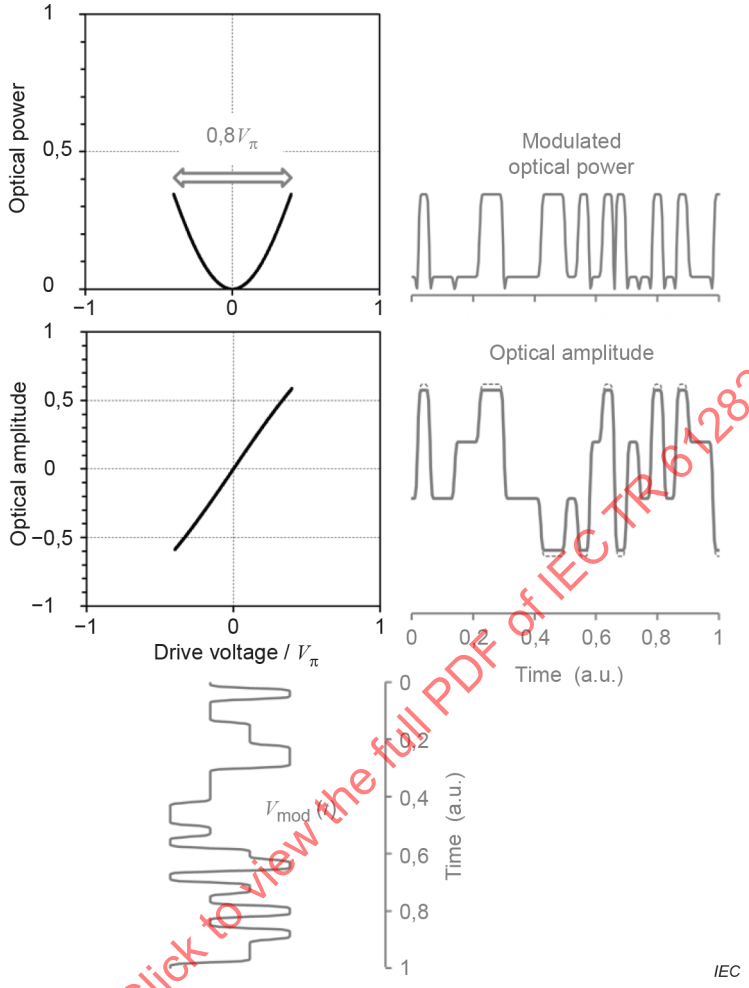


Figure 50 – Linear MZM operation for quaternary ASK signals

8.2.2 Modulator extinction ratio

In practice, the transfer function of an MZM often differs significantly from the ideal behaviour described by Formula (57) and Formula (59). Most noticeably, the optical output power of the unmodulated MZM does not completely vanish when V_b is set to $-V_\pi$. In this case, the MZM is said to have a finite extinction ratio (ER), which is usually expressed in dB. The ER is defined as the ratio between the maximal output power of the MZM, P_{max}, and the minimal output power, P_{min}, measured when V_{mod} is scanned over a range of at least $2V_\pi$. For on-off-keyed signals with $V_{mod}(t)$ swings between 0 and V_π, the extinction ratio can often be obtained directly from the maximal and minimal optical power in the modulated signal, as shown in the example of Figure 51.

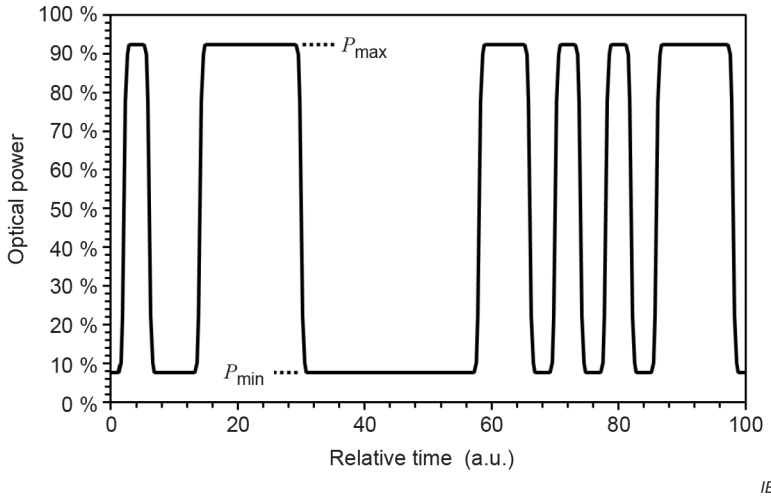


Figure 51 – Optical power variations in an NRZ-OOK signal with finite extinction ratio

Finite extinction ratios are often caused by unequal power splitting at the input and output branches of the MZM or by differential attenuation in the two modulator arms, so that for $V_b = -V_\pi$, the output amplitude of the MZM is given by Formula (61).

$$A_{out}(t) = \frac{A_{in}}{\sqrt{2}}\left\{\cos\left(\delta+\frac{\pi}{4}\right)\exp\left[j\,\frac{V_{mod}(t)+V_b}{V_\pi}\frac{\pi}{2}\right]+\sin\left(\delta+\frac{\pi}{4}\right)\exp\left[-j\,\frac{V_{mod}(t)+V_b}{V_\pi}\frac{\pi}{2}\right]\right\} = A_{in}\left\{\cos(\delta)\,\sin\left[\frac{V_{mod}(t)}{V_\pi}\frac{\pi}{2}\right]-j\sin(\delta)\cos\left[\frac{V_{mod}(t)}{V_\pi}\frac{\pi}{2}\right]\right\} \quad (61)$$

where

δ is a parameter that characterizes the power imbalance in the two arms of the MZM, in such a way that $[1-\sin(2\delta)]/[1+\sin(2\delta)]$ is the ratio of power imbalance in the two arms of the MZM.

For modulation in the linear regime, Formula (61) simplifies to Formula (62).

$$A_{out}(t) = A_{in}\cos(\delta)\,\frac{V_{mod}(t)}{V_\pi}\frac{\pi}{2} - j\,A_{in}\sin(\delta)\left\{1-\frac{1}{2}\left[\frac{V_{mod}(t)}{V_\pi}\frac{\pi}{2}\right]^2\right\} \quad (62)$$

The extinction ratio R_{ext} of the modulator is then given by Formula (63).

$$R_{ext} \equiv \frac{\max\left[P_{out}(V_{mod})\right]}{\min\left[P_{out}(V_{mod})\right]} = \frac{\cos^2(\delta)}{\sin^2(\delta)} = \cot^2(\delta) \quad (63)$$

Practical MZMs typically have extinctions ratios higher than 20 dB, corresponding to $\sin(|\delta|) < 0{,}07$. However, the magnitude of the undesired terms in Formula (61) and Formula (62) can be of the same order as that of the desired signal, especially in linear operation. Note that the undesired output signal is shifted in optical phase by 90° relative to the desired output signal. For operation in the linear regime, the output signal is predominantly an unmodulated optical carrier shifted by 90° in phase. It turns out that this unmodulated carrier can be compensated for when the MZM is employed as one of the two modulators in an I-Q modulator, as discussed in 8.3.2.

8.2.3 Adaptive bias control in Mach-Zehnder modulators

The bias voltage for the MZM is usually adjusted adaptively with the help of an automatic feedback control circuit and a blind search algorithm. Such adaptive bias control (ABC) often uses the average optical output power of the MZM as a feedback signal [92]. The average optical output power $<P_{out}(t)>$ of the modulator depends on the deviation ΔV_b from the desired bias point $V_b = \pm V_\pi$ as shown in Formula (64).

$$\left\langle P_{out}(t)\right\rangle = \left|A_{in}\right|^2 \left\langle \cos\left[\frac{V_{mod}(t)}{V_\pi}\pi\right]\right\rangle \sin\left(\frac{\Delta V_b}{V_\pi}\pi\right) + \frac{\left|A_{in}\right|^2}{2}\left\{1 - \left\langle \cos\left[\frac{V_{mod}(t)}{V_\pi}\pi\right]\right\rangle\right\} \quad (64)$$

where

<…> denotes time averaging over a large number of transmitted symbols.

A typical variation of $<P_{out}(t)>$ with bias voltage ΔV_b is depicted in Figure 52 (bold black curve). The magnitude of these variations depends on the total voltage swing ΔV_{mod} as well as on the waveform of the modulating voltage $V_{mod}(t)$, as described by the leftmost term on the right side of Formula (64).

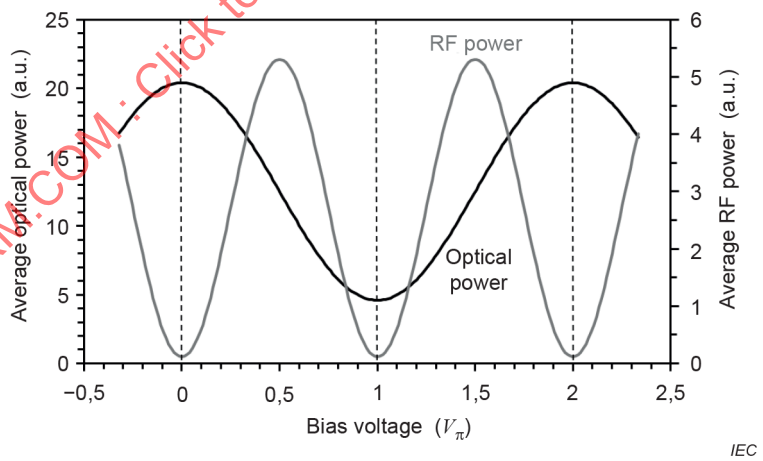


Figure 52 – Optical and RF output power versus bias voltage for linear operation

When the MZM is operated in the linear regime, with $\Delta V_{mod} < V_\pi$, the average output power is minimal at $V_b = V_\pi$, as shown in Figure 52, and it increases with increasing offset ΔV_b from the desired bias point. Thus, the desired bias point can be found by dithering the bias voltage slightly about its current value V_{cb}, as described by Formula (65).

$$V_b(t) = V_{cb} + \delta V \sin\left(2\pi f_d t\right) \tag{65}$$

where

f_d is the dither frequency, typically in the range between 10 Hz and 10 kHz, and

δV is the dither amplitude, typically less than 1 % of V_π.

These small bias voltage excursions introduce relatively small variations in the optical output power of the MZM, which can be detected by a low-speed monitor photo-diode connected to a monitor tap in the output of the MZM, as shown schematically in Figure 47. The magnitude and phase of the intensity variations at frequency f_d can serve as a feedback signal for adaptive adjustment of the bias voltage [92], [93].

Depending on the electrical circuitry connected to the monitor photodiode, the intensity variations could be in phase with the bias dither when $\Delta V_b > 0$ or in anti-phase. However, if they are in phase with the bias dither when $\Delta V_b > 0$, then they are in anti-phase when $\Delta V_b < 0$. In any case, there are no intensity variations at frequency f_d when $V_b = \pm V_\pi$. Phase-sensitive detection of the intensity variations at f_d thus provides an unambiguous error signal for adaptive adjustment of V_b. It should be noted that operation of the MZM at $V_b = +V_\pi$, instead of $V = -V_\pi$, inverts the polarity of the modulated optical output amplitude. This inversion has to be undone when the signal is decoded at the receiver.

When the MZM is operated with a voltage swing of $\Delta V_{mod} = 2V_\pi$, which is often the case for BPSK modulation, the average output power of the MZM is maximal at $V_b = \pm V_\pi$ (not minimal as in the case of linear modulation), so that the polarity of the error signal is opposite to the one used in the case of linear modulation.

It follows from the above observations that there is always a voltage swing ΔV_{mod} at which the optical output power of the MZM is insensitive to small variations in V_b and, hence, cannot be used as a feedback signal for adaptive bias control. The exact value of this voltage swing depends on the shape of the modulating waveform and is typically found at $\Delta V_{mod} > V_\pi$. If ΔV_{mod} is close to V_π, it is possible to derive an error signal from the magnitude of the high-speed modulation in the optical output signal. The high-speed variations of the modulated signal can be monitored with the help of a high-speed photodetector followed by a radio-frequency (RF) power detector [92]. Assuming linear modulation, the average RF power $\langle P_{RF}(t)\rangle$ in the high-speed photo-current varies with ΔV_b as described by Formula (66).

$$\left\langle P_{RF}(t)\right\rangle \propto \left\langle {P_{out}}^2(t)\right\rangle - \left\langle P_{out}(t)\right\rangle^2 = \frac{\left|A_{in}\right|^4}{4}\left\{1 - \left\langle \cos^2\left[\frac{V_{mod}(t)}{V_\pi}\pi\right]\right\rangle\right\} \sin\left(\frac{\Delta V_b}{V_\pi}\pi\right) + \frac{\left|A_{in}\right|^4}{4}\left\{\left\langle \cos^2\left[\frac{V_{mod}(t)}{V_\pi}\pi\right]\right\rangle - \left\langle \cos\left[\frac{V_{mod}(t)}{V_\pi}\pi\right]\right\rangle^2\right\} \tag{66}$$

where

A_{in} is the optical input amplitude to the MZM,

$P_{RF}(t)$ is the RF-power of the high-speed variations in the MZM output signal,

$P_{out}(t)$ describes the optical power variations in the MZM output signal, and

< ... > denotes time averaging over a large number of symbols.

The variations of $<P_{RF}(t)>$ with ΔV_b are plotted in Figure 52, which shows that $<P_{RF}(t)>$ is minimal when $V_b = \pm V_\pi$, as desired. However, the average RF power is not an unambiguous error signal, because it is also minimal when V_b is equal to 0 or $\pm 2V_\pi$. For this reason, it is necessary to set V_b to an initial value that is reasonably close to $-V_\pi$ or $+V_\pi$ before the control loop is closed. This could be accomplished, for example, by temporarily setting $V_{mod}(t) \equiv 0$ and adjusting V_b based on variations in the average optical output power of the MZM. However, once the control loop is closed, it automatically locks to the desired bias point. It should be noted that it is usually sufficient to analyse the RF-power in a small fraction of the total bandwidth of the high-speed modulation (e.g. up to 1 GHz), which means that the high-speed monitor detector does not require the same bandwidth as the high-speed photo detectors in the coherent optical receiver at the other end of the communication link [92].

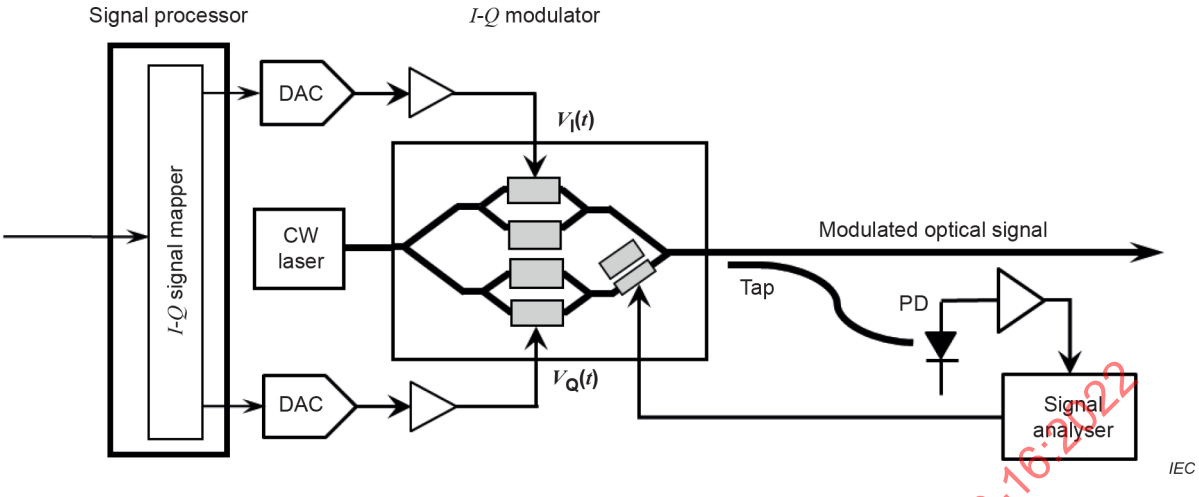
8.3 Dual Mach-Zehnder modulators

8.3.1 Quadrature-amplitude modulation

As indicated in 8.2.1, complex modulated signals are typically generated by two parallel MZMs whose outputs are combined in phase quadrature, i.e. with a relative optical phase shift of 90°, as shown in Figure 53. Neglecting optical losses in the interferometer waveguides, the combined output amplitude of these two MZMs is given by Formula (67).

$$A_{out}(t) = \frac{A_{in}}{\sqrt{2}}\left[A_I(t) + \mathrm{j}\,A_Q(t)\right] = \frac{A_{in}}{\sqrt{2}}\left\{\sin\left[\frac{V_I(t)}{V_\pi}\frac{\pi}{2}\right] + \mathrm{j}\sin\left[\frac{V_Q(t)}{V_\pi}\frac{\pi}{2}\right]\right\} \quad (67)$$

This dual MZM configuration facilitates the generation of "square" M-QAM signals with M^2 constellation points, such as the QPSK, 16QAM, and 64QAM signals, as shown in Figure 4. These signals can be generated by driving each modulator with an M-ary ASK signal, that is, with a binary, quaternary or octonary ASK signal, respectively. For example, two independently modulated binary ASK signals, generated as shown in Figure 49, can form a QPSK signal when combined in quadrature, whereas two quaternary ASK (QASK or 4ASK) signals, as shown in Figure 50, can form a 16QAM signal.



Key

CW Continuous wave laser

DAC Digital-to-analogue converter

PD Photodiode

Figure 53 – Dual MZM with adaptive *I-Q* phase control

The modulation amplitude of the two tributaries should be equal in order to obtain a perfectly square QAM signal. In the case of linear modulation, differences in the amplitudes of the two signals, which could be introduced by unequal splitting ratios or differential losses, can be compensated for by adjusting the magnitude of the voltage swing of either $V_I(t)$ or $V_Q(t)$.

Modulation formats with circular constellation diagrams, like M-ary PSK, can also be generated with this dual MZM configuration. However, such operation requires proper pre-distortion of the two drive signals $V_I(t)$ and $V_Q(t)$, which can be readily seen by expressing Formula (67) in polar coordinates, as shown in Formula (68).

$$A_{out}(t) \approx \frac{A_{in}}{\sqrt{2}} \frac{\pi}{2V_\pi} \sqrt{V_I(t)^2 + V_Q(t)^2} \exp\left\{ j \arctan\left[\frac{V_Q(t)}{V_I(t)} \right] \right\} \quad (68)$$

Hence, M-ary PSK can be generated by keeping the root-mean-square (RMS) amplitude of the two drive voltages constant and by varying only the ratio of the two voltages. Such pre-distortion of the drive signals can be readily performed in high-speed digital signal processors, where an I-Q signal mapper is used to assign the incoming binary data to the constellation points of the transmitted optical signal.

8.3.2 Compensation of finite extinction ratio

As discussed in 8.2.2, MZMs usually exhibit a finite extinction ratio (ER). When two parallel MZMs are operated as described in 8.3.1 and in the linear regime, each of the two MZMs can cross-compensate the undesired unmodulated carrier component resulting from the other's finite modulator extinction ratio. This compensation is accomplished in such a way that an offset in the bias voltage of the Q-modulator reduces the undesired carrier signal from the I-modulator and, likewise, an offset in the bias voltage in the I-modulator reduces the undesired carrier signal from the Q-modulator. Letting $V_{offs\ I}$ and $V_{offs\ Q}$ denote the additional bias offsets in the I- and Q-modulators, respectively, the undesired carrier signals are compensated when the two offsets are set to the values given in Formula (69).

$$V_{\text{offs I}} = -V_{\pi}\frac{2\sin\left(\delta_{\text{Q}}\right)}{\pi\cos\left(\delta_{\text{I}}\right)} \approx -\frac{2V_{\pi}}{\pi\sqrt{R_{\text{ext Q}}}}$$
$$V_{\text{offs Q}} = -V_{\pi}\frac{2\sin\left(\delta_{\text{I}}\right)}{\pi\cos\left(\delta_{\text{Q}}\right)} \approx -\frac{2V_{\pi}}{\pi\sqrt{R_{\text{ext I}}}} \quad (69)$$

where

δ_{I} and δ_{Q} characterise the finite extinction ratios of the individual I- and Q-modulators.

With the offsets of Formula (69) applied, the output amplitude of the IQ-modulator is described by Formula (70).

$$A_{\text{out}}(t) = \frac{A_{\text{in}}}{2}\left[\cos\left(\delta_{\text{I}}\right)\frac{V_{\text{I}}(t)}{V_{\pi}}\frac{\pi}{2} + \text{j}\cos\left(\delta_{\text{Q}}\right)\frac{V_{\text{Q}}(t)}{V_{\pi}}\frac{\pi}{2} + \sin\left(\delta_{\text{I}}\right)\varepsilon_{\text{Q}}(t) - \text{j}\sin\left(\delta_{\text{Q}}\right)\varepsilon_{\text{I}}(t)\right] \quad (70)$$

where the variables ε_{I} and ε_{Q} are defined as shown in Formula (71).

$$\varepsilon_{\text{I}}(t) = \frac{1}{2}\left[\frac{V_{\text{I}}(t)+V_{\text{offs I}}}{V_{\pi}}\frac{\pi}{2}\right]^2 - \left\langle\frac{1}{2}\left\{\frac{V_{\text{I}}(t)+V_{\text{offs I}}}{V_{\pi}}\frac{\pi}{2}\right\}^2\right\rangle$$
$$\varepsilon_{\text{Q}}(t) = \frac{1}{2}\left[\frac{V_{\text{Q}}(t)+V_{\text{offs Q}}}{V_{\pi}}\frac{\pi}{2}\right]^2 - \left\langle\frac{1}{2}\left\{\frac{V_{\text{Q}}(t)+V_{\text{offs Q}}}{V_{\pi}}\frac{\pi}{2}\right\}^2\right\rangle \quad (71)$$

The first two terms on the right side of Formula (70) describe the desired signal components, whereas the last two terms are residual signal perturbations, whose magnitudes are both much smaller than 1. Hence, the undesired carrier components in Formula (70) are substantially smaller than those in Formula (62). It should be noted that the additional bias offsets $V_{\text{offs I}}$ and $V_{\text{offs Q}}$ can be generated automatically by an adaptive bias control loop, as described in 8.3.1, provided that the feedback signal for bias control is obtained from the combined output of the two MZMs, as shown in Figure 53.

The parent (or outer) Mach-Zehnder interferometer, into which the two MZMs are imbedded, can also be imbalanced and, hence, exhibit finite extinction ratio. However, since the I- and Q-signals are shifted in phase by 90°, this interferometer imbalance will only lead to an imbalance of the amplitudes of the I- and Q-signals, but not to undesired signal components. The extinction ratio of the parent interferometer is typically better than 20 dB, just like that of the individual MZMs, so that the amplitude imbalance is at most 10 %. In the case of linear modulation, the imbalance can also be compensated for by an adjustment of the voltage swings of $V_{\text{I}}(t)$ or $V_{\text{Q}}(t)$.

8.3.3 Adaptive control of I-Q phase

The 90° optical phase shift between the output signals of the I- and Q-modulators is usually introduced by a variable optical phase shifter, instead of a fixed optical element as in the coherent receiver. The reason for using a variable optical phase shifter is that the high-speed

modulation of the MZMs often generates heat, which could alter the phase relationship between the I- and Q-modulated signals. Hence, the optical phase between these two signals, which is often referred to as the I-Q phase, needs to be adjusted adaptively. This is usually accomplished by using a blind search algorithm. There are two different methods to obtain a suitable feedback signal for automatic I-Q phase adjustment [92], [93].

In the first method, the I-Q phase is dithered slightly about its current value, at a frequency somewhere in the range between 10 Hz and 10 kHz, which leads to small variations of the average RF power $\langle P_{RF}(t)\rangle$ in the combined outputs of the two MZMs [92], [94]. It turns out that the average RF power is minimal when the I-Q phase is equal to 90°. Thus, when the I-Q phase is detuned from 90°, for example by an offset $\Delta\varphi$, $\langle P_{RF}(t)\rangle$ increases proportionally with $\sin^2(\Delta\varphi)$, as described by Formula (72) and also plotted in Figure 54.

$$\begin{aligned}\left\langle P_{\mathrm{RF}}(t)\right\rangle \propto \left|A_{\mathrm{in}}\right|^4 &\left\langle \sin^2\left[\frac{V_{\mathrm{I}}(t)}{V_\pi}\frac{\pi}{2}\right]\right\rangle \left\langle \sin^2\left[\frac{V_{\mathrm{Q}}(t)}{V_\pi}\frac{\pi}{2}\right]\right\rangle \sin^2(\Delta\varphi) + \\ &\frac{\left|A_{\mathrm{in}}\right|^4}{4}\left\{\left\langle \sin^4\left[\frac{V_{\mathrm{I}}(t)}{V_\pi}\frac{\pi}{2}\right]\right\rangle - \left\langle \sin^2\left[\frac{V_{\mathrm{I}}(t)}{V_\pi}\frac{\pi}{2}\right]\right\rangle^2\right\} + \\ &\frac{\left|A_{\mathrm{in}}\right|^4}{4}\left\{\left\langle \sin^4\left[\frac{V_{\mathrm{Q}}(t)}{V_\pi}\frac{\pi}{2}\right]\right\rangle - \left\langle \sin^2\left[\frac{V_{\mathrm{Q}}(t)}{V_\pi}\frac{\pi}{2}\right]\right\rangle^2\right\}\end{aligned} \tag{72}$$

Formula (72) is derived with the assumption that the I- and Q-signals are modulated with statistically independent data, which is usually the case when considering a large number of symbols. The RF power increases with the magnitude of $\Delta\varphi$, because the output signals of the two MZMs interfere coherently when $\Delta\varphi \neq 90°$ and, hence, introduce additional high-speed power variations in the combined signal. This RF-power-based control method is generally applicable to all operating modes of the MZMs, as long as the modulating signals $V_{\mathrm{I}}(t)$ and $V_{\mathrm{Q}}(t)$ are statistically independent. It should be noted that $\langle P_{RF}(t)\rangle$ is also minimal when $\Delta\varphi = -90°$, in which case $A_{\mathrm{out}}(t) = A_{\mathrm{I}}(t) - \mathrm{j}A_{\mathrm{Q}}(t)$, so that the phase of $A_{\mathrm{Q}}(t)$ is inverted. This polarity inversion cannot be avoided and, hence, needs to be corrected in the coherent receiver, as discussed in 8.2.3.

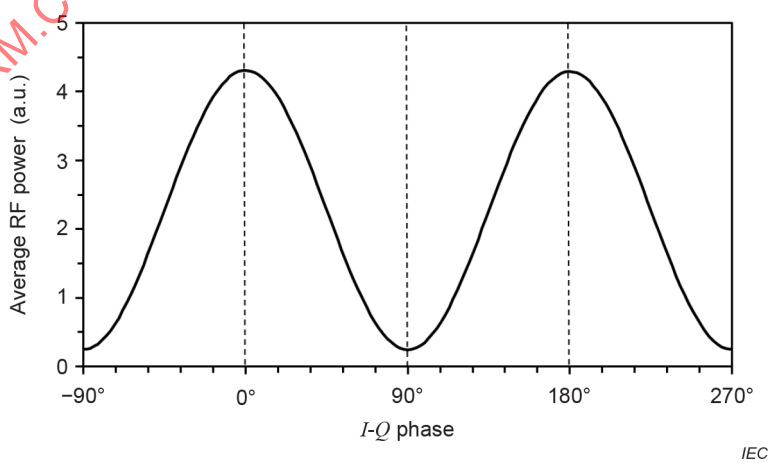


Figure 54 – RF output power of I-Q modulator versus I-Q phase

The second method for adaptive I-Q phase control is based on a secondary effect resulting from the bias dither in the two I-Q modulators [93], [95]. Therefore, this method does not require

dithering of the I-Q phase. However, it requires that the voltage swing of $V_{\text{mod}}(t)$ is significantly lower than $2V_{\pi}$, for example less than $1{,}6V_{\pi}$ [93]. Hence, this method is particularly well suited for modulation in the linear regime, which is often used to generate vector-modulated signals of higher cardinality. In the following, it is assumed that the bias dither in the I- and Q-modulator has the sinusoidal form described by Formula (73).

$$\begin{aligned} V_{\text{bias I}}(t) &= V_{\text{bI}} + \Delta V_{\text{I}} \sin\left(\omega_{\text{I}} t\right) \\ V_{\text{bias Q}}(t) &= V_{\text{bQ}} + \Delta V_{\text{Q}} \sin\left(\omega_{\text{Q}} t\right) \end{aligned} \tag{73}$$

where

$V_{\text{bias I}}$ and $V_{\text{bias Q}}$ are the dithered bias voltages with offsets V_{bI} and V_{bQ},

ΔV_{I} and ΔV_{Q} denote the dither excursions in the I- and Q-modulators, and

ω_{I} and ω_{Q} are the corresponding angular dither frequencies.

When the two MZMs are operated at their optimal bias points and at an I-Q phase of either +90° or –90°, the average optical output power of the I-Q modulator varies periodically at twice the dither frequencies ω_{I} and ω_{Q}, as discussed in 8.2.3. However, when the I-Q phase is detuned from +90° or –90°, then there are also optical power variations at the sum and difference frequencies $\omega_{\text{I}} + \omega_{\text{Q}}$ and $\omega_{\text{I}} - \omega_{\text{Q}}$ of the bias dither, as described by Formula (74).

$$\begin{aligned} \left\langle P_{\text{out}}(t)\right\rangle \approx \frac{\left|A_{\text{in}}\right|^2}{4}\Bigg\{ &\left\langle\left[\frac{V_{\text{I}}(t)}{V_{\pi}}\frac{\pi}{2}\right]^2\right\rangle + \left\langle\left[\frac{V_{\text{Q}}(t)}{V_{\pi}}\frac{\pi}{2}\right]^2\right\rangle + \left[\frac{\Delta V_{\text{I}}}{V_{\pi}}\frac{\pi}{2}\right]^2 \sin^2\left(\omega_{\text{I}} t\right) + \\ &\left[\frac{\Delta V_{\text{Q}}}{V_{\pi}}\frac{\pi}{2}\right]^2 \sin^2\left(\omega_{\text{Q}} t\right) - 2\sin(\Delta\varphi)\,\frac{\Delta V_{\text{I}}\,\Delta V_{\text{Q}}}{{V_{\pi}}^2}\frac{\pi^2}{4}\sin\left(\omega_{\text{I}} t\right)\sin\left(\omega_{\text{Q}} t\right)\Bigg\} \end{aligned} \tag{74}$$

The sum and difference frequencies are contained in the right-most term of Formula (74), which can be seen from Formula (75).

$$\sin\left(\omega_{\text{I}} t\right)\sin\left(\omega_{\text{Q}} t\right) = \frac{1}{2}\left[\cos\left(\omega_{\text{I}} t - \omega_{\text{Q}} t\right) - \cos\left(\omega_{\text{I}} t + \omega_{\text{Q}} t\right)\right] \tag{75}$$

The right-most term of Formula (74) disappears when the I-Q phase is ±90° but increases with increasing offset $\Delta\varphi$ from 90°. Therefore, an unambiguous error signal for adaptive phase control can be obtained by phase sensitive detection of the optical power variations at angular frequencies $\omega_{\text{I}} + \omega_{\text{Q}}$ and $\omega_{\text{I}} - \omega_{\text{Q}}$, which can be accomplished by multiplying the average optical output power at the monitor detector $\langle P_{\text{out}}(t)\rangle$ with Formula (75) and integrating the result over many dither periods, as described by Formula (76).

$$P_{\text{err}} = \int_0^T \left\langle P_{\text{out}}(t)\right\rangle \sin\left(\omega_{\text{I}} t\right)\sin\left(\omega_{\text{Q}} t\right) dt \tag{76}$$

where

P_{err} denotes the error signal, and

T is the integration time with $T >> 2\pi/\omega_I$ and $T >> 2\pi/\omega_Q$.

Figure 55 displays the dependence of P_{err} on the I-Q phase. Since the variations in $<P_{out}(t)>$ scale with the square of the bias dither excursion ΔV_I and ΔV_Q, the error signal P_{err} from Formula (76) tends to be fairly small. Hence, when the feedback signal $<P_{out}(t)>$ is fairly noisy, long integration times T should be used to obtain a robust error signal (e.g. $T > 200\pi/\omega_I$ or $T > 200\pi/\omega_Q$).

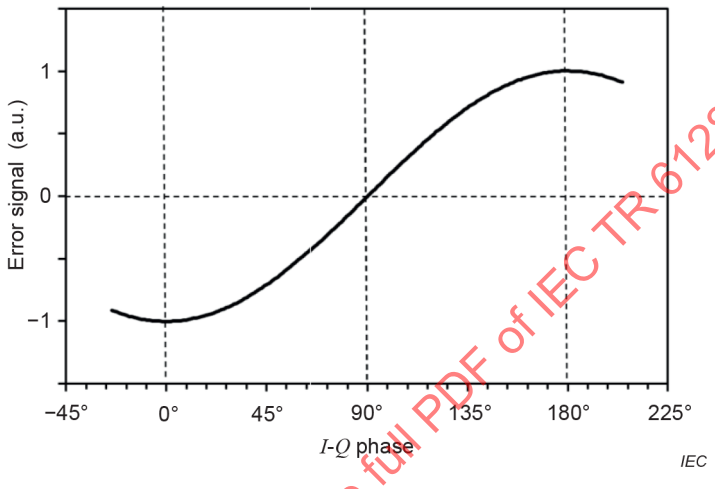


Figure 55 – Error signal for I-Q phase control derived from MZM bias dither

Instead of using two different dither frequencies for the two MZMs, it is also possible to dither the bias voltages of the two MZMs at the same angular frequency ω but in phase quadrature, so that Formula (73) becomes Formula (77) [95], [96].

$$V_{bias\,I}(t) = V_{bI} + \Delta V_I \sin(\omega t)$$
$$V_{bias\,Q}(t) = V_{bQ} + \Delta V_Q \cos(\omega t) \quad (77)$$

In this case, the error signal can be obtained by detecting the power variations occurring at frequency 2ω and being phase synchronous with $\sin(2\omega t)$.

It should be noted that an undesired offset in P_{err} can occur when one or both of the MZMs exhibit relatively low ER and when the phase dither is introduced in only one of the two arms of the MZMs, as shown in the example of Figure 56, which displays P_{err} as a function of $\Delta\varphi$. If not corrected, this offset can lead to a significant phase error when the I-Q phase is automatically adjusted as described above. This phase error is clearly visible in Figure 56, where the I-Q phase at $P_{err} \approx 0$ deviates substantially from the desired ±90°.

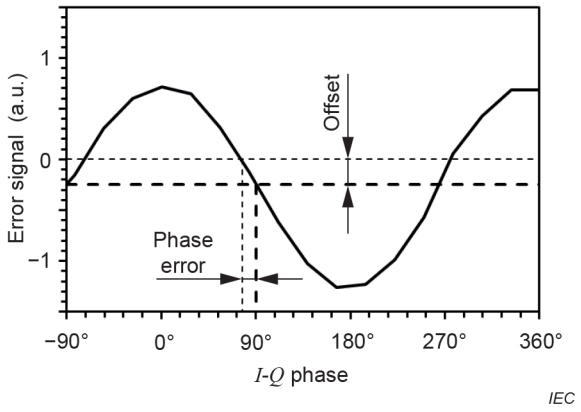


Figure 56 – *I-Q* phase error resulting from offset in feedback signal

However, the fact that the above-described second method does not require dithering of the *I-Q* phase means that the signal distortions caused by bias dithering are considerably lower than in the first method [95]. In the case of linear modulation, one can easily calculate the impact of bias and phase dithering in the output amplitude of the *I-Q* modulator, which is given by Formula (78).

$$A_{\text{out}}(t) \approx \frac{A_{\text{in}}}{4}\frac{\pi}{V_{\pi}}\left[V_{\text{I}}(t) + \text{j}V_{\text{Q}}(t) + \Delta V_{\text{I}}\sin\left(\omega_{\text{I}}t\right) + \text{j}\Delta V_{\text{Q}}\sin\left(\omega_{\text{Q}}t\right) - V_{\text{Q}}(t)\Delta\varphi\sin\left(\omega_{\text{IQ}}t\right)\right] \quad (78)$$

where

$\Delta\varphi$ denotes the magnitude of the phase dither, and

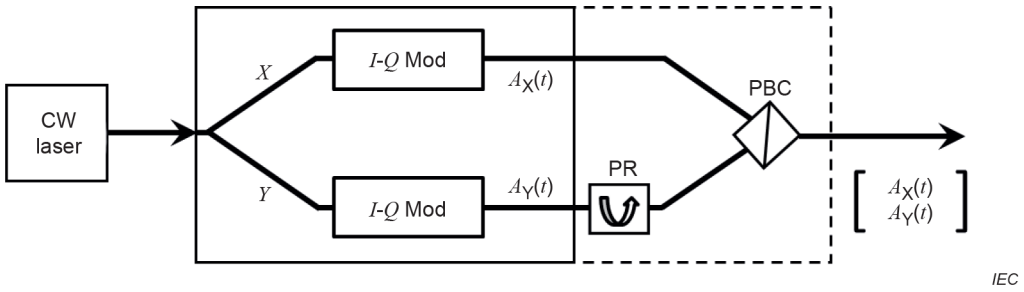
ω_{IQ} denotes the angular frequency of the *I-Q* phase dither.

In the case of the second *I-Q* phase adjustment method, the right-most term in Formula (78) vanishes, because $\Delta\varphi = 0$.

8.4 Quadruple modulators for polarization-multiplexed signals

Polarization-multiplexed (or dual-polarization) signals are typically generated by two pairs of *I-Q* modulators, one of which generates a modulated signal in the *X* polarization state and the other a differently modulated signal in the *Y* polarization state. This arrangement thus creates two complex-modulated amplitudes $A_{\text{X}}(t)$ and $A_{\text{Y}}(t)$, which are transmitted in two orthogonal polarization states.

However, most modulators are polarization-dependent, in the sense that their modulation is most efficient in one particular polarization state. For this reason, the two signals are usually created by splitting the output of the (single-polarized) transmitter laser between two identical *I-Q* modulators, which are integrated on the same chip, as shown in Figure 57, so that the modulated signals $A_{\text{X}}(t)$ and $A_{\text{Y}}(t)$ initially have the same polarization state, which is typically a linear state, either parallel with or orthogonal to the surface of the modulator substrate. The polarization state of one of the signals, for example $A_{\text{Y}}(t)$, is then rotated by 90°, with the help of a static polarization rotator (PR), and subsequently combined with the orthogonally polarized signal $A_{\text{X}}(t)$ in a polarization beam combiner (PBC), as shown in Figure 57. The PR and PBC are sometimes integrated with the modulators on the same substrate or, alternatively, added as separate components at the modulator output, using either integrated-optic waveguides or discrete micro-optic components.



Key

CW Continuous wave

Mod Modulator

PBC Polarization beam combiner

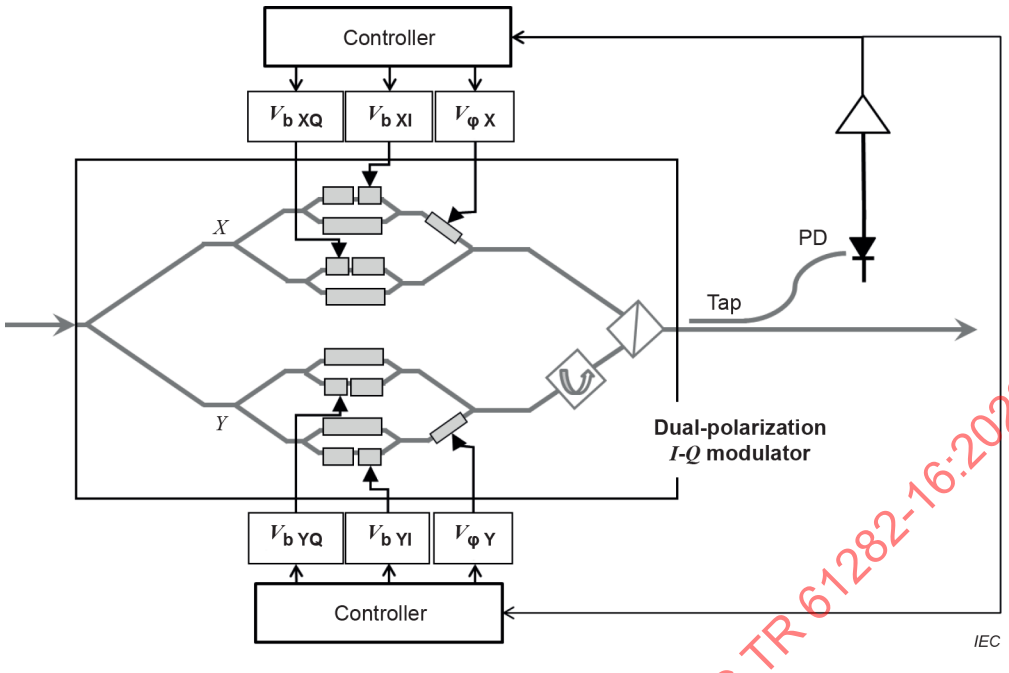
PR Polarization rotator

Figure 57 – Dual *I-Q* modulators for polarization multiplexing

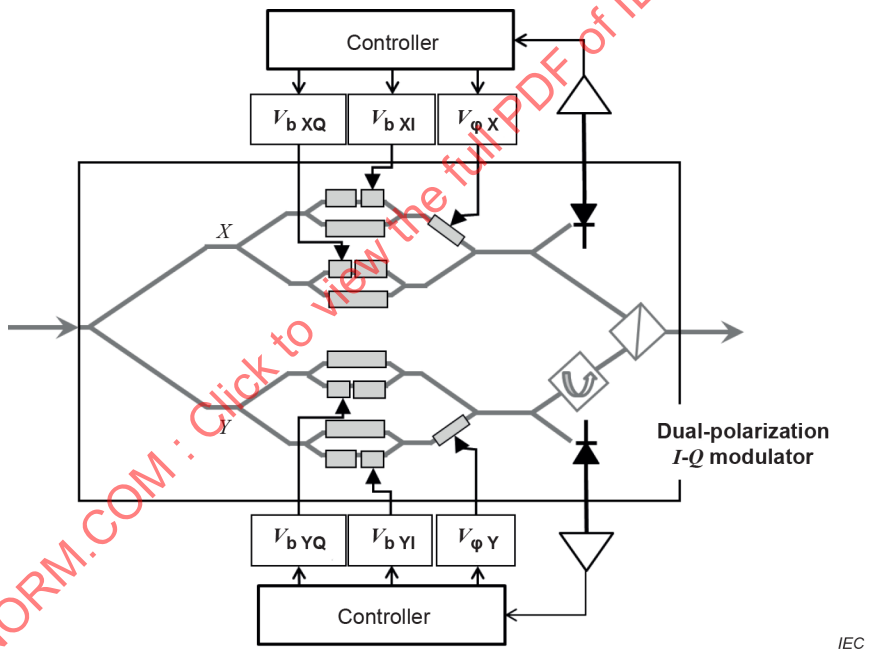
In either case, the output signal carries two pairs of quadrature signals, $A_X(t)$ and $A_Y(t)$, which are transmitted in mutually orthogonal polarization states. Both signals should have the same average optical power at the output of the modulator. This is especially important for optically amplified communication systems, in which optical noise is added to the signals. If one of the two orthogonally polarized signals is weaker than the other one, then the weaker signal component arrives at the receiver with a lower OSNR than the stronger signal component.

The quadruple MZM structure shown in Figure 58(a) and Figure 58(b) requires a total of six bias voltages to be adjusted:

- $V_{b\ XI}$ and $V_{b\ XQ}$ for the setting the biases in the two MZMs generating $A_X(t)$;
- $V_{b\ YI}$ and $V_{b\ YQ}$ for the setting the biases in the other two MZMs generating $A_Y(t)$; and
- $V_{\varphi\ X}$ and $V_{\varphi\ Y}$ for adjusting the *I-Q* phases in $A_X(t)$ and $A_Y(t)$.



(a) With a single external monitor detector



(b) With two internal monitor detectors

Figure 58 – Adaptive bias and phase control in dual-polarization *I-Q* modulator

These voltages can be adaptively adjusted in a similar fashion as discussed in 8.3.3 for the dual-modulator arrangement of Figure 53. Often the six voltages are adjusted in a time-sequential fashion, where

1) $V_{b\,XI}$ is dithered for a certain time period and subsequently re-adjusted; then
2) $V_{b\,XQ}$ is dithered for a certain time period and subsequently re-adjusted; then
3) $V_{\varphi\,X}$ is dithered (or alternatively $V_{b\,XI}$ and $V_{b\,XQ}$ are dithered simultaneously) for a certain time period and $V_{\varphi\,X}$ is subsequently re-adjusted; then
4) $V_{b\,YI}$ is dithered for a certain time period and subsequently re-adjusted; then

5) $V_{b\ YQ}$ is dithered for a certain time period and subsequently re-adjusted; and finally

6) $V_{\varphi\ Y}$ is dithered (or alternatively $V_{b\ YI}$ and $V_{b\ YQ}$ are dithered simultaneously) for a certain time period and $V_{\varphi\ Y}$ is subsequently re-adjusted.

This sequence is then repeated again and again for as long as the modulator is in service, so that these quasi-continuous adjustments track small changes or drifts that can occur in the bias voltages. The speed of these drifts depends on the particular implementation of the modulators but in general is of the order of seconds and, hence, fairly slow.

It is also possible to dither all six voltages simultaneously (or when using method two for phase adjustment, dither only $V_{b\ XI}$, $V_{b\ XQ}$, $V_{b\ YI}$, and $V_{b\ YQ}$ simultaneously) and adjust the six voltages continuously and in parallel. This continuous dither and adjustment has the advantage that the integration times used to calculate the various error signals are maximal, so that the sensitivity of each control loop is maximal, too. However, if all or some of the voltages are dithered simultaneously, one should use sufficiently different dither frequencies for the various bias voltages, so that the effect of a given bias dither in the feedback signal can be unambiguously identified. This is particularly important when only a single monitor detector at the output of the modulator structure is used, as shown in Figure 58(a).

When two independent monitor detectors are used for the X- and Y-polarized signals, as shown in Figure 58(b), the same set of dither frequencies for the X- and Y-related bias voltages can be used. Such internal monitor photodetectors are not uncommon, but they sometimes measure the optical power in a complementary output port of the beam combiner at the output of an I-Q modulator, as indicated in Figure 58(b). These signals can be used for feedback control of the bias voltages if they represent an accurate complement of the output signal of the I-Q modulator. Any significant deviation from this condition could impair the quality of the error signals used for adaptive bias control.

9 Digital signal processing in transmitters for vector-modulated signals

9.1 Pre-distortion of optical signals

9.1.1 General

In coherent optical communication systems, the drive voltages for the modulator(s) are usually generated by high-speed digital-to-analogue converters (DACs), which are integrated with the DSP used for de-convolution of the coherently received signals, as shown schematically in Figure 46. Hence, digital signal processing can be applied to the digital input signals of these DACs, similar to the digital signal processing applied to the received signals discussed in 7.2. In particular, digital signals processing can be used to pre-distort the drive voltages in such a way that they equalize the frequency response of the optical front end of the transmitter, which includes the DACs, the modulator driver amplifier, the optical modulator, and the interconnecting strip lines [89]. The most important transmitter impairments to be pre-distorted are

- magnitude and phase of the transmitter frequency response;
- skew between the XI, XQ, YI, and YQ lanes (i.e. I-Q skews and X-Y skew);
- amplitude imbalance between the XI, XQ, YI, and YQ lanes;
- nonlinear response of the modulator.

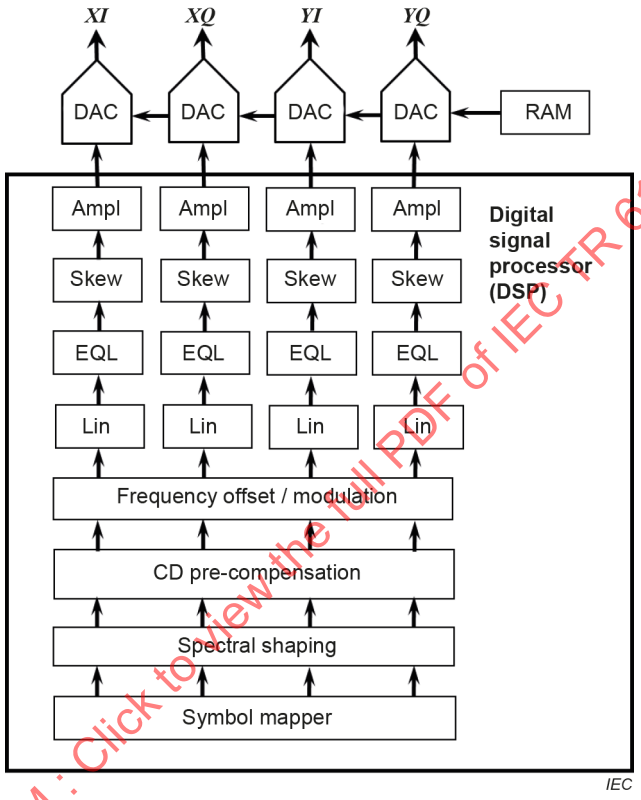
Moreover, digital signal processing is a very effective method to shape the spectrum of the transmitted signal, for example to minimize the bandwidth of the transmitted signal via Nyquist filtering with root-raised-cosine (RRC) spectral shape, as shown in Figure 24.

The various types of signal processing listed above can be performed separately in each of the transmitted lanes. However, when I-Q modulators are employed, it is also possible to pre-compensate chromatic dispersion in the fibre optic transmission link, similar to the CD post-

compensation in the receiver discussed in 6.3.3, or to add a frequency shift to the transmitted signal, similar to the frequency offset removal in the coherent receiver discussed in 6.3.2 [6].

Digital pre-distortion of the signals to be transmitted should be applied in the reverse order of how the impairments are introduced in the transmitter, because some of these operations do not commute (e.g. spectral shaping and nonlinearity compensation).

For the same reason, the digital information signal has to be mapped into the constellation points of the transmitted modulation format prior to applying pre-distortion, as shown schematically in Figure 59.



Key

Ampl	Amplitude adjustment and DC offset removal
DAC	Digital-to-analogue converter
EQL	Equalizer filter for compensation of transmitter frequency response
Lin	Linearizer for compensation of nonlinear modulator response
RAM	Random access data memory
XI	In-phase component of X-polarized signal
XQ	Quadrature component of X-polarized signal
YI	In-phase component of Y-polarized signal
YQ	Quadrature component of Y-polarized signal

Figure 59 – Typical digital signal processing steps in the transmitter

9.1.2 Pre-compensation of linear transmitter impairments

The frequency response of the optical front end of the transmitter can be pre-compensated with the help of fractionally-spaced equalizers (FSEs) as discussed in 6.3.5. Likewise, I-Q skew and X-Y skew can be compensated in the same FSEs or, alternatively, in separate shorter FSEs, as discussed in 6.3.4. The number of taps employed in these FSEs usually is larger than 9. The

more complex the frequency variations are, the more taps are needed. However, the more taps are used, the higher the power consumption of these filters is.

Moreover, spectral shaping of the transmitted signal can be performed by a sufficiently long FSE, as discussed in 6.3.5 [12]. The number of taps required for spectral shaping can range from 19 (for RRC shaping with a roll-off factor of 0,4) to more than 33 (for RRC shaping with a roll-off factor of 0,1; see Figure 24).

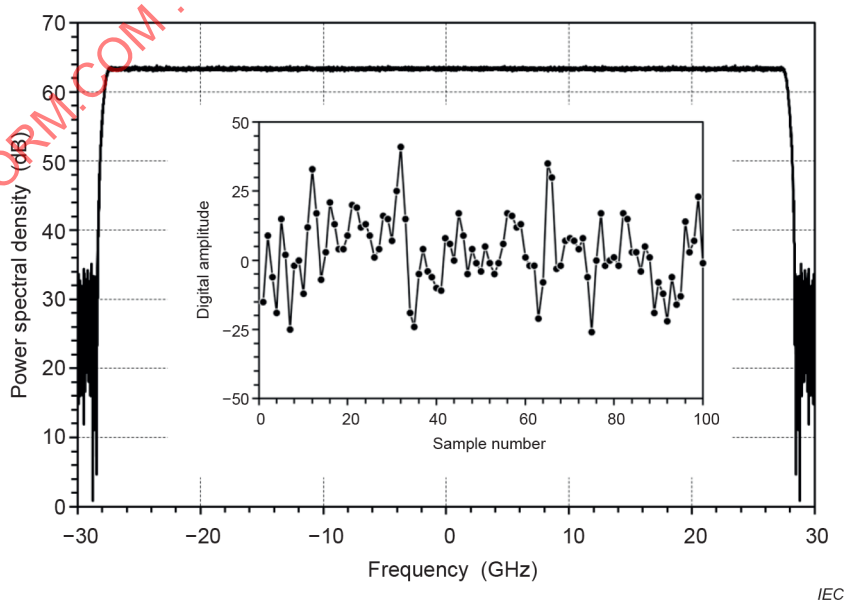
If there is no compensation of nonlinear modulator response (see 9.1.5), then spectral shaping, frequency response equalization, and skew compensation could be combined in just one single FSE, which can lead to a large number of taps to be employed in the FSE. When the number of taps exceeds a certain value, it becomes computationally more efficient to perform the spectral shaping and frequency equalization in the frequency domain, as described in 6.3.3.

If nonlinear modulator response is compensated, then spectral shaping has to be done before the nonlinearity compensation, and the frequency roll-off compensation has to be performed after nonlinearity compensation, as shown in Figure 59.

9.1.3 Determination of transmitter frequency response

Obviously, the frequency response in the transmitter front end has to be known before it can be properly compensated. It turns out that a DSP can facilitate measurements of frequency roll-off and skew. A popular method is to utilize random-access memory (RAM) inside the DSP to transmit predetermined test patterns (so-called test vectors) through the optical front end. RAMs are often attached to the high-speed DACs, as shown in Figure 59, and thus allow the insertion of digital test vectors directly into the DACs. The test vectors are composed of a large number of samples (typically more than 4 000), which can be repeated continuously for the duration of the tests. The samples do not need to represent modulated symbols and, for example, can be arbitrary sequences of random noise.

It is thus possible to transmit a test vector, whose digital spectrum is completely flat (i.e. “white”), as shown in the example of Figure 60. When this white signal is sent through the transmitter front end, it will be reshaped according to the frequency response of the entire front end (including DAC and optical modulator). The magnitude of the transmitter frequency response can then be directly seen in the optical output spectrum of the modulator (measured for example with a high-resolution OSA), as shown in the example of Figure 61.



NOTE The main graph displays the spectrum of the test signal, whereas the smaller graph, which is inserted in the centre of the main graph, displays the first 100 samples of the test signal.

Figure 60 – Spectrum and samples of a white test vector

It is often possible to transmit such a test vector in only one of the four lanes and leave the other three lanes unmodulated. If white test vectors are to be transmitted simultaneously in all four lanes, the samples in the vectors should be statistically independent, so as not to interfere with the I-Q phase control (see 8.3.3). In Figure 61, only one test vector was transmitted in the XI lane. The slight asymmetry in this spectrum can be attributed to electrical crosstalk between the I- and Q-signals.

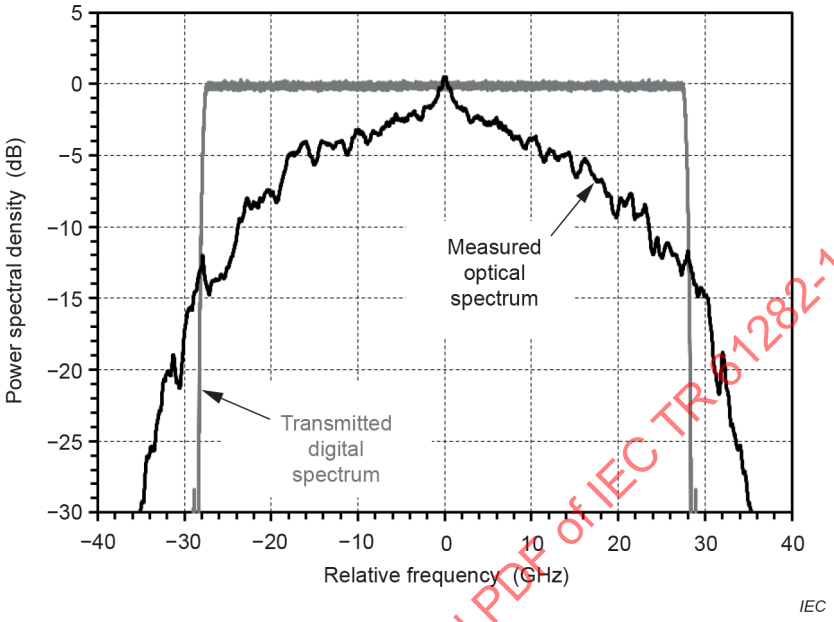


Figure 61 – Magnitude of transmitter frequency response

The frequency response of Figure 61 can be used to calculate the tap coefficients for the pre-compensation filter in the DSP. However, it is often also important to pre-compensate for the frequency-dependent phase delays in the optical front end. These phase delays cannot be easily measured with an OSA, because such an instrument typically is designed to measure the magnitude of the various frequency components but not their relative phases.

However, the relative phase delays can be measured with the help of a well-calibrated receiver, for example with an optical modulation analyser (OMA). This measurement requires transmission of a known test vector, as described above, and recording of the received signal in the calibrated receiver.

NOTE The DSPs used in coherent optical communication systems often have random access memory (RAM) connected to the high-speed ADCs at the receiver side, as shown in Figure 43, which can be used to record the received ADC samples.

If the received signal is recorded over the same time period as the length of the test-vector and sampled at the same rate, then the transfer function of the transmitter can be determined with the following procedure.

1) Calculate the complex spectrum of the transmitted signal via Fourier transformation.
2) Calculate the complex spectrum of the received signal via Fourier transformation.
3) Correct the received spectrum with the known transfer function of the receiver (if needed).
4) Calculate the spectral transfer function of the transmitter from the ratio of the corrected received spectrum and the transmitted spectrum (received spectrum divided by transmitted spectrum).
5) Remove any time delay between the transmitted and received signals from the spectrum.

The frequency dependence of the phase delays in the transmitted signal can then be directly calculated from the complex elements of the above-described transfer function. An example of