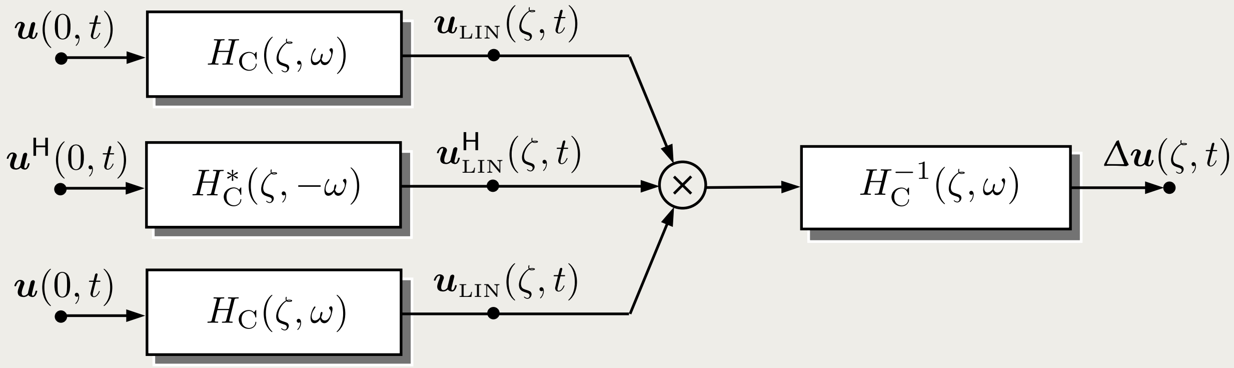
ulm university universität ulm

Felix Frey

A System Theoretical View on Nonlinear Fiber Propagation



Band 4
Schriftenreihe des Instituts für Nachrichtentechnik

Felix Frey

A System Theoretical View on Nonlinear Fiber Propagation

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.dnb.de abrufbar.

Dissertation, Universität Ulm,
Fakultät für Ingenieurwissenschaften, Informatik und Psychologie, 2022

Impressum

Universität Ulm
Institut für Nachrichtentechnik
Prof. Dr.-Ing. Robert Fischer
Albert-Einstein-Allee 43
89081 Ulm

Eine Übersicht über alle Bände der Schriftenreihe finden Sie unter
https://www.uni-ulm.de/ntschriften

Institut für Nachrichtentechnik der Universität Ulm, 2023

ISBN 4978-3-948303-12-9

ulm university universität uulm

A System Theoretical View on Nonlinear Fiber Propagation

DISSERTATION

zur Erlangung des akademischen Grades eines

DOKTOR-INGENIEURS

(Dr.-Ing.)

der Fakultät für Ingenieurwissenschaften, Informatik
und Psychologie der Universität Ulm

von

Felix Frey

aus Heidelberg

Gutachter: Prof. Dr.-Ing. Robert Fischer
Prof. Dr.-Ing. Laurent Schmalen

Amtierende Dekanin: Prof. Dr. Anke Huckauf

Ulm, 21. Juli 2022

Acknowledgment

This work was supported by the Deutsche Forschungsgemeinschaft (German Research Foundation) under grant FI 982/14-1.

Abstract

In communication theory, discrete-time end-to-end channel models play a fundamental role in developing advanced transmission and equalization schemes. Most notable, the discrete-time *linear*, dispersive channel with additive white Gaussian noise (AWGN) is often used to model *point-to-point* transmission scenarios. In the last decades, numerous transmission methods for such linear channels have emerged and are now applied in many digital transmission standards. With the advent of high-speed CMOS technology, those schemes have also been adopted in applications for *fiber-optical* transmission with digital-coherent reception.

Many of the applied techniques (e.g., coded modulation, signal shaping, and equalization) are still designed for linear channels whereas the fiber-optical channel is inherently *nonlinear*. A channel model which obtains the (discrete-time) output symbol sequence from a given (discrete-time) input symbol sequence by an *explicit* input/output relation is highly desirable to make further advances in developing strategies optimized for fiber-optical transmission.

In the past two decades, considerable effort was spent developing channel models for fiber-optical transmission with good trade-offs between computational complexity and numerical accuracy. Most of the early work is, however, concerned with the phenomenology in the *optical* domain alone, i.e., both the source and the effect of fiber nonlinearity is studied in the *continuous-time*, optical domain not considering the transmitter and/or receiver front-ends. Here, one promising strategy is, e.g., based on the so-called *perturbative* approach where nonlinear effects are considered as small perturbations to the optical signal. By now, the optical community is faced with a vast number of models based on the perturbation premise, each model with its own assumptions, simplifications, and objectives. The connection between already existing models using complementary views (e.g., one in time-, the other in frequency-domain) is often unclear. A detailed and rigorous derivation for some prominent models is still pending. E.g., the transition from the original continuous-time to the more relevant discrete-time end-to-end model lacks a comprehensive, system-theoretic analysis.

Hence, in this dissertation, nonlinear fiber propagation is assessed from a *systems-theoretic* point of view with applications for communication systems. Based on the theory of nonlinear systems, the present work aims to connect the dots between various, existing channel models, unifying and comparing the different approaches. To that end, the perturbation approach in continuous-time is revisited with a special emphasis on the dual representations of nonlinear systems in time and frequency. From that, a discrete-time end-to-end fiber-optical channel model is derived which includes the transmit-side pulse shaping, the receive-side matched filtering, and T-spaced sampling. As before, two complementary representations of the now time-discretized end-to-end model are present—one in (discrete) time domain, the other in $1/T$-periodic continuous-frequency domain. The time-domain formulation coincides with the well-known *pulse-collision picture*. The novel frequency-domain picture incorporates the sampling operation via an aliased and hence $1/T$-periodic formulation of the nonlinear system. This gives rise to an alternative perspective on the end-to-end input/output relation between the spectrum of the discrete-time transmit symbol sequence and the spectrum of the receive symbol sequence. Both views can be extended from a regular, i.e., solely additive model, to a combined regular-logarithmic model to take the multiplicative nature of certain distortions into consideration. A novel algorithmic implementation of the discrete and periodic frequency-domain model is presented. The derived end-to-end model requires only a single computational step and shows good agreement in the mean-squared error sense compared to the oversampled and inherently sequential split-step Fourier method.

Contents

Abstract **vii**

1 Introduction and Outline **1**
1.1 State of the Art . . . 2
1.2 The Present Work . . . 10
1.3 Outline of the Dissertation . . . 12

2 Foundations and Basic Concepts **15**
2.1 Notation and Basic Definitions . . . 16
2.1.1 Discrete- and Continuous-Time Signals . . . 18
2.1.2 Integral Transforms . . . 20
2.1.3 Linear- and Nonlinear Time-Invariant Systems . . . 24
2.2 The Linear System Model . . . 35
2.2.1 Transmitter Front-end . . . 35
2.2.2 2×2 MIMO AWGN Channel . . . 40
2.2.3 Receiver Front-end . . . 42

3 Principles of Fiber-Optic Transmission **47**
3.1 The Optical Field Envelope . . . 47
3.2 Wavelength-Division Multiplexing . . . 53
3.3 The Fundamental Evolution Equation . . . 56
3.3.1 Power Profile and Amplification Noise . . . 56
3.3.2 Chromatic Dispersion Profile . . . 61
3.3.3 Linear Channel Transfer Function . . . 70
3.3.4 Optimum Receive Filter for the Linear Fiber Channel . . . 70
3.3.5 Fiber Nonlinearity . . . 71
3.3.6 The Sampled Manakov Equation . . . 74

4 The Continuous-Time Perturbation Approach **77**
4.1 The First-Order Regular Solution in Frequency Domain . . . 79
4.2 The First-Order Regular Solution in Time Domain . . . 95
4.3 The (Analog) Baseband End-to-End Channel . . . 101

5 The Discrete-Time Perturbation Approach **107**
5.1 The First-Order Regular Solution . . . 108
5.1.1 Kernel Energy and Relation to the GN-Model . . . 122
5.1.2 Discrete-Frequency Algorithmic Implementation . . . 123

5.2 The First-Order Regular-Logarithmic Solution 125
5.2.1 Regular-Logarithmic Model in Time Domain 126
5.2.2 Regular-Logarithmic Model in Frequency Domain 130
5.3 Numerical Results 133

6 Concluding Remarks **141**

A Appendix **143**
A.1 The Jones and Stokes Formalism 143
A.2 Proof of the Fourier Relation in (4.62), (4.63) 156
A.3 Alternative Derivation of the (Analog) Baseband End-to-End RP Method . . . 158
A.4 Alternative Derivation of the Discrete-Time End-to-End RP Method 162

B Notation **167**
B.1 Abbreviations 167
B.2 Mathematical Operators 171
B.3 Mathematical Symbols 172

Bibliography **177**

Publications, Preprints, and Presentations of the Author **187**

1. Introduction and Outline

Modeling nonlinear fiber propagation spans over a variety of disciplines, ranging from applied math, over physics, to nonlinear system theory, and (optical) communication theory. As a result, large portions of the existing theory on this topic is scattered over numerous, often disconnected publications. There are only a few attempts that try to string together the different views into a single, cohesive picture. Instead, the connection between already existing models is often left unclear.

One of the main objectives of this thesis is to establish a unifying picture of the numerous models for fiber-optical transmission. What made this task particularly difficult is the lack of a common notation in use—people dealing with fiber nonlinearities have a strong background in physics, and their notation is often different to the notation used in the field of communication theory. Additionally, within the community that is concerned with modeling nonlinear fiber transmission, the notation may also vary significantly over different groups and over time, considering that the field is more than 30 years old. We decided to adopt the notation commonly practiced at my alma mater. This made it necessary to (re-) write large sections of the basic theory on fiber-optic transmission using the notational framework from [Hub92, Fis02]. We tried to keep the number of relevant equations at a minimum, while making this thesis as self-contained and complete as possible.

After reading this text, the reader will have a good understanding on perturbative models for nonlinear fiber transmission. The reader will establish a notion of the description for such nonlinear systems in both time and frequency domain, and apply this to both intra-channel and inter-channel effects. A particular focus is put on modeling the end-to-end relation between discrete-time transmit and receiver symbols. For some exemplary system scenarios, different implementations of the derived models will be compared to the split-step Fourier method. Despite the many assumptions and simplifications to arrive at the (relatively) simple models, the obtained results are remarkably accurate. The greater benefit of this work lies in the potential application in a variety of related fields such as model learning, performance monitoring, or nonlinear compensation.

1.1 State of the Art

In this section, we aim to give an overview on existing channel models for fiber-optical transmission systems. The diagram in Figure 1.1 gives a pictorial representation of the different classes of channel models and how they are related. Each block represents a certain model class; it contains a reference to a selected publication from the literature that represents that class. In anticipation of the following chapters, a set of central equations is already given which the reader may revisit when studying the respective sections[1]. We separate between the analytical approach, aka. *ansatz*, to a model class (top pane) and its numerical realization, i.e., the algorithmic implementation of the former (bottom pane).

Analytical Channel Model Within the top pane, we have the following situation: the analytic approach is either formulated in continuous or in discrete time, depending on the *domain* of the communication signal. Due to physical properties of the communication signal, the continuous-time domain is evidently relevant in the *optical* domain (e.g., considering the waveform of the electro-magnetic field of light), but also in the *analog/electrical* domain (e.g., considering the electric voltage/current at the analog transmitter and receiver components). In contrast, the discrete-time approach is present when a symbol sequence is transmitted with a rate of $R_\mathrm{s} \stackrel{\text{def}}{=} 1/T$ (where T is the duration of the modulation interval) and received over an (optical) communication channel. We then speak of a T-spaced discrete-time, end-to-end channel model. Discrete-time channel models play a fundamental role in the field of information theory, e.g., Shannon's noisy-channel coding theorem [Sha48] only applies to discrete-time channels.

The above continuous-time and discrete-time approaches differ in a fundamental aspect, as detailed in the following. Since all communication signals are bandlimited when transmitted over a *linear* channel, analog communication signals can be processed in discrete-time domain without loss of information when obeying the Nyquist-Shannon sampling theorem. Since the bandwidth of a single communication signal is typically larger than the symbol rate R_s, the sampling frequency has to be larger than $1/T$, so-called oversampling. A continuous-time, linear channel and its oversampled, discrete-time surrogate become equivalent from an information-theoretic point of view when obeying the sampling theorem. However, in any digital receiver for T-spaced pulse-amplitude modulation, T-spaced sampling and further T-spaced discrete-time signal processing are performed. Thus, for communication signals the sampling theorem is in general *not* fulfilled. In fact, aliasing of frequency components is an essential part in recovering the data. Hence, all discrete-time end-to-end channel models have to incorporate this sampling step in order to fully capture the effects.

When a communication signal is transmitted over a *nonlinear* channel the situation depends on the class of nonlinear system representing that channel. Some of the above considerations from linear channels can be generalized to the class of nonlinear systems which are time-invariant and have no (temporal) feedback (aka. non-recursive or open-loop systems). Those systems, if stable, can be mathematically expressed by a *Volterra series* [Sch80], reviewed

[1]We will also already make some (limited) use of the notational framework which will only be formally introduced in the course of this work.

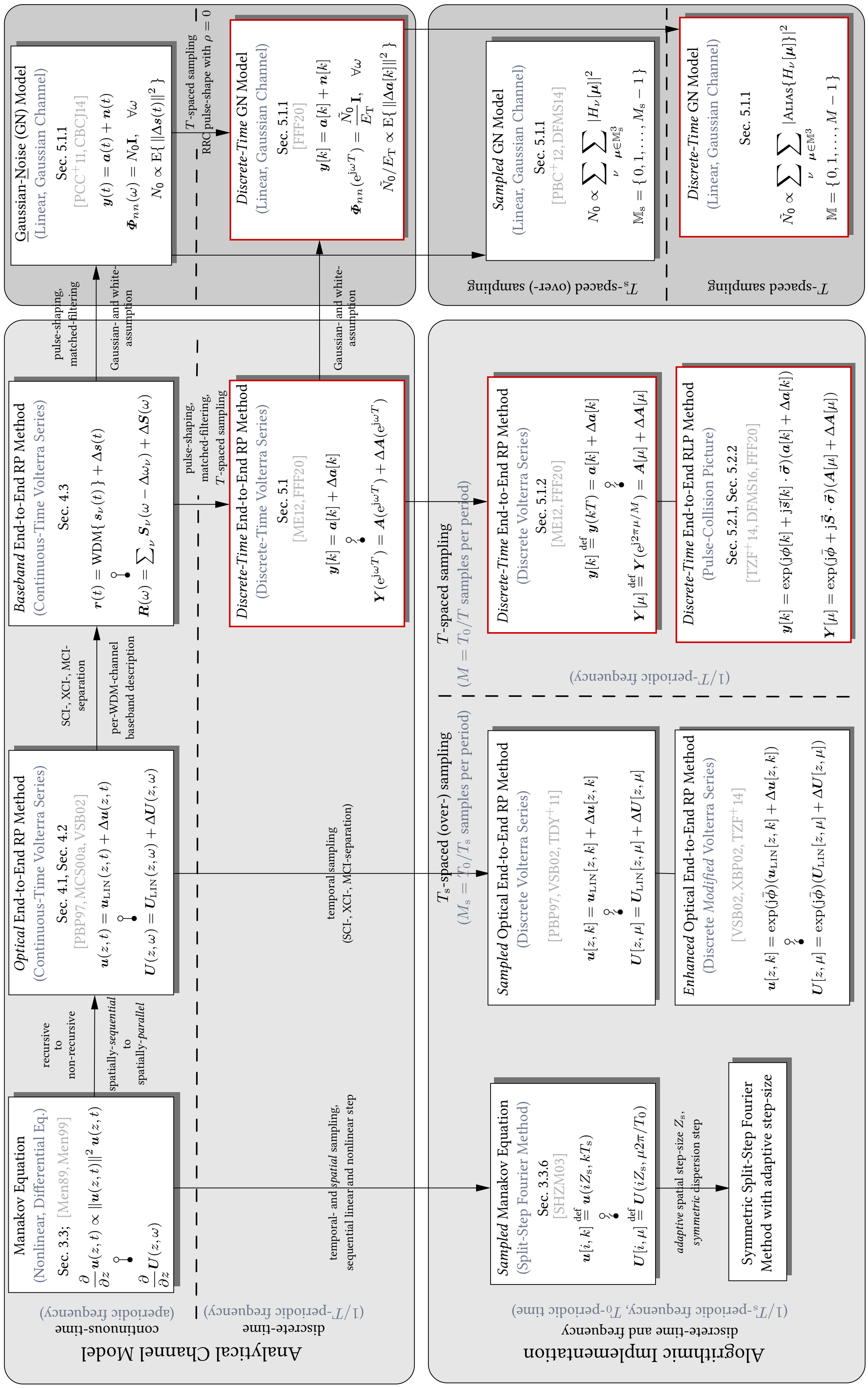


Figure 1.1: Overview and classification of fiber-optical channel models and their algorithmic implementation. See, e.g., Figure 3.2 or 4.1 for variable definitions.

in Section 2.1.3.2. In general, however, a continuous-time, nonlinear channel and its sampled derivative are no longer equivalent from an information-theoretic perspective.

Algorithmic Implementation For the reasons described in the preceding paragraph, we will draw a distinction between a continuous-time (nonlinear) channel and its discrete-time counterpart. The sampled representation of a (general) nonlinear channel is often just an approximation of the original model. Heuristic modifications can be implemented to improve the match between the sampled representation and the original model. Therefore, in the bottom pane of Figure 1.1, we collect common *numerical realizations* of the above analytical channel models. Here, discretization is typically performed not only in time domain but also in the transform domain (i.e., frequency domain for most practical cases) to allow for efficient algorithmic implementations.

Manakov Equation In fiber-optical transmission the nonlinear channel is described by two coupled nonlinear Schrödinger equations—the so-termed *Manakov equation* [Men89], see Figure 1.1 (top pane, left). The Manakov equation is a nonlinear, partial differential equation which is nonlinear in the (continuous-time) unknown function[2] $\boldsymbol{u}(z,t) \in \mathbb{C}^2$, the communication signal[3]. It represents the optical field envelope of the electromagnetic field in two polarizations at position z and time t, see Section 3.1. The rate of change of the signal $\boldsymbol{u}(z,t)$ in the spatial direction z is proportional to the nonlinear term $\|\boldsymbol{u}(z,t)\|^2\,\boldsymbol{u}(z,t)$. From a systems perspective, this may be visualized as a *closed-loop* system (over the spatial domain), i.e., one with the nonlinear expression in the feedback path. In this case, the communication signal $\boldsymbol{u}(z,t)$ may no longer be strictly bandlimited in a Nyquist-Shannon sense. This is because nonlinear transmission may generate nonlinear mixing products outside the communication signal's initial bandwidth, so-called *spectral broadening* [Agr06, Ch. 4]. Those mixing products may again interact nonlinearly to broaden the spectrum even further in a recursive manner. For that reason, a universal definition of *bandwidth* does not exist for the continuous-time, nonlinear fiber channel [Agr17]. Yet, the Nyquist-Shannon sampling theorem can only be applied to strictly bandlimited signals. This poses one of the central problems in determining the channel capacity of the fiber channel.

Sampled Manakov Equation When evaluating and implementing the Manakov equation, e.g., for numerical simulations, high oversampling is required to account for the effect of spectral broadening. However, compared to linear channels, the sampled, discrete-time surrogate channel (i.e., the *sampled* Manakov equation) is in general not equivalent to the continuous-time, nonlinear channel—independent of the oversampling ratio T/T_s (where T_s is the duration of the sampling period). Some of the information is lost, and, hence, the continuous- and discrete-time model are no longer equivalent. A practical solution is to set the oversampling to a value much larger than required for, e.g., linear channels, to keep the numerical error small with respect to the continuous-time channel.

[2]In calculus, the *unknown function* $\boldsymbol{u}$ is called the *dependent variable* and the variables it depends of, i.e., space z and time t, are called *independent* variables.

[3]For simplicity, we only consider the noiseless part of the communication signal and also neglect any wideband noise in the evolution equation.

The de-facto standard method for simulating (i.e., approximating) continuous-time fiber-optical transmission is the split-step Fourier method (SSFM) [SHZM03], see Figure 1.1 (bottom pane, left). The SSFM performs sampling of both the temporal domain t at integer multiples k of the sampling period T_s, and the spatial domain z at integer multiples i of the step-size Z_s. Additionally, the discretized optical field envelope $\boldsymbol{u}[i,k] \overset{\text{def}}{=} \boldsymbol{u}(iZ_\mathrm{s}, kT_\mathrm{s})$ is processed sequentially in z when evaluating the spatial recurrence relation (i.e., *difference equation*) of the sampled Manakov equation, iterating between a linear and a nonlinear step[4]. The linear step is represented by a linear, discrete-time system which is responsible for signal scaling (to account for gain/loss variations) and for signal dispersion (to account for the system memory induced mainly by chromatic dispersion). The nonlinear step is represented by a memoryless discrete-time system which causes a rotation of the signal phase proportional to $\|\boldsymbol{u}[i,k]\|^2$. The details of SSFM are reviewed in Section 3.3.6. In addition to time-discretization, discretization in frequency $\omega = 2\pi f$ at integer multiples μ of the spectral resolution $2\pi/T_0$ is performed, as it enables the use of efficient filter operations in the linear step. Discretization in both domains results in periodic sequences in both time and frequency domain. For instance, the fundamental period T_0 of the time-periodic signal must be chosen sufficiently large to reduce the effect of cyclic artifacts when the system has large memory, i.e., the number of simulated samples within the fundamental period $M_\mathrm{s} = T_0/T_\mathrm{s}$ must be large compared to the (discrete-time) channel memory. Vice-versa, the sampling frequency $1/T_\mathrm{s}$ is typically chosen (at least) *three* times the Nyquist frequency of the optical communication signal. The reason for this choice will become clear when considering the Volterra series to approximate the Manakov equation, see below.

To improve the accuracy of the SSFM, a variation of the implementation is the *symmetrized* SSFM with *adaptive step-size* [SHZM03]. Here, the nonlinear step is sandwiched between two linear steps, each accounting for half the differential system memory. Additionally, the spatial domain is not sampled uniformly, but instead, the step-size $Z_\mathrm{s}(z)$ is a function of the position z and adapted according to the (in general non-uniform) power profile to reduce the relative error or to reduce the required number of spatial steps without a decrease in accuracy.

Optical End-to-End RP Method Starting again from the Manakov equation, an approximate solution to the continuous-time *optical* end-to-end channel (top pane in Figure 1.1) can be obtained following the *perturbation theory* [Zwi98, VSB02].

In short, perturbation theory performs a power series expansion of an unknown function in the parameter ϵ with small value $\epsilon \ll 1$. The first term, i.e., the zero$^{\text{th}}$-order solution, must be a solvable system, while successive terms (i.e., the *perturbative* terms) in the series expansion typically become smaller with increasing order. If the power series converges asymptotically to the exact solution, the perturbation problem is called a regular perturbation (RP) problem. Applied to the Manakov equation, the unknown function is again the optical field envelope $\boldsymbol{u}(z,t)$, the small parameter ϵ is a proportionality coefficient to the nonlinear term (later introduced as the *nonlinearity coefficient*), and the zero$^{\text{th}}$-order solution is $\boldsymbol{u}_{\text{LIN}}(z,t)$ which solves

[4] The spatial, recursive recurrence relation is "unrolled" over z, and the linear and nonlinear step together form the joint differential step between two discrete positions i and $i+1$.

the corresponding end-to-end system comprising only *linear* effects. The main approximation using the *first-order* RP method is that the optical field envelope $\boldsymbol{u}(z,t)$, subject to nonlinear distortions, can be understood as the sum of the *linearly* propagating signal $\boldsymbol{u}_{\mathrm{LIN}}(z,t)$ and an additive perturbation signal $\Delta\boldsymbol{u}(z,t)$ representing the *nonlinear* distortions accumulated from the input of the transmission link up to position z.

The solution to the linear signal $\boldsymbol{u}_{\mathrm{LIN}}(z,t)$ is straightforward. The solution to the perturbation signal $\Delta\boldsymbol{u}(z,t)$ is more involved. Mathematically, it can also be described under the equivalent framework of *Volterra series*—either in time domain [MCS00b] or in frequency domain [PBP97]. In the first-order RP approximation, the only *source* of a nonlinear perturbation is the linear signal $\boldsymbol{u}_{\mathrm{LIN}}(z,t)$. Second and higher-order solutions can be obtained in a recursive manner by taking the first and higher-order solutions as source term. Due to the increasing numerical complexity, typically, only the first-order solution is considered. Then, the first-order solution, i.e., the perturbation signal $\Delta\boldsymbol{u}(z,t)$ itself, does not interact again in a nonlinear fashion to generate second- or higher-order nonlinear distortions—no nonlinear recursion is present. This is in contrast to the Manakov equation and the SSFM derived thereof. While the Manakov equation is inherently recursive in z, sequential processing is adequate, e.g., when implementing the SSFM. Instead, when considering a truncated RP solution, this constraint is relaxed, and the sequential view turns into a *parallel* view over the spatial domain, where each local perturbation is generated independently of other local perturbations (see, e.g., the *parallel fiber model* in [VSB02, Sec. IV] or Figure 4.2). Similar to the Manakov equation, the optical end-to-end RP method does not impose any constraints on the transmitter or receiver front-ends (including, e.g., pulse-shape or matched filter) by acting directly on the optical field envelope.

The RP method in frequency domain can be naturally derived by expressing the Manakov equation in frequency domain. It had its first appearance in the pioneering work of [PBP97] under the framework of a Volterra series. In [VSB02], the equivalence between the frequency-domain n^{th}-order RP solution and the (truncated) frequency-domain $(2n+1)^{\mathrm{th}}$-order Volterra series, aka. Volterra series transfer function (VSTF), was shown. The equivalence is also true for the time-domain solution of the RP method and the time-domain Volterra series. For that reason, we will use the terms *RP method* and *Volterra series* synonymously in this text. Independently of the work above mostly concerned with the approach in frequency domain, the time-domain analysis was first introduced in a series of publications in the early 2000s by A. Mecozzi *et al.* [MCS00a, MCS00b, MCS$^+$01]. The method was applied to transmission schemes that were practical at that time (e.g., dispersion-managed transmission, intensity-modulation, direct-detection, and Gaussian pulse-shapes). The details of the theory and its derivation were published more recently in [Mec11].

The terminology used in the optical communications community today, see, e.g., [CGK$^+$17], still deceptively implies that models based on the RP method and on Volterra theory relate to two different solutions. Therein, the RP method often refers to a time-domain approach, and VSTF refers to a frequency-domain approach. The connection between the time- and frequency-domain view was already pointed out in the literature, e.g., in [Wei06, BSO08] using the theory from *dispersion-managed* nonlinear Schrödinger equation (NLSE) [GT96, AB98, AH02a, AH02b] originally designed for soliton transmission.

Sampled RP Method In order to obtain numerical results from the derived RP methods, a *sampled* representation of the continuous-time optical end-to-end RP solution (i.e., a *discrete-time* Volterra series, see bottom pane in Figure 1.1) is required. In contrast to the Manakov equation, see discussion above, the truncated, continuous-time RP model can be sampled without loss of information. This is true under the assumption that the input signal is strictly bandlimited and that the Volterra system has finite order and finite memory [BC85]. Then, the Nyquist sampling theorem can be applied to the *output* of the Volterra system. For our application, the spectral support of the system's output depends on the spectral width of the (bandlimited) linear communication signal $\boldsymbol{u}_{\text{LIN}}(z,t)$ (i.e., including all wavelength channels) and on the *order* of the RP method. E.g., the first-order RP method is equivalent to a third-order nonlinear system (i.e., mathematically expressed by a third-order Volterra operator) and hence requires a sampling rate of (at least) *three* times the Nyquist frequency of $\boldsymbol{u}_{\text{LIN}}(z,t)$ [PBP97]. This is due to the fact that nonlinear mixing products originating from the linear signal spectrum must fall within three times the Nyquist region given a third-order nonlinear system. Since the system is not recursive over z, no (out-of-band) higher-order mixing products are generated.

In later years, the methodology of the above optical end-to-end models was applied in the context of fiber nonlinearity compensation. Here, methods for fiber nonlinearity compensation (NLC) in (discrete) frequency domain can be directly derived from the frequency-domain RP formulation. Commonly, such a functional block is embedded into the receiver-side digital signal processing (DSP) just after the analog-to-digital conversion and *before* matched filtering [LLH+12, GP13, BCRC16], i.e., still in the oversampled domain of the receiver DSP. A similar approach was also applied for fiber NLC using a *time*-domain variant of the RP method implemented at the transmitter side as a pre-distortion algorithm [TDY+11] or at the receiver side prior to linear equalization [GAMP15].

Already in the early 2000s, the authors of both [XBP02] and [VSB02] published an *enhanced* version of the RP method (aka. *eRP* or *modified* VSTF [XBP01, GRTP11], see bottom pane in Figure 1.1) where the time-invariant (i.e., average) nonlinear phase rotation is explicitly carried out as a unitary rotation, see also [SB15]. This modification greatly improved the accuracy of the numerical results using a frequency-domain, first-order RP solution. For the same reason, the authors of [FDT+12, TZF+14] proposed a so-termed *additive-multiplicative* (A-M) model where a subset of time-domain perturbations (also corresponding to the average nonlinear phase rotation) were implemented as a multiplicative distortion. In [FS05, SF12, SFP13], a logarithmic perturbation (LP) model was derived which is exact in the limit of zero-dispersion links. On the other hand in dispersion-uncompensated (DU) links (which is the default in current optical networks), as pointed out in [SB13], the LP method yields a log-normal distribution of the nonlinear distortion which is inconsistent with observations from simulations and experiments.

Baseband End-to-End RP Method Going back to the original continuous-time, optical end-to-end RP method in Figure 1.1 (top pane), we introduce an intermediate model—the *baseband* formulation of the former—in the derivation of the *discrete*-time end-to-end RP method. The reason for this is as follows: the optical transmit signal at the input of the transmission system is constituted by a number of wavelength channels with channel index $\nu = 1, 2, \ldots, N_{\text{ch}}$ by means of wavelength-division multiplexing (WDM), see Section 3.2. Using a baseband formulation of the RP method, the end-to-end relation is now established between the per-WDM-channel baseband transmit signals $\boldsymbol{s}_\nu(t)$ and the joint receive signal $\boldsymbol{r}(t)$ before channel selection and matched filtering. This allows to separate the received baseband perturbation signal $\Delta\boldsymbol{s}(t)$ in the spectral support of a probe channel (i.e., a channel under test) into contributions originating from self-channel interference (SCI), cross-channel interference (XCI), and multi-channel interference (MCI). Here, SCI relates to the portion of the perturbation $\Delta\boldsymbol{s}(t)$ which the probe channel inflicts on itself, whereas XCI and MCI relate to the portion of $\Delta\boldsymbol{s}(t)$ that depends on the nonlinear interaction between the probe channel and one or more other wavelength channels with channel separation $\Delta\omega_\nu$.

This baseband view and, in particular, the separation of the perturbation signal according to its origin is also, in part, already considered in some of the published works associated to the optical end-to-end RP method.

Discrete-Time End-to-End RP Method In the seminal paper by A. Mecozzi and R.-J. Essiambre [ME12] the former work on continuous-time RP methods was extended by including the transmit pulse-shape, the matched filter, and T-spaced sampling after ideal coherent detection. This work constitutes the first true T-spaced *discrete-time* end-to-end formulation using a time-domain approach (see Figure 1.1, top pane). It provides a first-order approximation of the *per-modulation-interval* equivalent end-to-end input/output relation assuming a multi wavelength channel, i.e., a WDM, scenario. One central result is the integral formulation of the (Volterra) *kernel*[5] in time domain. The complementing view in periodic frequency domain which takes the aliasing of frequency components properly into account was unknown at the time and is part of the present work [FFF20], see discussion in Section 1.2.

The transition to the discrete-time algorithmic implementation (see Figure 1.1, bottom pane) is straightforward considering the time-domain view. Discretization in frequency domain requires proper processing using the overlap-save method to address cyclic effects in time domain, also part of the present work.

Based on [ME12], R. Dar *et al.* [DFMS13, DFMS14, DFMS16, DFM$^+$15] derived the so-called *pulse-collision picture* of the nonlinear fiber-optical channel. Here, the properties of XCI were properly associated with certain types of so-called *pulse collisions* in time domain. In particular, the importance of separating additive and multiplicative distortions, similar to the *enhanced* RP method or the A-M model, was discussed. A similar approach to the discrete frequency-domain model is part of the present work [FFF20].

The work from [ME12] and the follow-up work on the pulse-collision picture spawned a renewed interest in perturbation-based fiber NLC. In contrast to prior implementations (also perturbation-based, but in frequency domain under the label of VSTF, see, e.g., [LLH$^+$12]),

[5]The term *kernel* is conceptually used as an extension of an *impulse response* or *transfer function* to nonlinear systems described by a Volterra series.

the present approach allows carrying out compensation algorithms on the transmit or receive *symbol sequence* rather than on an oversampled representation thereof. For this reason, the functional NLC block can now be placed at the transmitter side *before* pulse-shaping or at the receiver side *after* matched-filtering in the T-spaced domain of the DSP. The derived NLC methods are based on variants of the *time*-domain RP solution, see also [MCS00a] and [TDY$^+$11], and have been demonstrated by various groups in slightly different flavors, cf. [GCKY13, ONO$^+$14, GCK$^+$14, ZRB$^+$14]. All of them consider, however, only a subset of time-domain kernel coefficients[6] due to an assumption sometimes termed *temporal matching* constraint, see [ME12, Eq. (84)]. In [FES$^+$18], we present an improved method taking into account the full set of time-domain kernel coefficients.

Gaussian-Noise Model Another major class of channel models is known as Gaussian-noise (GN) model [SKP93,PCC$^+$11] which belongs to the space of *stochastic* models (see Figure 1.1 top pane, right). GN-models are widely applied as network planning tools, e.g., for provisioning of resources in a loaded network. For that, GN-models estimate the link-delivered signal quality of a specific probe channel in a point-to-point transmission which is determined by the joint impact of noise (e.g., from optical amplification) and nonlinear distortions. For simplicity, both effects are combined into a single metric, e.g., combined into a received signal-to-interference ratio including noise and distortions[7]. Estimation of the nonlinear distortion's power (or variance) is thus at the core of all GN-models. The underlying theory is based on the first-order RP solution of the continuous-time baseband end-to-end channel.

In terms of a fiber-optical channel model, the GN-model is essentially just a *linear*, memoryless channel[8] with an additive random variable (RV) $\boldsymbol{n}(t)$ and a probability density function (PDF) associated to it. The random variable $\boldsymbol{n}(t)$ represents the combination of additive white Gaussian noise (AWGN) and deterministic nonlinear distortions. For that purpose, the main simplification of the GN-model is that nonlinear distortions have, similar to the AWGN assumption, a Gaussian distribution and can be considered white. Then, the power spectral density (PSD) $\boldsymbol{\Phi}_{nn}(\omega)$ associated with $\boldsymbol{n}(t)$ is constant over all frequencies ω with equal value N_0 per equivalent complex baseband (ECB) domain. The magnitude of N_0 is by assumption proportional to the sum of the link-delivered noise power and the power of the perturbation signal $\Delta\boldsymbol{s}(t)$. The objective of GN-models is hence to (efficiently) evaluate the expectation $\mathrm{E}\{\, \|\Delta\boldsymbol{s}(t)\|^2 \,\}$ of the perturbation signal.

Notably, the GN-model is derived on the basis of the *continuous-time* RP solution—not taking T-spaced sampling into account. This is, however, achieved by what we refer to as *discrete-time* GN-model which is part of the present work, see discussion in the following Section 1.2.

Sampled GN-Model Similar as for the continuous-time RP methods, the underlying channel model is formulated in continuous-time (and frequency), while the algorithmic implementa-

[6] This subset of kernel coefficients is often labeled by the variable $C_{m,n}$ in the literature cited above.

[7] In the community, this metric is often termed *nonlinear* optical signal-to-noise ratio (OSNR) [PJ17, Sec. II A] to imply that nonlinear distortions are included in the ratio.

[8] Both the (noiseless) Manakov equation and the derived RP methods provide a *deterministic* relation between the transmit and receive communication signal (aka. *waveform channel model*).

tion is performed on a *sampled* representation thereof, see, e.g., [PBC+12].

The expectation $\mathrm{E}\{\, \|\Delta\boldsymbol{s}(t)\|^2 \,\}$ can be approximated by summing over the squared magnitude of the *sampled* frequency-domain kernel $H_\nu[\boldsymbol{\mu}] \overset{\text{def}}{=} \textsc{Sample}_{1/T_0}\{\, H_\nu(\boldsymbol{\omega}) \,\}$. The kernel $H_\nu(\boldsymbol{\omega})$ is the third-order Volterra kernel which describes the nonlinear interaction between the probe channel and the ν^{th} wavelength channel in baseband description[9]. In practice, the sample operation is more likely performed on the periodic kernel $H_\nu(\mathrm{e}^{\mathrm{j}\boldsymbol{\omega}T_\mathrm{s}})$ representing the *oversampled* system, e.g., a two-fold oversampled system description (and the deduced integration bounds) is often considered to properly capture the signal pulse-shape. The triple integration is either performed by straightforward numerical integration [PBC+12] or more efficiently by Monte-Carlo integration [DFMS14]. The details on the sampled GN-model will be discussed in Section 5.1.1.

A comprehensive book chapter which provides a good overview on the current state of the art on both time-domain perturbative models and the related GN-models has recently been published in [BDS+20].

1.2 The Present Work

The preceding section provided an overview on the current state of the art on fiber-optical channel models summarized by the block diagram in Figure 1.1. The present work contributes to the existing theory, specifically, to the classes of channel models which are highlighted by a colored frame (—) in Figure 1.1. Beyond that, this work aims to connect the existing classes of channel models in a unified and structured manner. In the following, we summarize the main innovations.

Discrete-Time End-to-End RP Method The existing view on T-spaced end-to-end channel models for optical transmission systems is complemented by an equivalent *frequency*-domain description of the first-order RP method. The frequency-domain system formulation is inherently related to the discrete-time formulation from [ME12] by a discrete-time Fourier transform (DTFT). Following basic system-theoretic principles, this frequency-domain system description is necessarily $1/T$-periodic due to the time discretization with symbol spacing T. This is fundamentally different from prior frequency-domain RP methods, e.g. [PBP97,VSB02], which are based on the continuous-time end-to-end system description, see above.

The fundamental requirement on the complementary frequency-domain system description of a T-spaced channel is as follows: frequency components that appear—due to aliasing—at the sampled output of the continuous-time model when probed with a communication signal[10] must also be produced by the discrete-time Volterra system by an equivalent (frequency-domain) kernel. This property was first exploited in [Fra96] for a discrete-time Volterra system with oversampling. In the present case, the discrete-time (third-order) Volterra system is realized by aliasing the frequency-domain kernel of the continuous-time system (which has spectral support that spans over a three-dimensional frequency basis) into its Nyquist region (per

[9] For notational simplicity, we neglect MCI which is typically included in GN-models.

[10] The continuous-time communication signal is also $1/T$-periodic in frequency domain, but is weighted by the shape of the transfer function of the transmit-pulse.

dimension of the spectral support). In doing so, a frequency-domain RP solution is obtained which relates the periodic spectrum of the transmit *symbol* sequence to the periodic spectrum of the receive *symbol* sequence, i.e., after (linear) channel matched filtering and aliasing to frequencies within the Nyquist interval. Remarkably, the *frequency matching* which is a result of the general four-wave mixing (FWM) process in the continuous-time optical domain[11] is now realized in the $1/T$-periodic frequency domain where the fourth frequency is modulo reduced into the Nyquist interval, see Section 5.1.

Beyond the contribution to the basic theory, we present algorithmic implementations of the discrete-time Volterra system in *discrete* frequency domain. Similar to the implementations of the continuous-time model, see, e.g., the *enhanced* RP method, a subset of the perturbative contributions can be considered as multiplicative (including both a common phase and polarization rotation) and may also be implemented as such in $1/T$-periodic, discrete frequency domain. This motivates the extension of the original *regular* perturbation model in frequency domain to a combined *regular-logarithmic*, similar to the pulse-collision model in time domain.

We believe that both the existing time-domain end-to-end channel model according to the pulse collision picture, and the novel $1/T$-periodic frequency-domain end-to-end model have potential application in a variety of fields. Among those is the application as a *forward* channel model for the optimization of detection schemes that operate on a per-symbol basis, e.g., recovery of phase distortions or determination of symbol likelihood values. Similarly, in a *backward-propagation-sense*, both models can find application in fiber nonlinearity compensation which requires *real-time* processing using fixed-point arithmetic, i.e., implementation and computational complexity is of particular interest. In both forward and reverse applications, the kernel coefficients can be pre-calculated or pre-trained, whereas for the latter case, i.e., for fiber nonlinearity compensation, additionally adaptation of the kernel coefficients is required. Those application-driven considerations will, however, not be covered within the present work.

Discrete-Time GN-Model The discrete-time GN-model extends the conventional GN-model by taking the T-spaced sampling of the continuous-time communication signal into account. In this regard, the objective of the discrete-time GN-model is to evaluate the expectation $\mathrm{E}\{\, \|\Delta \boldsymbol{a}[k]\|^2 \,\}$, i.e., the amount of perturbative distortion to the transmit *symbols*. This quantity has, in fact, much more relevance for the overall system performance compared to the result of the conventional GN-model as it measures the amount of distortion that is present at the symbol decision (or soft metric calculation) on the receive-side. Only for the special case of a root-raised cosine (RRC) pulse-shape with zero roll-off, the discrete-time and the conventional GN-model will produce the same result[12]. The discrete-time GN-model follows the same conceptual idea as the conventional one but takes the *aliased* third-order Volterra kernel as a basis for estimating the magnitude of the nonlinear distortion.

[11] Frequency matching (here, of third order) is generally a property of nonlinear systems described by a (third-order) Volterra operator.

[12] After matched filtering with respect to the transmit RRC shape with zero roll-off, no spectral support of the communication signal is present outside the Nyquist interval, and, hence, no aliasing occurs.

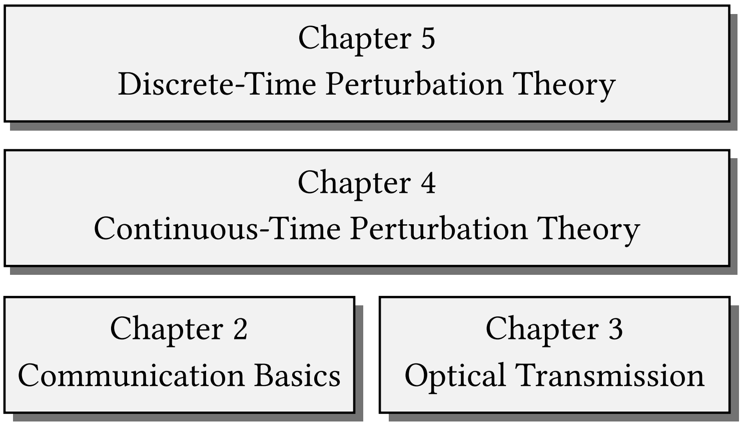


Figure 1.2: Structure of the dissertation.

The proposed discrete-time GN-model may supplement existing GN-models, where the communication signal's pulse-shape may not be neglected.

1.3 Outline of the Dissertation

The structure of the thesis is outlined in Figure 1.2 in a bottom-up representation. The foundational chapters cover the basic theory on communication systems and fiber-optical transmission. Building upon that, first, the continuous-time perturbation theory for nonlinear fiber propagation, and then, based thereon, the discrete-time perturbation theory of the respective end-to-end communication system is derived.

Specifically, in Chapter 2, the notation is briefly introduced and the theory of linear and nonlinear time-invariant systems is reviewed. For the latter, we study the theory of *Volterra series* to model nonlinear systems with memory. In the second part of Chapter 2, the general concepts in point-to-point communication systems are established, exemplified for transmission over a complex-valued 2×2 multiple-input/multiple-output (MIMO) channel as common in coherent fiber optical transmission. For such a *linear* channel, the discrete-time T-spaced end-to-end model is reviewed, including both the time- and frequency-domain formulation.

In Chapter 3, the fundamentals of fiber-optic transmission are examined. We introduce the *equivalent complex baseband* model for signals and systems in the *optical* domain. The baseband model is used to express signal evolution on the basis of a *nonlinear* partial differential equation (PDE)—the *Manakov equation*. First, only the *linear* part of the Manakov equation is considered to assess the two most relevant linear transmission effects: chromatic dispersion and attenuation (including gain from optical amplification). We assess the system impact of chromatic dispersion in a WDM transmission scenario and how it determines the system memory of a single wavelength channel and the (temporal) interaction length between wavelength channels. Then, the *optical Kerr effect*, the primary nonlinear transmission effect, is included into the picture. We briefly recap how the Manakov equation is implemented for numerical simulation using the split-step method.

Having established the theoretical foundation in first two Chapters, we proceed with the regular perturbation method for the continuous-time domain in Chapter 4. Starting from the Manakov equation, the derivation of the optical end-to-end first-order RP solution is performed. The relevant system parameters, i.e., *memory* and *strength*, of the nonlinear response are identified which lead to design rules for potential applications. Further, we highlight the relation between the time and frequency representation and point out the connection to other well-known channel models. As an intermediate step to the discrete-time end-to-end relation, we also introduce the (analog) baseband end-to-end model, which describes the nonlinear interaction based on the baseband communication signals in WDM setting.

In Chapter 5, the theory of the RP method is translated to the discrete-time domain. The theoretical considerations are complemented by numerical simulations which are in accordance with results obtained by the SSFM. Here, the mean-squared error (MSE) between the T-spaced output sequences of the RP method and the SSFM is assessed to determine the match between the models.

Chapter 6 presents some conclusions and an outlook.

A significant part of Chapter 4 and 5, which is at the center of this work, has been published as a pre-print on *arXiv* [FFF20], and presented in part at workshops and conferences on optical communication. Some of the numeric implementations of the derived models have been applied in fiber-optic system experiments in the context of fiber-nonlinearity mitigation [FME$^+$17, FES$^+$18], which is not part of the present work.

2. Foundations and Basic Concepts

This chapter briefly introduces the notation and basic concepts that will be used throughout this work. The chapter is divided into two parts.

The first part starts by establishing the most fundamental conventions on the notation. It is then followed by Section 2.1.1 with a brief revision of discrete- and continuous-time signals including a comment on how signals are normalized in this work. The next Section 2.1.2 defines all the relevant integral transforms in a convenient, multi-dimensional fashion. This will turn out to be useful in Section 2.1.3 which recaps the fundamentals of linear and nonlinear time-invariant systems. Nonlinear, time-invariant systems can be represented by a sum of multi-dimensional convolutions, known as Volterra series, in time or equivalently in frequency domain. The Volterra kernels in both domains (i.e., a generalized form of impulse response and transfer function for nonlinear systems) are interrelated by multi-dimensional integral transforms. A third-order nonlinear systems, i.e., one with cubic nonlinearity is discussed in more detail as it is closely connected to the nonlinear process present during fiber-optical propagation. The theory is supplemented by two basic examples of a third-order Hammerstein and Wiener system.

In the second part in Section 2.2, the general system model is introduced. The basic concepts and nomenclature of point-to-point transmission are assessed. This is done on the background of fiber-optical transmission where a complex-valued 2×2 MIMO channel is present. The whole cascade of transmitter, channel, and receiver is discussed using a simplified setting. Here, only a noisy, non-dispersive channel is considered, while the next chapter addresses the fundamentals of fiber-optic transmission in greater detail. Both the continuous- and discrete-time end-to-end model for the case of a noisy, non-dispersive channel are established. A formal definition of the signal-to-noise ratio is provided and the Shannon capacity is introduced.

2.1 Notation and Basic Definitions

Non-bold italic letters, like x, are scalar variables, whereas non-bold Roman letters refer to constants, e.g., the speed of light is denoted by

$$\mathrm{c} = 299792458\ \mathrm{m/s}. \tag{2.1}$$

Number Sets A *number set* or *finite field* is typeset in blackboard typeface, e.g., the set of real numbers is $\mathbb{R}$, and the set of non-negative real numbers is $\mathbb{R}_{\geq 0}$. The set of complex numbers $\mathbb{C}$ is an extension of $\mathbb{R}$ with elements defined as

$$z = x + \mathrm{j}\, y\ \in \mathbb{C}\,, \tag{2.2}$$

with the imaginary number $\mathrm{j} \stackrel{\text{def}}{=} \sqrt{-1}$, and the Cartesian components $x = \mathrm{Re}\{z\} \in \mathbb{R}$ (aka. *real* part, or *inphase* component), and $y = \mathrm{Im}\{z\} \in \mathbb{R}$ (aka. *imaginary* part, or *quadrature* component). A complex number can also be expressed in Polar coordinates via its magnitude

$$r = |z| \stackrel{\text{def}}{=} \sqrt{\mathrm{Re}^2\{z\} + \mathrm{Im}^2\{z\}}, \tag{2.3}$$

and its phase (or angle)

$$\varphi = \arg\{z\} \stackrel{\text{def}}{=} \tan^{-1}(y/x), \tag{2.4}$$

to be combined in the Polar representation of complex numbers

$$z = r\, \mathrm{e}^{\mathrm{j}\,\varphi}. \tag{2.5}$$

Complex conjugation reverses the sign of the imaginary part, or equivalently the sign of the angle, and is denoted by

$$z^* = \mathrm{Re}\{z\} - \mathrm{j}\,\mathrm{Im}\{z\} = r\, \mathrm{e}^{-\mathrm{j}\,\varphi}. \tag{2.6}$$

The set of integers $\mathbb{Z}$ is a subset of the real numbers

$$\mathbb{Z} = \{\, -\infty, \ldots, -2, -1, 0, 1, 2, \ldots, \infty \,\}, \tag{2.7}$$

and the set of natural numbers is a subset of the integers obtained by taking all nonnegative elements

$$\mathbb{N} = \mathbb{Z}_{\geq 0} = \{\, 0, 1, 2, \ldots, \infty \,\} = \{\, 0 \,\} \cap \mathbb{N}_{>0}. \tag{2.8}$$

Elements from the binary Galois field $\mathbb{F}_2$ and derived variables thereof are written in *Fraktur font*. E.g., the two elements of $\mathbb{F}_2$ are $\{\, \mathfrak{0}, \mathfrak{1} \,\}$, and a variable form that field is $\mathfrak{q} \in \mathbb{F}_2$, where operations (defined on such variables) have algebraic properties associated with $\mathbb{F}_2$.

Other (non-special) sets are denoted with calligraphic letters, e.g., $\mathcal{A}$ will later denote the set of data symbols, i.e., the symbol *alphabet* or *signal constellation*.

Vectors and Matrices Bold letters, such as $\boldsymbol{x}$, indicate vectors or matrices (capital letters are reserved for variables in the transform domain, see next section). The difference between a vector and a matrix should be clear from the context. If not stated otherwise, a (complex) vector $\boldsymbol{x}$ of dimension N is a column vector and reads

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} = [x_1, x_2, \dots, x_N]^{\mathsf{T}} = [x_1^*, x_2^*, \dots, x_N^*]^{\mathsf{H}} \in \mathbb{C}^N, \tag{2.9}$$

where $(\cdot)^{\mathsf{T}}$ denotes transposition and $(\cdot)^{\mathsf{H}}$ denotes Hermitian transposition.

Similarly, boldface letters can also represent matrices, e.g., a (complex) matrix of dimension $N \times M$ reads

$$\boldsymbol{x} = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,M} \\ x_{2,1} & x_{2,2} & & x_{2,M} \\ \vdots & & & \vdots \\ x_{N,1} & x_{N,2} & \cdots & x_{N,M} \end{bmatrix} = \left[x_{n,m}\right] \in \mathbb{C}^{N\times M} \ , \quad n = 1, \dots, N \ , \quad m = 1, \dots, M \ , \tag{2.10}$$

with entries $x_{n,m}$. The same notation applies to vectors and matrices of different fields.

We define the *matrix exponential* of square matrices $\boldsymbol{x} \in \mathbb{C}^{N\times N}$ as a power series

$$\exp(\boldsymbol{x}) = \mathrm{e}^{\boldsymbol{x}} \stackrel{\text{def}}{=} \sum_{k=0}^{\infty} \frac{1}{k!} \boldsymbol{x}^k, \tag{2.11}$$

where $\boldsymbol{x}^0 = \mathbf{I} \in \mathbb{C}^{N\times N}$. If the matrix $\boldsymbol{x}$ is diagonal with elements $x_{n,n}$ on the main diagonal and zeros otherwise, then the matrix exponential is obtained by exponentiating each element on the main diagonal $\mathrm{e}^{x_{n,n}}$ for $n = 1, \dots, N$. As a result, for a diagonalizable matrix $\boldsymbol{x}$, we have

$$\boldsymbol{x} = \boldsymbol{u}\boldsymbol{d}\boldsymbol{u}^{-1}, \tag{2.12}$$

and $\boldsymbol{d} \in \mathbb{C}^{N\times N}$ is diagonal, then

$$\mathrm{e}^{\boldsymbol{x}} = \boldsymbol{u}\mathrm{e}^{\boldsymbol{d}}\boldsymbol{u}^{-1}. \tag{2.13}$$

In optical communication, vectorial variables often appear, e.g., within the so-called *Jones* or *Stokes* formalism, both formally introduced in Appendix A.1. To indicate variables from Jones or Stokes space we introduce the following notation.

A Jones variable $\boldsymbol{x} \in \mathbb{C}^2$ is a pair of complex numbers represented as a two-dimensional vector, where each dimension relates to one of the two polarizations of light, denoted as x- and y-polarization. To emphasis this association on occasion, we may subindex the components of $\boldsymbol{x}$ as follows

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_{\mathsf{x}} \\ x_{\mathsf{y}} \end{bmatrix} \in \mathbb{C}^2. \tag{2.14}$$

A Stokes variable $\vec{\boldsymbol{x}} \in \mathbb{R}^3$ is a 3-tuple of real numbers represented as a three-dimensional vector, where each dimension relates to one of the three polarization states, i.e., linearly-polarized, $45°$-polarized, and circularly-polarized light. We will use decorated bold letters $\vec{\boldsymbol{x}}$ to denote vectors in *Stokes space.*

2.1.1 Discrete- and Continuous-Time Signals

A (complex-valued) function $x(t) \in \mathbb{C}$ which depends on the real-valued time variable $t \in \mathbb{R}$, given in *seconds* (s), is called *continuous-time signal.* The support of a (time-domain) function is defined as the set of all elements $t \in \mathbb{R}$ for which $x(t)$ is non-zero, i.e.,

$$\operatorname{supp}(x) \overset{\text{def}}{=} \{\, t \in \mathbb{R} \mid x(t) \neq 0 \,\}\,. \tag{2.15}$$

A (complex-valued) N-dimensional continuous-time signal $x(t_1, t_2, \ldots, t_N) \in \mathbb{C}$ is a function of N time variables $t_n \in \mathbb{R}$ with $n = 1, \ldots, N$. We will write

$$x(\boldsymbol{t}) \overset{\text{def}}{=} x(t_1, t_2, \ldots, t_N), \tag{2.16}$$

with $\boldsymbol{t} = [t_1, t_2, \ldots, t_N]^{\mathsf{T}} \in \mathbb{R}^N$ for short. The reader should be able to deduce the dimensionality of the vector from the context of the equation.

Additionally, a space-dependent signal $x(z, t)$ depends on the location (or position) $z \in \mathbb{R}$ given in *meter* (m), usually relative to the position of the transmitter. In point-to-point optical communications, the variable z describes the (one-dimensional) path along the fiber, measured from the transmitter at $z = 0$ to the receiver at $z = L \in \mathbb{R}_{>0}$, i.e., in practical cases it only takes values in the interval $z \in [0, L]$.

A *discrete-time* sequence is obtained by sampling a continuous-time signal at intervals of, e.g., the (temporal) spacing T. We define the discrete-time signal

$$x[k] \overset{\text{def}}{=} x(kT), \tag{2.17}$$

with the discrete-time (index) variable $k \in \mathbb{Z}$. We use round parenthesis $(\,\cdot\,)$ to denote functions defined over continuous variables and square brackets $[\,\cdot\,]$ to denote functions defined over discrete variables. If the *whole* (finite- or infinite-length) sequence is treated, we will use the angled bracket notation $\langle\, x[k] \,\rangle$, whereas $x[k]$ is a single element of that sequence indexed at position k.

The support of a sequence is defined as the set of all elements $k \in \mathbb{Z}$ for which $x[k]$ is non-zero, i.e.,

$$\operatorname{supp}(x) \overset{\text{def}}{=} \{\, k \in \mathbb{Z} \mid x[k] \neq 0 \,\}\,. \tag{2.18}$$

The energy of a continuous- and discrete-time signal are defined as [PS08]

$$E_x \overset{\text{def}}{=} \int_{\mathbb{R}} |x(t)|^2 \,\mathrm{d}t, \qquad\qquad E_x \overset{\text{def}}{=} \sum_{k\in\mathbb{Z}} |x[k]|^2. \tag{2.19}$$

Signal Normalization Continuous-time signals are associated with meaningful physical units, e.g., the electrical field can be measured in units of *volts per meter* (V/m). In the mathematical treatment of those physical signals, units are often considered inconvenient. Consequently, a normalization is often performed to create dimensionless (in the sense of *unitless*) mathematical quantities which have the same orientation as the original physical object (see, e.g., discussion in [Fis02, P. 11] or [Kam08, P. 230]).

In the context of optical transmission, the propagation of the *real* optical field is typically modeled in complex Jones space over a quantity $\boldsymbol{u}(z,t) = [u_{\mathsf{x}}(z,t), u_{\mathsf{y}}(z,t)]^{\mathsf{T}} \in \mathbb{C}^2$ called *optical field envelope*, see the disccusion in Section A.1. The optical field envelope is the equivalent complex baseband (ECB) representation [Fis02] of the *real* optical signal, see Section 3.1. It has the same orientation as the associated electrical field, where the two vector elements $u_{\mathsf{x}}(z,t)$ and $u_{\mathsf{y}}(z,t)$ correspond to the two polarizations of the optical field.

In the literature, we encounter various conventions on how the optical field envelope is normalized. A common strategy is to normalize $\boldsymbol{u}(z,t)$ such that $\boldsymbol{u}^{\mathsf{H}}(z,t)\boldsymbol{u}(z,t)$ equals the instantaneous (and local) power given in *watts* (W), cf. [Agr06, Eq. (2.3.28)].

In this work, instead, any continuous-time signal, like $\boldsymbol{u}(z,t)$, is generally treated as dimensionless quantity. This considerably simplifies the notation when we move between the various signal domains, e.g., form discrete-time to continuous-time or from analog-electrical to analog-optical domain (and vice-versa). To this end, $\boldsymbol{u}^{\mathsf{H}}(z,t)\boldsymbol{u}(z,t)$ is normalized to be dimensionless. Consequently, the *nonlinearity coefficient* $\gamma(z)$, see Section 3.3, commonly given in $\mathrm{W}^{-1}\mathrm{m}^{-1}$ must also be normalized to have units of m^{-1} to be consistent with our signal definition. We will touch on this notational convention again in the relevant Sections where the above mentioned quantities will be formally introduced.

2.1.2 Integral Transforms

The Fourier transform and related integral transforms play a central role in the theory of signals and systems.

It turns out that, in the context of fiber-optic transmission (and for nonlinear systems in general), some relations between a transformation pair (e.g., time and frequency domain) appear in higher dimensions. Accordingly, all integral transforms are introduced for signals in N dimensions. In particular, the Fourier transform is introduced given a function $x(\boldsymbol{t}) \in \mathbb{C}$ which may depend on the N-dimensional variable $\boldsymbol{t} \in \mathbb{R}^N$ over which the transformation is performed. For the case $N = 1$, the following definitions will be equivalent to the commonly known one-dimensional transformations. Then, all subindices, e.g., as in t_1, are dropped for better readability.

All (continuous-) frequency domain variables are expressed in terms of the angular frequency $\boldsymbol{\omega} = 2\pi \boldsymbol{f} \in \mathbb{R}^N$ with each frequency variable $f_n \in \mathbb{R}$, $n = 1, \ldots, N$, measured in *Hertz* (Hz).

2.1.2.1 The Fourier Transform

We use lower-case letters for *time*-domain signals such as $x(\boldsymbol{t})$, and upper-case letters for *frequency*-domain signals such as $X(\boldsymbol{\omega})$. The upper-case notation in frequency domain may collide with the notation often used in the context of information-theoretic elements, where an upper-case letter will indicate a random variable (RV), and its realization is written in lower-case letters. The reader should be able to follow the text and understand the respective terminology from the present context of the text.

In the present work the so-called *engineering* notation of the Fourier transform with a negative sign in the complex exponential in the forward, i.e., time-to-frequency direction is used[1].

The N-dimensional *Fourier transform* of a continuous-time signal $x(\boldsymbol{t}) = x(t_1, t_2, \ldots, t_N)$ depending on the N-dimensional time vector $\boldsymbol{t} = [t_1, t_2, \ldots, t_N]^\mathsf{T} \in \mathbb{R}^N$ is defined as [OW83, Ch. 4]

$$X(\boldsymbol{\omega}) = \mathcal{F}\{\, x(\boldsymbol{t}) \,\} \stackrel{\text{def}}{=} \int\limits_{\mathbb{R}^N} x(\boldsymbol{t})\, \mathrm{e}^{-\mathrm{j}\boldsymbol{\omega}\cdot\boldsymbol{t}}\, \mathrm{d}^N\boldsymbol{t} \tag{2.20}$$

$$\bullet\!\!-\!\!\!\circ$$

$$x(\boldsymbol{t}) = \mathcal{F}^{-1}\{\, X(\boldsymbol{\omega}) \,\} = \frac{1}{(2\pi)^N} \int\limits_{\mathbb{R}^N} X(\boldsymbol{\omega})\, \mathrm{e}^{\mathrm{j}\boldsymbol{\omega}\cdot\boldsymbol{t}}\, \mathrm{d}^N\boldsymbol{\omega}. \tag{2.21}$$

Here, $X(\boldsymbol{\omega})$ is a function of (continuous-) angular frequencies $\boldsymbol{\omega} = [\omega_1, \omega_2, \ldots, \omega_N]^\mathsf{T} \in \mathbb{R}^N$.

[1] This has an immediate effect on the solution of the electro-magnetic wave equation (cf. *Helmholtz equation* in [Agr06, Eq. (2.1.18)]), and therefore also for the NLSE. In the optical community, there is no fixed convention with respect to the sign notation, e.g., some texts are written with the physicists' (e.g., [Agr06, Eq. (2.1.12)], [Agr10, Eq. (2.2.8)] or [ME12]) and others with the engineering (e.g., [Kam13], [Eng15, Eq. (A.4)]) notation in mind. Consequently, the derivations shown in the remaining chapters may differ marginally from some of the original sources.

In the exponential we made use of the dot product of vectors in $\mathbb{R}^N$ given by

$$\boldsymbol{\omega} \cdot \boldsymbol{t} \overset{\text{def}}{=} \omega_1 t_1 + \omega_2 t_2 + \cdots + \omega_N t_N \,. \tag{2.22}$$

The integral is an N-fold integral over $\mathbb{R}^N$ and the integration boundaries are at $-\infty$ and ∞ in each dimension. Note, since $x(\boldsymbol{t})$ is per-definition unitless, its Fourier-transform $X(\boldsymbol{\omega})$ has units seconds^N (s^N). For the differential, we use the shorthand notation

$$\mathrm{d}^N \boldsymbol{t} \overset{\text{def}}{=} \mathrm{d}t_1 \mathrm{d}t_2 \ldots \mathrm{d}t_N. \tag{2.23}$$

We may also write the correspondence as $x(\boldsymbol{t}) \circ\!\!-\!\!\bullet\, X(\boldsymbol{\omega})$ for short, where the transformation applies to the *whole* function, not just a single value. We may also indicate a change of variables by, e.g., $X(\boldsymbol{\omega}) = \mathcal{F}_{\boldsymbol{\omega}\leftrightarrow\boldsymbol{\tau}}\{\, x(\boldsymbol{\tau})\,\}$, if the transformation pair is defined w.r.t. to another argument.

If instead the transformation is carried out over multiple vector components of the function $\boldsymbol{x}(t) = [x_1(t), x_2(t), \ldots, x_L(t)]^\mathsf{T} \in \mathbb{C}^L$, then each vector component is transformed independently with $\boldsymbol{X}(\omega) = [X_1(\omega), X_2(\omega), \ldots, X_L(\omega)]^\mathsf{T} \in \mathbb{C}^L$ s.t. $x_l(t) \circ\!\!-\!\!\bullet\, X_l(\omega)$ with $l = 1, \ldots, L$.

We define the Dirac delta function (aka. Dirac impulse) implicitly using its Fourier transform

$$\mathcal{F}\{\, \delta(t)\,\} = \int_{\mathbb{R}} \delta(t) \mathrm{e}^{-\mathrm{j}\omega t}\, \mathrm{d}t = \mathrm{e}^{-\mathrm{j}\omega 0} = 1 \quad \forall \omega, \tag{2.24}$$

and the Heaviside step function (aka. unit step) via the Dirac impulse as

$$\varepsilon(t) \overset{\text{def}}{=} \int_{-\infty}^{t} \delta(\tau)\, \mathrm{d}\tau = \begin{cases} 1\,, & t > 0 \\ 0\,, & t < 0 \end{cases} . \tag{2.25}$$

2.1.2.2 The Discrete-Time Fourier Transform

The N-dimensional discrete-time Fourier transform (DTFT) of a discrete-time sequence $\langle\, x[\boldsymbol{k}]\,\rangle$ with $\boldsymbol{k} = [k_1, k_2, \ldots, k_N]^\mathsf{T} \in \mathbb{Z}^N$ is defined as [OW83, Ch. 5]

$$X(\mathrm{e}^{\mathrm{j}\boldsymbol{\omega}T}) = \hat{\mathcal{F}}\{\, x[\boldsymbol{k}]\,\} \overset{\text{def}}{=} \sum_{\boldsymbol{k}\in\mathbb{Z}^N} x[\boldsymbol{k}]\, \mathrm{e}^{-\mathrm{j}\boldsymbol{\omega}\cdot\boldsymbol{k}T} \tag{2.26}$$

$$x[\boldsymbol{k}] = \hat{\mathcal{F}}^{-1}\{\, X(\mathrm{e}^{\mathrm{j}\boldsymbol{\omega}T})\,\} = \left(\frac{T}{2\pi}\right)^N \int_{\mathbb{T}^N} X(\mathrm{e}^{\mathrm{j}\boldsymbol{\omega}T})\, \mathrm{e}^{\mathrm{j}\boldsymbol{\omega}\cdot\boldsymbol{k}T}\, \mathrm{d}^N\boldsymbol{\omega}. \tag{2.27}$$

where T is the (time-constant, temporal) spacing between two elements of the time-domain sequence. In this thesis, the variable T will be equivalent with the duration of the symbol interval of the modulated signal.

The notation $\sum_{\boldsymbol{k}\in\mathbb{Z}^N}$ is short for $\sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \cdots \sum_{k_N=-\infty}^{\infty}$. Given that the sample duration is T, it directly follows that the spectrum of the associated discrete-time sequence is periodic with $1/T$. This is indicated by the notation $X(\mathrm{e}^{\mathrm{j}\boldsymbol{\omega}T})$ which takes the $1/T$-periodic exponential as an argument. The set of frequencies within the fundamental period (aka. *Nyquist interval*) of the spectrum is defined as

$$\mathbb{T} \overset{\text{def}}{=} \{\, \omega \in \mathbb{R} \mid -\omega_{\text{Nyq}} \le \omega < \omega_{\text{Nyq}}\,\}, \tag{2.28}$$

with the Nyquist (angular) frequency $\omega_{\text{Nyq}} \stackrel{\text{def}}{=} 2\pi f_{\text{Nyq}} = 2\pi/(2T)$.

The periodic spectrum $X(\mathrm{e}^{\mathrm{j}\boldsymbol{\omega}T})$ is sometimes also termed *periodic continuation* of the aperiodic spectrum $X(\boldsymbol{\omega})$. The former one is obtained by aliasing all spectral components of the latter into the Nyquist interval. We define the aliasing operator as

$$\begin{aligned} X(\mathrm{e}^{\mathrm{j}\boldsymbol{\omega}T}) &= \textsc{Alias}_{\,\omega_{\text{Nyq}}}\{\, X(\boldsymbol{\omega})\,\} \stackrel{\text{def}}{=} \frac{1}{T^N} \sum_{\boldsymbol{m}\in\mathbb{Z}^N} X(\boldsymbol{\omega} - 2\omega_{\text{Nyq}}\,\boldsymbol{m}) && (2.29) \\ &\underset{\circ}{\overset{\bullet}{|}} \\ x[\boldsymbol{k}] &= \textsc{Sample}_{\,T}\{\, x(\boldsymbol{t})\,\} \stackrel{\text{def}}{=} x(\boldsymbol{k}T), && (2.30) \end{aligned}$$

corresponding to a sampling operation in time domain with $\boldsymbol{t} = \boldsymbol{k}T$ and $\boldsymbol{k} \in \mathbb{Z}^N$.

2.1.2.3 The Discrete Fourier Transform

The N-dimensional discrete Fourier transform (DFT) of a *finite*-length discrete-time sequence $\langle\, x[\boldsymbol{k}]\,\rangle$ (M samples per dimension) is defined as

$$\begin{aligned} X[\boldsymbol{\mu}] &= \ \mathrm{DFT}\{\, x[\boldsymbol{k}]\,\} \ \stackrel{\text{def}}{=} \sum_{\boldsymbol{k}\in\mathbb{M}^N} x[\boldsymbol{k}]\, \mathrm{e}^{-\mathrm{j}\frac{2\pi}{M}\boldsymbol{\mu}\cdot\boldsymbol{k}} && (2.31) \\ &\underset{\circ}{\overset{\bullet}{|}} \\ x[\boldsymbol{k}] &= \mathrm{DFT}^{-1}\{\, X[\boldsymbol{\mu}]\,\} = \frac{1}{M^N} \sum_{\boldsymbol{\mu}\in\mathbb{M}^N} X[\boldsymbol{\mu}]\, \mathrm{e}^{\mathrm{j}\frac{2\pi}{M}\boldsymbol{\mu}\cdot\boldsymbol{k}}, && (2.32) \end{aligned}$$

where we use the reduced number set defined as

$$\mathbb{M} \stackrel{\text{def}}{=} \{\, 0, 1, \ldots, M-1 \,\} = \mathbb{Z} \mod M. \tag{2.33}$$

Both integer indices $\boldsymbol{k}$ and $\boldsymbol{\mu}$ are from the finite set $\mathbb{M}^N$, where M coincides with the length of the sequence.

We may also write $x[\boldsymbol{k}] \circ\!\!-\!\!\bullet\, X[\boldsymbol{\mu}]$ for short where the correspondence always relates the whole sequence of discrete-time and discrete-frequency domain elements.

2.1.2.4 The z-Transform and Laplace-Transform

The N-dimensional z-transform of a discrete-time sequence $\langle\, x[\boldsymbol{k}]\,\rangle$ is defined as

$$X(\boldsymbol{z}) = \ \mathcal{Z}\{\, x[\boldsymbol{k}]\,\} \stackrel{\text{def}}{=} \sum_{\boldsymbol{k}\in\mathbb{Z}^N} x[\boldsymbol{k}]\, z_1^{-k_1} z_2^{-k_2} \ldots z_N^{-k_N} = \sum_{\boldsymbol{k}\in\mathbb{Z}^N} x[\boldsymbol{k}]\, \boldsymbol{z}^{\cdot(-\boldsymbol{k})}, \tag{2.34}$$

where $\boldsymbol{k} \in \mathbb{Z}^N$ is the N-dimensional integer time index and $\boldsymbol{z} = [z_1, z_2, \ldots, z_N]^{\mathsf{T}} \in \mathbb{C}^N$ is an N-dimensional complex number. Here, we use the generalization of the vector dot *product* to the vector dot *exponential* defined as

$$\boldsymbol{z}^{\cdot(-\boldsymbol{k})} \stackrel{\text{def}}{=} z_1^{-k_1} z_2^{-k_2} \ldots z_N^{-k_N}. \tag{2.35}$$

The z-Transform is related to the DTFT by evaluating the variable $\boldsymbol{z}$ at its unit circle (or unit hypersphere in N dimensions), i.e., by evaluating $\boldsymbol{z} = \mathrm{e}^{\mathrm{j}\boldsymbol{\omega}T}$.

For continuous-time variables, the N-dimensional Laplace transform is defined accordingly as

$$X_{\mathscr{L}}(\boldsymbol{s}) = \mathscr{L}\{\, x(\boldsymbol{t}) \,\} \stackrel{\text{def}}{=} \int_{\mathbb{R}^N} x(\boldsymbol{t})\, \mathrm{e}^{-\boldsymbol{s}\cdot\boldsymbol{t}}\, \mathrm{d}^N\boldsymbol{t}, \tag{2.36}$$

with $\boldsymbol{s} \stackrel{\text{def}}{=} \boldsymbol{\sigma} + \mathrm{j}\,\boldsymbol{\omega} \in \mathbb{C}^N$ and real-valued components $\boldsymbol{\sigma} \in \mathbb{R}^N$ and $\boldsymbol{\omega} \in \mathbb{R}^N$. The Laplace-transform is related to the Fourier transform by evaluating the variable $\boldsymbol{s}$ only at its imaginary part (i.e., setting the real part $\boldsymbol{\sigma}$ to zero).

2.1.3 Linear- and Nonlinear Time-Invariant Systems

This section provides a short review of linear- and nonlinear time-invariant systems. We discuss the most relevant system properties and give basic examples of nonlinear systems.

A continuous-time system translates a continuous-time input signal $x(t) \in \mathbb{C}$ to an output signal $y(t) \in \mathbb{C}$ via the transformation $y(t) = \mathcal{S}\{\, x(t) \,\}$. Here, $\mathcal{S}\{\, \cdot \,\}$ must be understood as an operator acting on the *whole* input function, and returning the complete output function, not just the function evaluated at a particular instant of time.

A *time-invariant* system is given if an arbitrary delay $t_0 \in \mathbb{R}$ of the input signal translates to the same delay at the output. Given the relation $y(t) = \mathcal{S}\{\, x(t) \,\}$, we find

$$y(t - t_0) = \mathcal{S}\{\, x(t - t_0) \,\}, \quad t_0 \in \mathbb{R}\,. \tag{2.37}$$

2.1.3.1 Linear Time-invariant Systems

The input/output relation of a stable, non-recursive (continuous-time) linear time-invariant (LTI) system can be mathematical expressed by the convolution of the input signal $x(t)$ with the (linear) *impulse response* of the system $h(t) \overset{\text{def}}{=} \mathcal{S}\{\, \delta(t) \,\} \in \mathbb{C}$, where $\delta(t)$ denotes the Dirac impulse. It is given by

$$y(t) = \mathcal{S}\{\, x(t) \,\} \quad = x(t) * h(t) \overset{\text{def}}{=} \int_{\mathbb{R}} x(\tau)h(t-\tau)\mathrm{d}\tau = \int_{\mathbb{R}} x(t-\tau)h(\tau)\mathrm{d}\tau \tag{2.38}$$

$$\circ\!\!-\!\!\bullet$$

$$Y(\omega) = \mathcal{S}\{\, X(\omega) \,\} = X(\omega)\, H(\omega)\,, \tag{2.39}$$

and equivalently, in frequency domain by multiplying the Fourier-transform (aka. spectrum) of the input signal $X(\omega) = \mathcal{F}\{\, x(t) \,\}$ with the system *transfer function* $H(\omega) \overset{\text{def}}{=} \mathcal{F}\{\, h(t) \,\} \in \mathbb{C}$ to obtain the Fourier-transform of the output signal $Y(\omega)$.

An LTI system is bounded-input bounded-output (BIBO) stable, if for any bounded input the system response is also bounded

$$|x(t)| < x_{\max} < \infty \quad \Rightarrow \quad |y(t)| < y_{\max} < \infty, \;\; t \in \mathbb{R}\,. \tag{2.40}$$

A necessary and sufficient condition for *stable* LTI systems with bounded input is

$$\int_{\mathbb{R}} |h(t)|\, \mathrm{d}t < \infty\,, \tag{2.41}$$

i.e., the impulse response must be a *transient* function.

For the output signal $y(t)$ to have the same units as the input signal $x(t)$, the continuous-time impulse response $h(t)$ must have units s^{-1}, and hence $H(\omega)$ is unitless.

The same considerations also hold for *discrete-time* LTI systems, where we find the relation

$$y[k] = \mathcal{S}\{\, x[k] \,\} \qquad = x[k] * h[k] \overset{\text{def}}{=} \sum_{\kappa \in \mathbb{Z}} x[\kappa]h[k-\kappa] = \sum_{\kappa \in \mathbb{Z}} x[k-\kappa]h[\kappa] \tag{2.42}$$

$$\circ\!\!-\!\!\bullet$$

$$Y(\mathrm{e}^{\mathrm{j}\omega T}) = \mathcal{S}\{\, X(\mathrm{e}^{\mathrm{j}\omega T}) \,\} = X(\mathrm{e}^{\mathrm{j}\omega T})\, H(\mathrm{e}^{\mathrm{j}\omega T}) = X(z)\, H(z)\Big|_{z=\mathrm{e}^{\mathrm{j}\omega T}}\,, \tag{2.43}$$

with the discrete-time transfer function (aka. *system function*) $H(z) \overset{\text{def}}{=} \mathcal{Z}\{\, h[k] \,\}$. Both, the discrete-time impulse response $h[k]$ and the transfer function $H(z)$ are unitless.

2.1.3.2 Nonlinear Time-invariant Systems

The input/output relation of a wide class of nonlinear time-invariant (NTI) systems can be represented using a *Volterra series* expansion — a sum of multi-dimensional convolutions. The Volterra series expansion can be considered as a Taylor series expansion for nonlinear systems with memory, whereas the usual Taylor series only applies to systems with instantaneous (i.e., *memoryless*) input/output relations.

The infinite (i.e., non-truncated) Volterra series representing a continuous-time, nonlinear system $\mathcal{S}\{\,\cdot\,\}$ without feedback (i.e., no recursion) is given by the relation [Sch80]

$$\begin{aligned} y(t) = \mathcal{S}\{\, x(t) \,\} = h_0 &+ \int_{\mathbb{R}} x(t-\tau_1)\, h_1(\tau_1)\, \mathrm{d}\tau_1 && (2.44)\\ &+ \int_{\mathbb{R}^2} x(t-\tau_1)x(t-\tau_2)\, h_2(\tau_1,\tau_2)\, \mathrm{d}\tau_1\mathrm{d}\tau_2 && (2.45)\\ &+ \int_{\mathbb{R}^3} x(t-\tau_1)x(t-\tau_2)x(t-\tau_3)\, h_3(\tau_1,\tau_2,\tau_3)\, \mathrm{d}\tau_1\mathrm{d}\tau_2\mathrm{d}\tau_3 && (2.46)\\ &+ \dots \\ = h_0 &+ \sum_{n=1}^{\infty} \mathfrak{y}_n(t)\,, && (2.47) \end{aligned}$$

where each summand $\mathfrak{y}_n(t)$ corresponds to an n^{th}-order convolution comprising a product of (time-delayed) input signals $x(t)$ weighted by a so-termed *Volterra kernel* $h_n(\boldsymbol{\tau})$—a generalized impulse response of order n. The summand $\mathfrak{y}_n(t) = \mathcal{H}_n\{\, x(t) \,\}$ is called n^{th}-order *Volterra operator* and its *standard* form is given by

$$\mathfrak{y}_n(t) = \mathcal{H}_n\{\, x(t) \,\} \stackrel{\text{def}}{=} \int_{\mathbb{R}^n} h_n(\boldsymbol{\tau}) \prod_{j=1}^{n} x(t-\tau_j)\, \mathrm{d}^n\boldsymbol{\tau}\,, \tag{2.48}$$

where the kernel $h_n(\boldsymbol{\tau}) = h_n(\tau_1, \tau_2, \dots, \tau_n) \in \mathbb{C}$ is a function $\mathbb{R}^n \mapsto \mathbb{C}$ depending on the order $n \in \mathbb{N}_{>0}$. We note, that the n^{th}-order Volterra kernel $h_n(\boldsymbol{\tau})$ has units s^{-n}.

The Volterra operator is sometimes also given in its alternative representation as

$$\mathcal{H}_n\{\, x(t) \,\} = \int_{\mathbb{R}^n} h_n(t-\mathbf{t}) \prod_{j=1}^{n} x(\mathrm{t}_j)\, \mathrm{d}^n\mathbf{t}\,, \tag{2.49}$$

with $\boldsymbol{\tau} = t - \mathbf{t}$, i.e., $\tau_j = t - \mathrm{t}_j, \forall\, j$. In particular, for $n = 1$ this recovers the two ways the convolution is expressed in (2.38).

In many practical cases, the infinite-length series in (2.47) is truncated to only a subset of summation terms. E.g., most relevant for nonlinear fiber communication are the Volterra terms $\mathfrak{y}_1(t)$ and $\mathfrak{y}_3(t)$—the linear part equivalent to an LTI system, and the third-order nonlinear part.

Symmetry A Volterra kernel is symmetric, if $h_n(\boldsymbol{\tau})$ is invariant under any reordering of the vector $\boldsymbol{\tau}$, i.e., we have [Sch80, Eq. (5.2-4)]

$$h_n(\boldsymbol{\tau}) = h_n(\, \pi_i(\boldsymbol{\tau})\,), \quad \forall i\,, \tag{2.50}$$

where the function $\pi_i(\cdot)$ performs the $i \in \{\, 1, 2, \dots, n!{-}1 \,\}$ (non-trivial) permutations w.r.t. the vector indices, e.g., for a symmetric second-order kernel we have $h_2(\tau_1, \tau_2) = h_2(\tau_2, \tau_1)$.

Causality A Volterra system is causal, if [Sch80, Eq. (5.5-2)]

$$h_n(\boldsymbol{\tau}) = h_n(\tau_1, \tau_2, \ldots, \tau_n) = 0, \qquad \text{for any } \tau_j < 0, \quad j = 1, 2, \ldots, n\,. \tag{2.51}$$

Stability The series in (2.47) does not necessarily converge for all classes of nonlinear systems and input signals $x(t)$. If the series does converge for any given input $x(t)$, we say that the system is *stable*.

In most practical cases, we consider only a truncated Volterra series. Then, the question of stability is shifted to the stability of the Volterra *operators* $\mathcal{H}_n\{\, x(t)\,\}$ under consideration. Similar to LTI systems, a Volterra operator is BIBO stable, if for any *bounded* input the operator response is also bounded

$$|x(t)| < x_{\max} < \infty \quad \Rightarrow \quad |\mathsf{y}_n(t)| < \mathsf{y}_{n,\max} < \infty, \;\; t \in \mathbb{R}, \forall n\,. \tag{2.52}$$

A sufficient (but not necessary) condition for the stability of Volterra operators is the straightforward extension of (2.41) to [Sch80, Eq. (5.6-3)]

$$\int_{\mathbb{R}^n} |h_n(\boldsymbol{\tau})| \,\mathrm{d}^n\boldsymbol{\tau} < \infty, \quad \forall n\,. \tag{2.53}$$

Note, that there are Volterra operators which are stable but not transient, i.e., do not satisfy (2.53). The condition is also necessary for the Fourier transform of $h_n(\boldsymbol{\tau})$ to exist, see below.

Volterra Series in Transform Domain Taking the Fourier transform of the system response $y(t)$ leads to the Volterra series representation in continuous-frequency domain, aka. Volterra series transfer function (VSTF). It relates the Fourier-transform at the input $X(\omega)$ to the Fourier-transform at the output $Y(\omega)$ of the nonlinear system via [Sch80, Ch. 6]

$$\begin{aligned} Y(\omega) = h_0\, \delta(\omega) &+ X(\omega)\, H_1(\omega) && (2.54)\\ &+ \frac{1}{2\pi}\int_{\mathbb{R}} X(\omega_1) X(\omega-\omega_1)\, H_2(\omega_1, \omega-\omega_1)\,\mathrm{d}\omega_1 && (2.55)\\ &+ \frac{1}{(2\pi)^2}\int_{\mathbb{R}^2} X(\omega_1)X(\omega_2)X(\omega-\omega_1-\omega_2)\, H_3(\omega_1,\omega_2,\omega-\omega_1-\omega_2)\,\mathrm{d}\omega_1\mathrm{d}\omega_2 && (2.56)\\ &+ \ldots \\ = h_0\,\delta(\omega) &+ \sum_{n=1}^{\infty} \mathsf{Y}_n(\omega)\,, && (2.57) \end{aligned}$$

where, due to the linearity of the Fourier transform, each series term, i.e., each Volterra operator is related by a 1D Fourier transform $\mathsf{y}_n(t) \circ\!\!-\!\!\bullet\, \mathsf{Y}_n(\omega)$. The general form of $\mathsf{Y}_n(\omega)$ is given by [Sch80, Ch. 6]

$$\mathsf{Y}_n(\omega) \stackrel{\text{def}}{=} \frac{1}{(2\pi)^{n-1}} \int_{\mathbb{R}^{n-1}} H_n(\boldsymbol{\omega}) \prod_{j=1}^{n} X(\omega_j)\, \mathrm{d}^{n-1}\boldsymbol{\omega}\,, \tag{2.58}$$

with the frequency-domain Volterra kernel (i.e., a generalized transfer function of order n) $H_n(\boldsymbol{\omega}) = H_n(\omega_1, \omega_2, \ldots, \omega_n) \in \mathbb{C}$. Here, the last component of $\boldsymbol{\omega}$ (depending on the order n) is defined as

$$\omega_n \stackrel{\text{def}}{=} \omega - \omega_1 - \cdots - \omega_{n-1}, \quad \text{with} \quad \omega_1 = \omega, \text{ for } n = 1\,. \tag{2.59}$$

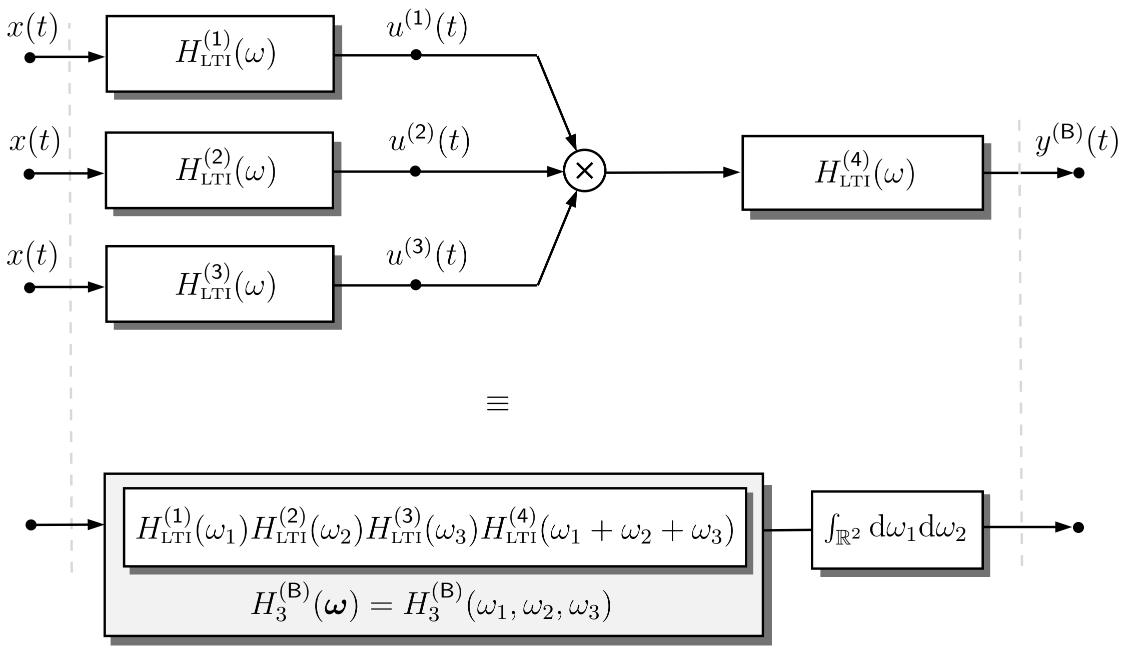


Figure 2.1: Block diagram of a *basic third-order Volterra system* with three (different) linear, dispersive systems at the input with transfer function $H^{(\mathsf{i})}_{\mathrm{LTI}}(\omega)$, and $i = 1, 2, 3$, and a single linear, dispersive system at the output with transfer function $H^{(4)}_{\mathrm{LTI}}(\omega)$. The equivalent block diagram at the bottom highlights the Volterra view in frequency domain implementing a two-fold convolution of the input transform with the Volterra kernel $H_3^{(\mathsf{B})}(\boldsymbol{\omega})$.

In fiber-optical transmission, this relation with $n = 3$ will translate to the frequency matching constraint for four-wave mixing, see Section 3.3.5.

The Volterra kernel in time and frequency domain are related by an n-dimensional Fourier transform [Sch80, Eq. (6.1-4)]

$$H_n(\boldsymbol{\omega}) = \mathcal{F}_{\boldsymbol{\omega}\leftrightarrow\boldsymbol{\tau}}\{\, h_n(\boldsymbol{\tau})\,\}\,. \tag{2.60}$$

From a symmetric time-domain kernel, a symmetric frequency-domain kernel follows [Sch80, P. 118]

$$h_n(\boldsymbol{\tau}) = h_n(\pi_i(\boldsymbol{\tau})), \quad \forall i \quad \Rightarrow \quad H_n(\boldsymbol{\omega}) = H_n(\pi_i(\boldsymbol{\omega})), \quad \forall i\,, \tag{2.61}$$

with $i \in \{\,1, 2, \ldots n!-1\,\}$. For a real-valued time-domain kernel, the frequency-domain kernel is conjugate symmetric with

$$h_n(\boldsymbol{\tau}) \in \mathbb{R} \quad \Rightarrow \quad H_n(\boldsymbol{\omega}) = H_n^*(-\boldsymbol{\omega})\,. \tag{2.62}$$

Basic Nonlinear Systems We consider a *basic third-order Volterra system* [Sch80, Fig. 6.4-1] which is of interest to nonlinear fiber propagation. The block diagram is shown in Figure 2.1.

It consists of a nonlinear, memoryless system with cubic input/output relation, sandwiched between four (potentially) different LTI systems characterized by $h^{(\mathsf{i})}_{\mathrm{LTI}}(t) \circ\!\!-\!\!\bullet\ H^{(\mathsf{i})}_{\mathrm{LTI}}(\omega)$ with $i = 1, 2, 3, 4$.

Following the third-order Volterra approach, see above, the output of the basic third-order

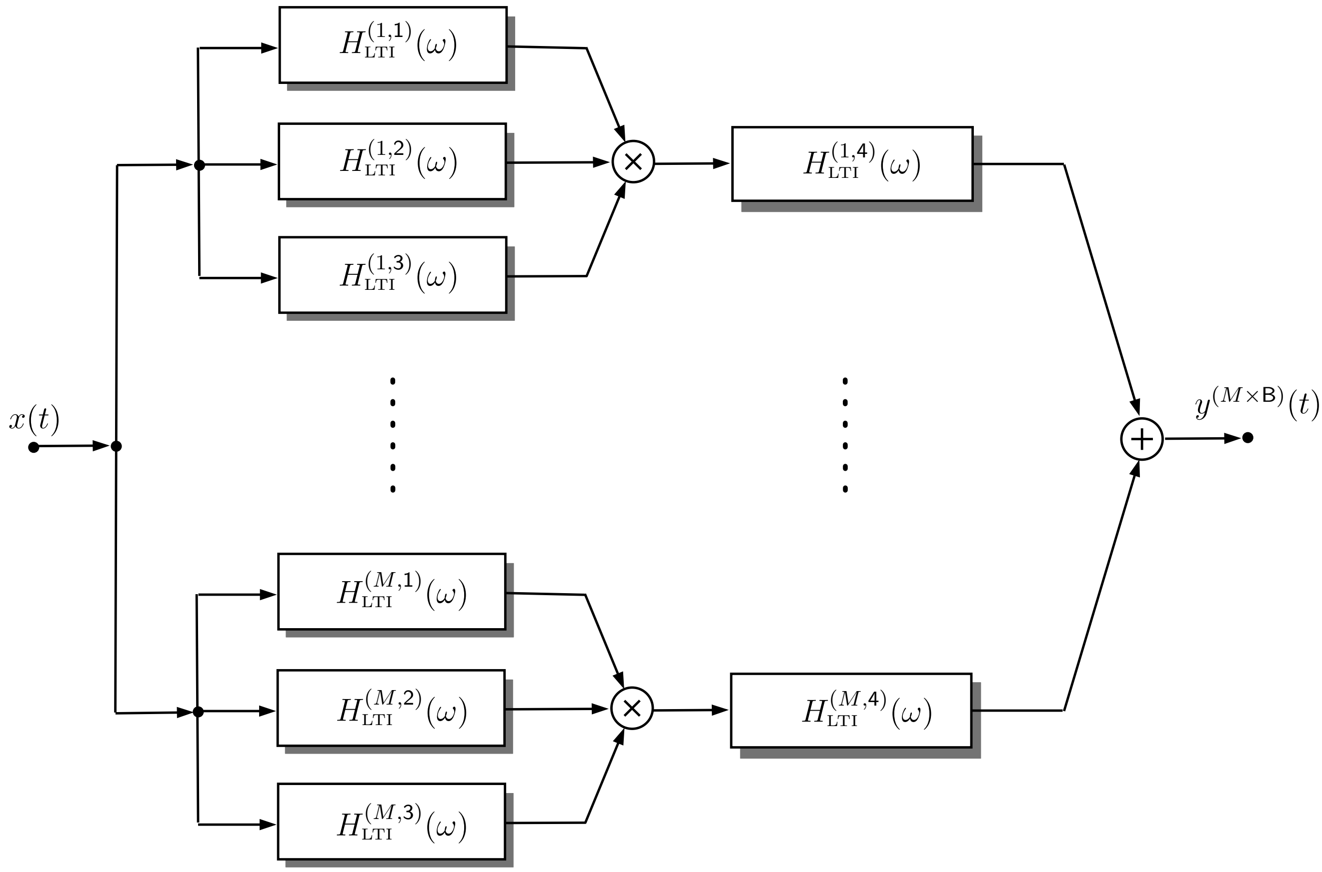


Figure 2.2: Block diagram of M basic third-order Volterra systems in parallel.

system is given by a closed-form expression in both time and frequency domain

$$y^{(\mathsf{B})}(t) = \int_{\mathbb{R}^3} x(t-\tau_1)x(t-\tau_2)x(t-\tau_3)\, h_3^{(\mathsf{B})}(\tau_1,\tau_2,\tau_3)\,\mathrm{d}\tau_1\mathrm{d}\tau_2\mathrm{d}\tau_3 \tag{2.63}$$

$$\circ\!\!-\!\!\bullet$$

$$Y^{(\mathsf{B})}(\omega) = \frac{1}{(2\pi)^2}\int_{\mathbb{R}^2} X(\omega_1)X(\omega_2)X(\underbrace{\omega-\omega_1-\omega_2}_{\omega_3})\, H_3^{(\mathsf{B})}(\omega_1,\omega_2,\underbrace{\omega-\omega_1-\omega_2}_{\omega_3})\,\mathrm{d}\omega_1\mathrm{d}\omega_2\,, \tag{2.64}$$

with the (frequency-matching) constraint $\omega = \omega_1+\omega_2+\omega_3$ from (2.59).

The time- and frequency-domain third-order kernel can be given explicitly by [Sch80, Eq. (4.4-3), (6.4-2)]

$$h_3^{(\mathsf{B})}(\boldsymbol{\tau}) = \int_{\mathbb{R}} h_{\mathrm{LTI}}^{(1)}(\tau_1-\sigma)\, h_{\mathrm{LTI}}^{(2)}(\tau_2-\sigma)\, h_{\mathrm{LTI}}^{(3)}(\tau_3-\sigma)\, h_{\mathrm{LTI}}^{(4)}(\sigma)\,\mathrm{d}\sigma \tag{2.65}$$

$$\circ\!\!-\!\!\bullet\ \mathcal{F}_{\boldsymbol{\tau}\leftrightarrow\boldsymbol{\omega}}$$

$$H_3^{(\mathsf{B})}(\boldsymbol{\omega}) = H_{\mathrm{LTI}}^{(1)}(\omega_1)\, H_{\mathrm{LTI}}^{(2)}(\omega_2)\, H_{\mathrm{LTI}}^{(3)}(\omega_3)\, H_{\mathrm{LTI}}^{(4)}(\underbrace{\omega_1+\omega_2+\omega_3}_{\omega})\,. \tag{2.66}$$

Note, that the system function $H_{\mathrm{LTI}}^{(4)}(\cdot)$ with argument $(\omega_1+\omega_2+\omega_3)$ is the one following the triple multiplier, i.e., at the output of the combined system, whereas the system functions at the input have arguments involving either ω_1, ω_2, or ω_3 alone.

One can also deduce, that both time- and frequency-domain kernel become symmetric, if the input LTI systems are identical, i.e., we have

$$h_{\mathrm{LTI}}^{(1)}(t) = h_{\mathrm{LTI}}^{(2)}(t) = h_{\mathrm{LTI}}^{(3)}(t) \quad \Rightarrow \quad h_3^{(\mathsf{B})}(\boldsymbol{\tau}) = h_3^{(\mathsf{B})}(\pi_i(\boldsymbol{\tau})), \quad \forall i\,. \tag{2.67}$$

We now consider a third-order Volterra system consisting of M basic systems connected in parallel. The block diagram of such a general third-order Volterra system is shown in Figure 2.2. The output $y^{(M\times\mathsf{B})}(t)$ is simply the sum of the output of each basic system. Consequently, the kernel of the parallel system is the sum of the M basic third-order kernels [Sch80, Eq. (4.4-10)]

$$h_3^{(M\times\mathsf{B})}(\boldsymbol{\tau}) = \sum_{m=1}^{M} \int_{\mathbb{R}} h_{\mathrm{LTI}}^{(m,1)}(\tau_1 - \sigma)\, h_{\mathrm{LTI}}^{(m,2)}(\tau_2 - \sigma)\, h_{\mathrm{LTI}}^{(m,3)}(\tau_3 - \sigma)\, h_{\mathrm{LTI}}^{(m,4)}(\sigma)\, \mathrm{d}\sigma \tag{2.68}$$

$$\circ\!\!-\!\!\bullet\ \mathcal{F}_{\boldsymbol{\tau}\leftrightarrow\boldsymbol{\omega}}$$

$$H_3^{(M\times\mathsf{B})}(\boldsymbol{\omega}) = \sum_{m=1}^{M} H_{\mathrm{LTI}}^{(m,1)}(\omega_1)\, H_{\mathrm{LTI}}^{(m,2)}(\omega_2)\, H_{\mathrm{LTI}}^{(m,3)}(\omega_3)\, H_{\mathrm{LTI}}^{(m,4)}\underbrace{(\omega_1 + \omega_2 + \omega_3)}_{\omega}\,. \tag{2.69}$$

This view will later become relevant when the so-termed *parallel* fiber model is considered. In that case, the summation in (2.68)–(2.69) over a *discrete* number of parallel systems will asymptotically turn into an integral over a *continuous* variable, i.e., infinitely many parallel systems.

Examples of Third-order Volterra Systems In the following, we discuss two examples of a basic third-order nonlinear system. To that end, we introduce two simple realizations of a nonlinear system—the *Hammerstein* and the *Wiener* system, shown in Figure 2.3.

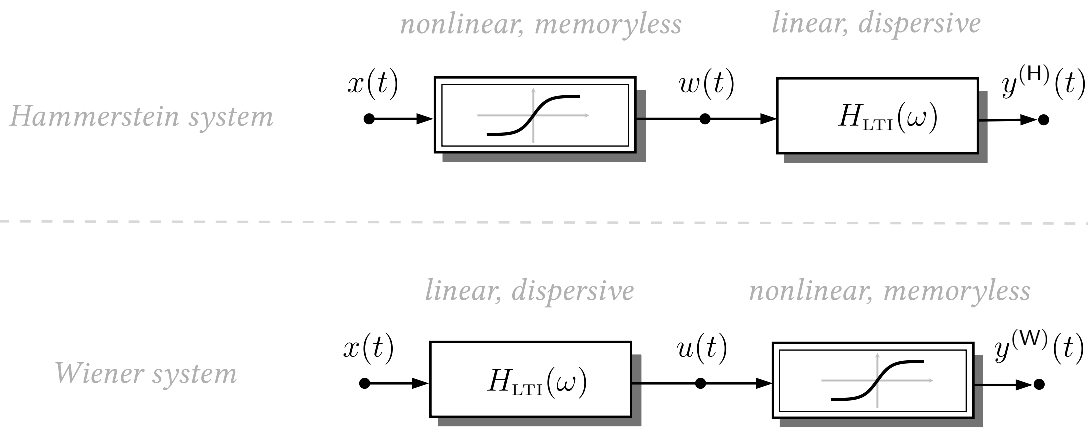


Figure 2.3: Block diagram of a Hammerstein (top) and Wiener system (bottom). A Hammerstein system is a nonlinear, memoryless mapping at the input, in series with a linear, dispersive system characterized by its transfer function $H_{\mathrm{LTI}}(\omega)$. The Wiener system is a reverse-concatenated Hammerstein system.

The Hammerstein system is a concatenation of a nonlinear, memoryless transfer characteristic and a linear, dispersive system. The Wiener system reverses the concatenation so that the linear system is at the input and the nonlinear part at the system output. Both systems can be combined, e.g., the *Hammerstein-Wiener* system consists of a linear, dispersive system sandwiched between two nonlinear, memoryless systems.

Example 2.1: A simple 3^{rd}-order Hammerstein system

We consider the following pair of nonlinear differential equations [Boy85, Ch. 1]

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}u(t) &= -u(t) + x^3(t)\\ y^{(\mathsf{H})}(t) &= u(t)\,,\end{aligned}$$

which relates the system input $x(t)$ to the system output $y^{(\mathsf{H})}(t)$ of the third-order Hammerstein system including an intermediate variable $u(t)$. We assume a causal signal $x(t)$ with $x(t) = 0$ for $t \leq 0$.

A block diagram of the system is shown in Figure 2.4. It consists of the cubic function

$$w(t) = x^3(t)\,,$$

at the input and an LTI system at the output with transfer function

$$\begin{aligned}H_{\mathrm{LTI}}(\omega) = H_{\mathrm{LTI},\mathscr{L}}(s)\big|_{s=\mathrm{j}\omega} = \frac{1}{1+s}\bigg|_{s=\mathrm{j}\omega} &= \frac{1}{1+\mathrm{j}\omega}\\ \bullet\!\!-\!\!\circ\quad & \\ h_{\mathrm{LTI}}(t) = \varepsilon(t)\,\mathrm{e}^{-t}\,. &\end{aligned}$$

Here, $H_{\mathrm{LTI}}(\omega)$ is a first-order low pass with a single pole at $s_\infty = -1$.

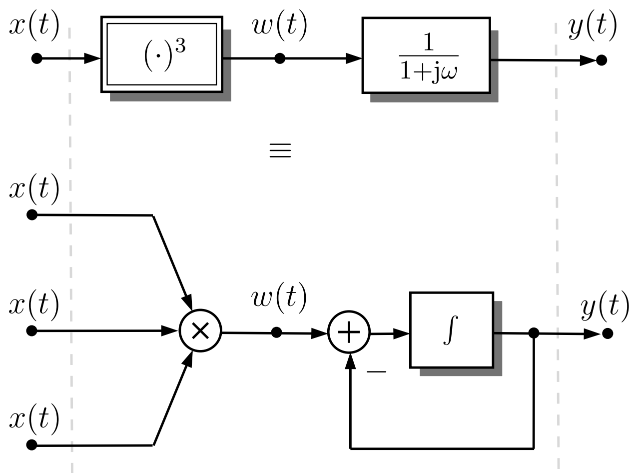


Figure 2.4: Block diagram of the third-order Hammerstein system from Example 2.1. The top shows the block diagram representation of the nonlinear mapping $(\cdot)^3$ and the LTI system characterized by its transfer function $H_{\mathrm{LTI}}(\omega)$. The bottom shows the equivalent representation using elementary building blocks.

In time domain, we obtain the solution of the differential equations by using the method of integrating factors for $t > 0$ as

$$\begin{aligned}y^{(\mathsf{H})}(t) &= \int_0^\infty \mathrm{e}^{-\tau} x^3(t-\tau)\,\mathrm{d}\tau\\ &= \int_{\mathbb{R}^3} \varepsilon(\tau_1)\varepsilon(\tau_2)\varepsilon(\tau_3)\delta(\tau_1-\tau_2)\delta(\tau_2-\tau_3)\,\mathrm{e}^{-\tau_1}\,x(t-\tau_1)x(t-\tau_2)x(t-\tau_3)\;\mathrm{d}^3\boldsymbol{\tau}\,,\end{aligned}$$

where the Heaviside step function $\varepsilon(t)$ and the Dirac impulse $\delta(t)$ are used to bring the right-hand side into a form similar to the third-order time-domain Volterra representation in (2.45). Comparing the result with (2.45), we find the 3^{rd}-order Volterra kernel as

$$\begin{aligned} h_3^{(\mathsf{H})}(\boldsymbol{\tau}) = h_3^{(\mathsf{H})}(\tau_1,\tau_2,\tau_3) &= \varepsilon(\tau_1)\varepsilon(\tau_2)\varepsilon(\tau_3)\delta(\tau_1-\tau_2)\delta(\tau_2-\tau_3)\,\mathrm{e}^{-\tau_1} \\ &= h_{\mathrm{LTI}}(\tau_1)\,\delta(\tau_1-\tau_2)\delta(\tau_2-\tau_3)\,, \end{aligned}$$

which has support only for $\tau_1=\tau_2$, $\tau_2=\tau_3$, and $\tau_1,\tau_2,\tau_3\geq 0$, and is hence causal.
Alternatively, the kernel can also be obtained using the general solution from (2.65) resulting in

$$\begin{aligned} h_3^{(\mathsf{H})}(\boldsymbol{\tau}) &= \int_{\mathbb{R}} h_{\mathrm{LTI}}^{(1)}(\tau_1-\sigma)\,h_{\mathrm{LTI}}^{(2)}(\tau_2-\sigma)\,h_{\mathrm{LTI}}^{(3)}(\tau_3-\sigma)\,h_{\mathrm{LTI}}^{(4)}(\sigma)\,\mathrm{d}\sigma \\ &= \int_{\mathbb{R}} \delta(\tau_1-\sigma)\delta(\tau_2-\sigma)\delta(\tau_3-\sigma)\,\varepsilon(\sigma)\mathrm{e}^{-\sigma}\,\mathrm{d}\sigma \\ &= h_{\mathrm{LTI}}(\tau_1)\,\delta(\tau_1-\tau_2)\delta(\tau_2-\tau_3)\,, \end{aligned}$$

with $h_{\mathrm{LTI}}^{(1)}(t)=h_{\mathrm{LTI}}^{(2)}(t)=h_{\mathrm{LTI}}^{(3)}(t)=\delta(t)$ and $h_{\mathrm{LTI}}^{(4)}(t)=\varepsilon(t)\,\mathrm{e}^{-t}$.
We use the Fourier relation between time- and frequency-domain kernels from (2.60) to obtain the 3^{rd}-order Volterra kernel in frequency domain as

$$H_3^{(\mathsf{H})}(\boldsymbol{\omega}) = H_3^{(\mathsf{H})}(\omega_1,\omega_2,\omega_3) = \mathcal{F}_{\boldsymbol{\omega}\leftrightarrow\boldsymbol{\tau}}\{\,h_3^{(\mathsf{H})}(\boldsymbol{\tau})\,\} = \frac{1}{1+\mathrm{j}(\omega_1+\omega_2+\omega_3)}\,,$$

by taking a 3D-Fourier transform.
Similarly, the same result can also be obtained using the general solution in frequency domain from (2.66) to arrive at

$$\begin{aligned} H_3^{(\mathsf{H})}(\boldsymbol{\omega}) &= H_{\mathrm{LTI}}^{(1)}(\omega_1)\,H_{\mathrm{LTI}}^{(2)}(\omega_2)\,H_{\mathrm{LTI}}^{(3)}(\omega_3)\,H_{\mathrm{LTI}}^{(4)}(\omega_1+\omega_2+\omega_3) \\ &= \frac{1}{1+\mathrm{j}(\omega_1+\omega_2+\omega_3)}\,, \end{aligned}$$

with $H_{\mathrm{LTI}}^{(1)}(\omega)=H_{\mathrm{LTI}}^{(2)}(\omega)=H_{\mathrm{LTI}}^{(3)}(\omega)=1$ and $H_{\mathrm{LTI}}^{(4)}(\omega)=\frac{1}{1+\mathrm{j}\omega}$.
Next, we substitute ω_3 using the frequency constraint $\omega_3=\omega-\omega_1-\omega_2$ to arrive at

$$H_3^{(\mathsf{H})}(\omega_1,\omega_2,\omega-\omega_1-\omega_2) = \frac{1}{1+\mathrm{j}\omega}\,,$$

We can now give the final result of the frequency-domain input/output relation as

$$\begin{aligned} Y^{(\mathsf{H})}(\omega) &= \frac{1}{(2\pi)^2}\int_{\mathbb{R}^2} X(\omega_1)X(\omega_2)X(\omega-\omega_1-\omega_2)\,H_3^{(\mathsf{H})}(\omega_1,\omega_2,\omega-\omega_1-\omega_2)\,\mathrm{d}^2\boldsymbol{\omega} \\ &= \frac{1}{1+\mathrm{j}\omega}\,\frac{1}{(2\pi)^2}\int_{\mathbb{R}^2} X(\omega_1)X(\omega_2)X(\omega-\omega_1-\omega_2)\,\mathrm{d}^2\boldsymbol{\omega} \\ &= \frac{1}{1+\mathrm{j}\omega}\,\left(X(\omega)*X(\omega)*X(\omega)\right). \end{aligned}$$

Note, that the third-order Hammerstein kernel is symmetric, and since it does not depend on ω_1 and ω_2, the kernel multiplication and the two-fold convolution are separable.

Example 2.2: A simple 3^{rd}-order Wiener system

For the Wiener system, we consider the following set of differential equations

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}u(t) &= -u(t) + x(t)\\ y^{(\mathsf{W})}(t) &= u^3(t)\,.\end{aligned}$$

where the order of the nonlinearity $(\cdot)^3$ and the LTI system with impulse response $h_{\mathrm{LTI}}(t)$ is exchanged w.r.t. the Hammerstein system from our last example. Again, with a causal signal $x(t)$ and $x(0) = 0$, the solution is given by

$$\begin{aligned}y^{(\mathsf{W})}(t) &= \left(\int_0^\infty \mathrm{e}^{-\tau} x(t-\tau)\,\mathrm{d}\tau\right)^3\\ &= \int_{\mathbb{R}^3} \varepsilon(\tau_1)\varepsilon(\tau_2)\varepsilon(\tau_3)\mathrm{e}^{-(\tau_1+\tau_2+\tau_3)} x(t-\tau_1)x(t-\tau_2)x(t-\tau_3)\ \mathrm{d}^3\boldsymbol{\tau}\,,\end{aligned}$$

where the right-hand side is written in a form similar to (2.45). The Volterra kernel in time domain directly follows as

$$h_3^{(\mathsf{W})}(\boldsymbol{\tau}) = \varepsilon(\tau_1)\varepsilon(\tau_2)\varepsilon(\tau_3)\mathrm{e}^{-(\tau_1+\tau_2+\tau_3)} = h_{\mathrm{LTI}}(\tau_1)h_{\mathrm{LTI}}(\tau_2)h_{\mathrm{LTI}}(\tau_3)\,.$$

Alternatively, the kernel is derived using (2.65) to find

$$\begin{aligned}h_3^{(\mathsf{W})}(\boldsymbol{\tau}) &= \int_{\mathbb{R}} h_{\mathrm{LTI}}^{(1)}(\tau_1-\sigma)\,h_{\mathrm{LTI}}^{(2)}(\tau_2-\sigma)\,h_{\mathrm{LTI}}^{(3)}(\tau_3-\sigma)\,h_{\mathrm{LTI}}^{(4)}(\sigma)\ \mathrm{d}\sigma\\ &= \int_{\mathbb{R}} \varepsilon(\tau_1-\sigma)\mathrm{e}^{-(\tau_1-\sigma)}\varepsilon(\tau_2-\sigma)\mathrm{e}^{-(\tau_2-\sigma)}\varepsilon(\tau_3-\sigma)\mathrm{e}^{-(\tau_3-\sigma)}\ \delta(\sigma)\ \mathrm{d}\sigma\\ &= h_{\mathrm{LTI}}(\tau_1)h_{\mathrm{LTI}}(\tau_2)h_{\mathrm{LTI}}(\tau_3)\,,\end{aligned}$$

with $h_{\mathrm{LTI}}^{(1)}(t) = h_{\mathrm{LTI}}^{(2)}(t) = h_{\mathrm{LTI}}^{(3)}(t) = \varepsilon(t)\mathrm{e}^{-t}$ and $h_{\mathrm{LTI}}^{(4)}(t) = \delta(t)$. In contrast to the Hammerstein system, the Wiener kernel has support for any $\tau_1, \tau_2, \tau_3 \geq 0$, and is also causal.
A 3D-Fourier transform gives the 3^{rd}-order Volterra kernel in continuous-frequency as

$$H_3^{(\mathsf{W})}(\boldsymbol{\omega}) = \mathcal{F}_{\boldsymbol{\omega}\leftrightarrow\boldsymbol{\tau}}\{\,h_3^{(\mathsf{W})}(\boldsymbol{\tau})\,\} = \frac{1}{(1+\mathrm{j}\omega_1)\,(1+\mathrm{j}\omega_2)\,(1+\mathrm{j}\omega_3)}\,.$$

or using again the general formula from (2.66) to find

$$\begin{aligned}H_3^{(\mathsf{W})}(\boldsymbol{\omega}) &= H_{\mathrm{LTI}}^{(1)}(\omega_1)\,H_{\mathrm{LTI}}^{(2)}(\omega_2)\,H_{\mathrm{LTI}}^{(3)}(\omega_3)\,H_{\mathrm{LTI}}^{(4)}(\omega_1+\omega_2+\omega_3)\\ &= \frac{1}{(1+\mathrm{j}\omega_1)\,(1+\mathrm{j}\omega_2)\,(1+\mathrm{j}\omega_3)}\,,\end{aligned}$$

with $H_{\mathrm{LTI}}^{(1)}(\omega) = H_{\mathrm{LTI}}^{(2)}(\omega) = H_{\mathrm{LTI}}^{(3)}(\omega) = \frac{1}{1+\mathrm{j}\omega}$ and $H_{\mathrm{LTI}}^{(4)}(\omega) = 1$.
The frequency matching constraint $\omega_3 = \omega - \omega_1 - \omega_2$ is used again to rewrite the kernel

$$H_3^{(\mathsf{W})}(\omega_1, \omega_2, \omega-\omega_1-\omega_2) = \frac{1}{1+\mathrm{j}\omega_1}\ \frac{1}{1+\mathrm{j}\omega_2}\ \frac{1}{1+\mathrm{j}(\omega-\omega_1-\omega_2)}\,,$$

We obtain the final frequency-domain input/output relation of the third-order Wiener system with

$$
\begin{aligned}
Y^{(\mathsf{W})}(\omega) &= \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} X(\omega_1) X(\omega_2) X(\omega - \omega_1 - \omega_2)\, H_2^{(\mathsf{W})}(\omega_1, \omega_2, \omega - \omega_1 - \omega_2)\, \mathrm{d}^2\boldsymbol{\omega} \\
&= \left(\frac{X(\omega)}{1 + \mathrm{j}\omega}\right) * \left(\frac{X(\omega)}{1 + \mathrm{j}\omega}\right) * \left(\frac{X(\omega)}{1 + \mathrm{j}\omega}\right) \\
&= U(\omega) * U(\omega) * U(\omega)\,,
\end{aligned}
$$

where we used that the frequency-domain kernel itself is symmetric and factorable, i.e., we have

$$H_3^{(\mathsf{W})}(\boldsymbol{\omega}) = H_3^{(\mathsf{W})}(\omega_1, \omega_2, \omega_3) = H_{\mathrm{LTI}}(\omega_1)\, H_{\mathrm{LTI}}(\omega_2)\, H_{\mathrm{LTI}}(\omega_3)\,,$$

and the same is also true for the time-domain kernel

$$h_3^{(\mathsf{W})}(\boldsymbol{\tau}) = h_3^{(\mathsf{W})}(\tau_1, \tau_2, \tau_3) = h_{\mathrm{LTI}}(\tau_1)\, h_{\mathrm{LTI}}(\tau_2)\, h_{\mathrm{LTI}}(\tau_3)\,.$$

2.2 The Linear System Model

In this work we consider *point-to-point* coherent optical transmission over two planes of polarization in a single-mode fiber. To this end, the analog transmit and receive signals, $\boldsymbol{s}(t)$ and $\boldsymbol{r}(t)$, are modeled in the two-dimensional (2D) complex plane $\mathbb{C}^2$, commonly known in optics as the *Jones space* [Jon41], to account for the two states of polarization. A generic block diagram of such a point-to-point communication system is shown in Figure 2.5.

The transmission system is fed with equiprobable source bits. The binary source generates uniform i.i.d. *information bits* $\mathfrak{q}[\kappa] \in \mathbb{F}_2$ at each discrete-time index $\kappa \in \mathbb{Z}$. The rate at which the source generates information bits is called *information rate* and is defined as the inverse temporal interval between two successive information bits; hence, the data (information bit) rate is defined as $R_\mathrm{T} \stackrel{\text{def}}{=} 1/T_\mathrm{b}$. The transmitter translates the *discrete-time* source sequence into a modulated *continuous-time* transmit signal $\boldsymbol{s}(t) \in \mathbb{C}^2$, later associated with the two polarizations of the *optical* signal. After transmission over the channel, the receive signal $\boldsymbol{r}(t) \in \mathbb{C}^2$ is processed by the receiver to obtain the recovered bit sequence $\hat{\mathfrak{q}}[\kappa] \in \mathbb{F}_2$. Typically, the receiver employs techniques from signal equalization and error-correction coding (i.e., channel coding) to recover the exact source bit sequence.

In the course of this work, we will establish a framework to determine the distortions originating from nonlinear interaction of the optical signal with itself (i.e., *signal-signal* nonlinear interference (NLI)[2]) while propagating along the fiber (channel). However, signal equalization as part of a state-of-the-art *digital coherent receiver* [Sav10] (e.g., polarization demultiplexing, channel estimation and equalization, carrier- and phase recovery, etc.) as well as error correction coding (i.e., forward error correction (FEC)) will not be considered explicitly. Instead, we assume ideal channel state information (CSI) such that polarization demultiplexing and linear equalization can be incorporated into the end-to-end channel description as part of the end-to-end system description.

In the remainder of this chapter we discuss the relevant building blocks of the transmitter and receiver front-end necessary to establish a *linear* point-to-point communication system. We will particularly highlight the transitions between the discrete-time and continuous-time domain and discuss how *aliasing* of frequency components is typically modeled in discrete-time end-to-end models for linear channels. Most of the consideration and the notational framework for linear channels can also be found in the text book [Fis02, Ch. 2].

2.2.1 Transmitter Front-end

In the optical community, the term *front-end* is often associated with the electrical (and/or optical) interface at the transmitter or receiver, e.g., the digital-to-analog converter (DAC) and analog-to-digital converter (ADC) (possibly, plus additional optical components). In the present work, the front-end characterizes the continuous-time (i.e., analog) part of the trans-

[2]The denomination *distortion*, *disturbance*, and *interference* will be used synonymously. In the literature, the term *nonlinear interference noise* is also used since the source of the nonlinear interference is often another out-of-band modulated signal, i.e., non-accessible and hence *random* to the channel of interest.

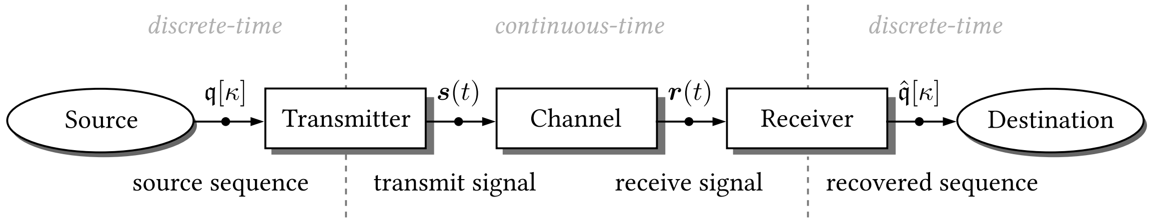


Figure 2.5: Block diagram of a *point-to-point* communication system. The transmitter and receiver constitute the interface between the discrete- and continuous-time domain.

mitter and receiver. In practice, the analog front-end (using our terminology) is typically realized as a composition of the real *physical* interface and some oversampled *digital* processing to obtain a desired target transfer characteristic. Accordingly, we prefer to conceptually incorporate some of the building blocks, often realized as part of the DSP chain (such as *pulse-shaping* and/or *matched filtering*), as part of the continuous-time domain, since this significantly simplifies the notation[3]. Similarly, we would consider, e.g., the low-pass characteristic of the DAC or ADC as part of the *transmission channel*.

In Figure 2.6, a block diagram of an uncoded (i.e., without error correction coding) communication system is shown. In coherent optical transmission, two (independent) planes of polarization are present, each with inphase and quadrature component in the ECB. Then, transmission over two polarizations (x and y) with (polarization-) *diversity reception* results in a complex-valued 2×2 MIMO transmission which is typically used for multiplexing. Within the modulated bandwidth of a single transmitter/receiver pair, we can consider the *optical end-to-end* MIMO channel as frequency-flat if we neglect the effects of bandlimiting devices (e.g., switching elements in a routed network). The topic of polarization demultiplexing and equalization of, e.g., polarization-mode dispersion (PMD) (i.e., equalization of the dispersive MIMO channel) is discussed in numerous research papers (c.f. [Sav10] for a comprehensive overview). We will not give an exhaustive description of it, but will rather concentrate on basic concepts. Hence, for the intent of this work, aspects related to *linear* equalization and timing synchronization are already incorporated into the optical channel as suited *linear* transfer characteristics. E.g., see the following chapter, the average group delay of the probe channel will be canceled from the propagation equation using the concept of a *retarded time frame* in the equivalent complex baseband. This is equivalent to a linear phase response, which can also be applied as part of the receiver DSP.

For coherent optical transmission the natural choice of *real* signal dimensions D (i.e., the dimensionality of the *signal space*) is *four* because each polarization offers two independent degrees of freedom, cf. [Kam13, P. 831]. The evolution of the *optical field* is typically described using a *complex-valued* baseband representation via a set of coupled differential equations—the so-called *Manakov equation*, cf. Section 3.3. Therefore, we treat data symbols as a pair of complex-valued symbols corresponding to the polarizations x and y. If channel coding is present, the signal space depends on the coded-modulation strategy and is possible dif-

[3]When obeying the sampling theorem, both discrete- and continuous-time representations are equivalent when considering linear systems.

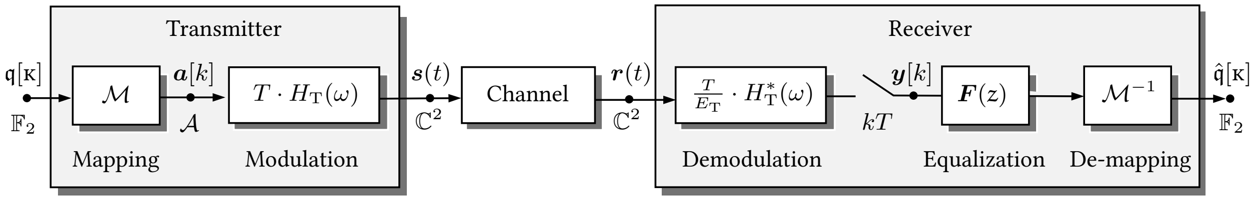


Figure 2.6: Block diagram of an uncoded, complex-valued 2×2 MIMO communication system in the *equivalent complex baseband* (ECB) representation. In the transmitter, binary source symbols $\mathfrak{q}[\kappa] \in \mathbb{F}_2$ are mapped ($\mathcal{M}$) to the data symbols $\boldsymbol{a}[k]$ taken from the signal constellation $\mathcal{A}$. Then, the data symbols are modulated via the transmit pulse $H_\mathrm{T}(\omega)$ to the ECB transmit signal $\boldsymbol{s}(t) \in \mathbb{C}^2$. On the receive-side, the cascade is passed in reverse order. The receive signal $\boldsymbol{r}(t) \in \mathbb{C}^2$ is demodulated using the (matched filter) receiver frond-end $H_\mathrm{R}(\omega) = \frac{T}{E_\mathrm{T}} \cdot H^*_\mathrm{T}(\omega)$ and sampled at $t = kT$ to obtain the discrete-time receive symbols $\boldsymbol{y}[k]$. Residual intersymbol interference induced by the channel is equalized by the discrete-time part $\boldsymbol{F}(z) \in \mathbb{C}^{2\times 2}$. The processed receive symbols are de-mapped (i.e., inverse mapping) to recover an estimate of the source sequence $\hat{\mathfrak{q}}[\kappa]$.

ferent from the one used to describe signal evolution. Throughout the thesis, we use four-dimensional (4D) *modulation formats* ($D = 4$). Important digital modulation formats in the 2D signal space ($D = 2$) such as quadrature amplitude modulation (QAM) will be included as special cases if we consider independent modulation in two polarizations, namely polarization-division multiplex (PDM).

Bit-to-Symbol Mapping The binary source sequence $\langle\, \mathfrak{q}[\kappa] \,\rangle$ is partitioned into binary tuples of length R_m, such that

$$\boldsymbol{\mathfrak{q}}[k] = [\, \mathfrak{q}_1[k], \mathfrak{q}_2[k], \ldots, \mathfrak{q}_{R_\mathrm{m}}[k] \,]^\mathsf{T} \in \{0, 1\}^{R_\mathrm{m}} , \tag{2.70}$$

where $k \in \mathbb{Z}$ is the discrete-time index of the *data symbols* $\boldsymbol{a}[k]$. Here, R_m is called the rate of the modulation and will be equivalent to the number of bits per transmitted data symbol (neglecting channel coding and assuming that the size of the symbol set is a power of two). Each R_m-tuple is associated with one of the possible data symbols $\boldsymbol{a} = [a_1, a_2]^\mathsf{T} = [a_\mathsf{x}, a_\mathsf{y}]^\mathsf{T} \in \mathcal{A} \subset \mathbb{C}^2$, i.e., with one of the *constellation* points. In other words, the binary R_m-tuples are mapped to the data symbols $\boldsymbol{a} \in \mathcal{A}$ by a bijective mapping rule $\mathcal{M} : \boldsymbol{\mathfrak{q}} \mapsto \boldsymbol{a}$.

The size of the data symbol set is $M \overset{\text{def}}{=} |\mathcal{A}| = 2^{R_\mathrm{m}}$ and we can write the alphabet as set

$$\mathcal{A} \overset{\text{def}}{=} \{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_M\} \subset \mathbb{C}^2 . \tag{2.71}$$

The symbol set is zero mean if not stated otherwise, that is $\mathrm{E}\{\, \boldsymbol{a} \,\} \overset{!}{=} \mathbf{0}$, and we deliberately normalize, without loss of generality, the variance of the symbol set to unity

$$\sigma_a^2 \overset{\text{def}}{=} \mathrm{E}\{\, \|\boldsymbol{a} - \mathrm{E}\{\, \boldsymbol{a} \,\}\|^2 \,\} = \mathrm{E}\{\, \|\boldsymbol{a}\|^2 \,\} \overset{!}{=} 1 , \tag{2.72}$$

where expectation is denoted by $\mathrm{E}\{\, \cdot \,\}$ and the Euclidean vector norm is $\|\boldsymbol{a}\|^2 = a_\mathsf{x} a_\mathsf{x}^* + a_\mathsf{y} a_\mathsf{y}^*$. If we consider WDM, cf. Section 3.2, we denote the data symbols of interfering *wavelength channels*[4] by $\boldsymbol{b}_\nu[k]$ where the integer subscript $\nu \in \{\, 1, 2, \ldots, N_\mathrm{ch} \,\}$ is the wavelength channel number, and N_ch is the number of wavelength channels.

[4]The term *channel* can refer to the physical (fiber) channel, i.e., the transmission medium, or in the context of WDM transmission to a wavelength channel, i.e., the transmitter/receiver pair at different center wavelengths/frequencies.

If *signal shaping* [Fis02] is not present and $\langle\, \mathfrak{q}[\kappa] \,\rangle$ is i.i.d. uniform, all data symbols occur with the same probability $\Pr\{\, \boldsymbol{a}_m \,\} = 1/M,\ \forall\, m$ and (2.72) simplifies to

$$\sigma_a^2 = \frac{1}{M} \sum_{m=1}^{M} \|\boldsymbol{a}_m\|^2 \stackrel{!}{=} 1\,, \tag{2.73}$$

where the variance of the sequence $\langle\, \boldsymbol{a}[k] \,\rangle$ is equal to the variance of the constellation $\mathcal{A}$.

Example 2.3: PDM 16-ary QAM

Transmission of PDM 16-ary QAM with a 4D cardinality $M = 256$ is considered. The two polarizations are transmitted independently, i.e., tuples of $R_\mathrm{m} = \log_2(M) = 8$ bits are mapped into two 16-QAM sets using a binary-reflected Gray labeling (BRGL) rule for each set [SA15, Fig. 2.14 (b)].

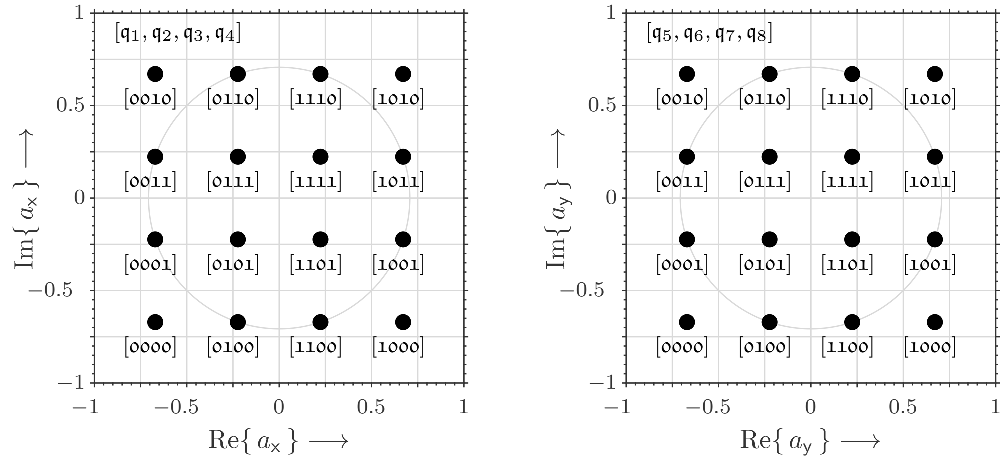


Figure 2.7: Two-dimensional projections of the data symbols $\boldsymbol{a} = [a_\mathsf{x}, a_\mathsf{y}]^\mathsf{T} \in \mathcal{A}$ with PDM 16-ary QAM and Gray-labeling per polarization. The variance of the signal constellation σ_a^2 is normalized to unity variance, no shaping assumed.

In Figure 2.7, the two-dimensional projections of the data symbols are shown. The symbols in the x-polarizations are addressed by $[\mathfrak{q}_1, \mathfrak{q}_2, \mathfrak{q}_3, \mathfrak{q}_4]^\mathsf{T}$, and the symbols in the y-polarization are addressed independently by $[\mathfrak{q}_5, \mathfrak{q}_6, \mathfrak{q}_7, \mathfrak{q}_8]^\mathsf{T}$. Since no shaping is considered, we see that the variance of the signal constellation is $\sigma_a^2 = 1$ (the circle in Figure 2.7 has radius $1/\sqrt{2}$).

Modulation The discrete-time data symbols $\boldsymbol{a}[k]$ are converted to the continuous-time transmit signal $\boldsymbol{s}(t)$ by means of *pulse-shaping* constituting the digital-to-analog (D/A) transition. The (linear) properties of the real, physical converter may conceptually be included in the pulse-shape. We can express the transmit signal $\boldsymbol{s}(t) = [s_1(t), s_2(t)]^\mathsf{T} = [s_\mathsf{x}(t), s_\mathsf{y}(t)]^\mathsf{T} \in \mathbb{C}^2$ as a function of the data symbols with [Fis02, (2.1.1)]

$$\boldsymbol{s}(t) = T \cdot \sum_{k \in \mathbb{Z}} \boldsymbol{a}[k] h_\mathrm{T}(t - kT)\,, \tag{2.74}$$

where $\boldsymbol{s}(t)$ is a superposition of a time-shifted (with symbol period T) *basic pulses* $h_\mathrm{T}(t)$ weighted by the data symbols.

Since $h_{\mathrm{T}}(t)$ has units s^{-1} (cf. Section 2.1.3), the pre-factor T is required to preserve a dimensionless signal in the continuous-time domain (cf. [Fis02, P. 11] or [Kam08, P. 230]). For the transmit (and respectively for the receive filter) we may also use the dimensionless entity—the basic pulse *shape*, given as

$$g_{\mathrm{T}}(t) \stackrel{\text{def}}{=} T \cdot h_{\mathrm{T}}(t)\,. \tag{2.75}$$

In a WDM scenario, the transmitter of interest (i.e., the *probe* channel) transmits at the *symbol rate* $R_{\mathrm{s}} \stackrel{\text{def}}{=} 1/T$. To keep the derivations in the following chapters tractable, all other wavelength channels transmit at the same symbol rate as the probe channel[5].

The pulse energy E_{T} is given by [Fis02, Eq. (2.2.22)]

$$E_{\mathrm{T}} = \int_{-\infty}^{\infty} |T \cdot h_{\mathrm{T}}(t)|^2 \mathrm{d}t = \frac{1}{2\pi} \int_{-\infty}^{\infty} |T \cdot H_{\mathrm{T}}(\omega)|^2 \mathrm{d}\omega\,. \tag{2.76}$$

The pulse energy E_{T} has the unit *seconds* due to the normalization, see above.

Using the symbol energy $E_{\mathrm{s}} \stackrel{\text{def}}{=} \sigma_a^2 E_{\mathrm{T}}$, the average signal power P (later equivalent to the optical signal power in two polarizations) calculates to [Fis02, Eq. (4.1.1)]

$$P \stackrel{\text{def}}{=} \frac{1}{T} \int_0^T \mathrm{E}\{\, \|\boldsymbol{s}(t)\|^2 \,\}\, \mathrm{d}t = \frac{\sigma_a^2}{T} E_{\mathrm{T}} = \frac{E_{\mathrm{s}}}{T}\,. \tag{2.77}$$

Here, we use the *cyclo-stationary* property of $\boldsymbol{s}(t)$ and average the expectation over a single symbol period. Since, see above, the variance of the data symbols σ_a^2 is fixed to 1, the transmit power P is directly adjusted via the pulse energy E_{T}. The corresponding quantities related to one of the other wavelength channels is indicated by the subscript ν.

We assume that the transmit pulse has $\sqrt{\text{Nyquist}}$ property. In frequency domain, the transmit pulse shape $g_{\mathrm{T}}(t) \circ\!\!-\!\!\bullet\, G_{\mathrm{T}}(\omega)$ has Nyquist property if the squared magnitude of the *periodic continuation* $\sum_{m\in\mathbb{Z}} |G_{\mathrm{T}}(\omega - \frac{2\pi m}{T})|^2$ sums up to a constant for all $\omega \in \mathbb{R}$ [Fis02, Eq. (2.2.4)]. This can be expressed in both frequency and time domain as

$$\sum_{m\in\mathbb{Z}} |T \cdot H_{\mathrm{T}}(\omega - \frac{2\pi m}{T})|^2 = \sum_{m\in\mathbb{Z}} |G_{\mathrm{T}}(\omega - \frac{2\pi m}{T})|^2 \stackrel{!}{=} 2\pi\, E_{\mathrm{T}} T = \text{const.} \tag{2.78}$$

$$\varphi_{gg}(\lambda T) \stackrel{\text{def}}{=} g_{\mathrm{T}}(t) * g_{\mathrm{T}}^*(-t)|_{t=\lambda T} \stackrel{!}{=} E_{\mathrm{T}}\, \delta[\lambda]\,, \qquad \lambda \in \mathbb{Z}. \tag{2.79}$$

In time domain, the $\sqrt{\text{Nyquist}}$ property of $g_{\mathrm{T}}(t)$ translates to the (first) Nyquist criterion applied to the *autocorrelation* $\varphi_{gg}(\lambda T)$ of the pulse $g_{\mathrm{T}}(t)$. That is, the autocorrelation $\varphi_{gg}(t)$, evaluated at integer multiples $\lambda \in \mathbb{Z}$ of the symbol period T, equals the *Dirac sequence* $\delta[\lambda]$, i.e., it is always *zero* unless $\lambda = 0$.

The root-raised cosine (RRC) shape fulfills the properties in (2.78), (2.79) and can be obtained from the raised cosine (RC) by $H_{\mathrm{RRC}}(\omega) = \sqrt{H_{\mathrm{RC}}(\omega)}$ with [PS08, P. 608]

$$H_{\mathrm{RC}}(\omega) = c_{\mathrm{e}} \cdot \begin{cases} 1, & 0 \le |\omega| \le \frac{(1-\rho)\pi}{T} \\ \frac{1}{2}\left(1 + \cos\left(\frac{T}{2\rho}(|\omega| - \frac{1-\rho}{2T})\right)\right), & \frac{(1-\rho)\pi}{T} \le \omega \le \frac{(1+\rho)\pi}{T} \\ 0, & |\omega| > \frac{(1+\rho)\pi}{T} \end{cases}, \tag{2.80}$$

[5]The considerations in the following are not restricted to the assumption of equal symbol rates over all wavelength channels; the derived results can generalized.

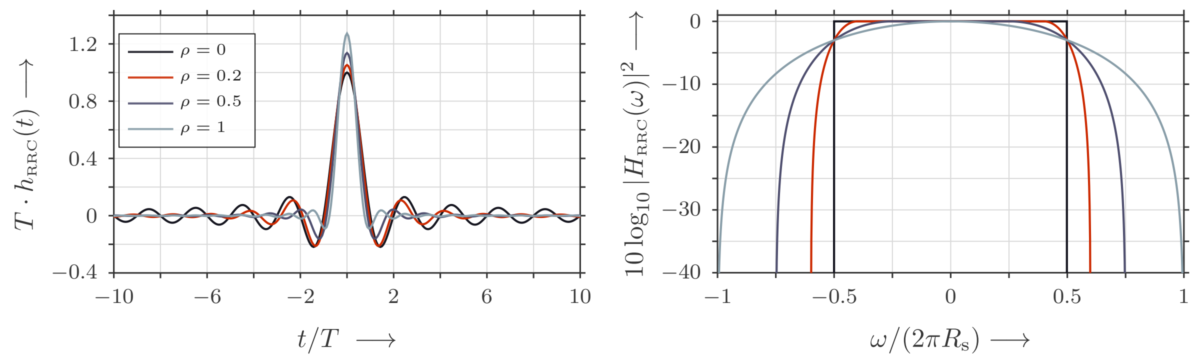


Figure 2.8: Transmit pulse shape for the (normalized) RRC pulse shape in time and frequency domain with roll-off factor $\rho = \{0, 0.2, 0.5, 1\}$.

where $0 \le \rho \le 1$ is the *roll-off* factor, which can be used to adjust the *bandwidth* of the probe signal. The leading constant c_{e} is a normalization constant that can be varied to adjust the pulse energy E_{T} (or later the optical launch power).

The signal bandwidth is defined as the width of the *spectral support* of the communication signal, i.e., the transmission *band*

$$\mathcal{B} \stackrel{\text{def}}{=} \{\omega \,\big|\, H_{\mathrm{T}}(\omega) \neq 0\}, \tag{2.81}$$

which yields for RC and RRC pulse shapes

$$B \stackrel{\text{def}}{=} R_{\mathrm{s}}(1+\rho). \tag{2.82}$$

Accordingly, in the next chapter, the bandwidth B will be equivalent to the bandwidth of the carrier-modulated signal at the beginning of the transmission link in the optical domain.

In Figure 2.8, the RRC pulse $h_{\mathrm{RRC}}(t) \circ\!\!-\!\!\bullet\, H_{\mathrm{RRC}}(\omega)$ is shown for a number of different roll-off factors. A small roll-off factor comes at the expense of stronger overshoots during the symbol transition and thus a continuous-time signal with higher *peak-to-average power ratio*. On the other hand, in the frequency domain, it is apparent that a small roll-off factor increases the *spectral efficiency* (also *bandwidth efficiency*), defined as

$$\mathit{\Gamma} \stackrel{\text{def}}{=} \frac{R_{\mathrm{T}}}{B} \stackrel{\text{uncoded}}{=} \frac{R_{\mathrm{m}}}{1+\rho} \qquad \left[\frac{\mathrm{bit/s}}{\mathrm{Hz}}\right], \tag{2.83}$$

where the (gross) *data rate* R_{T} is equal to $R_{\mathrm{m}} R_{\mathrm{s}}$ for transmission without error-correction coding and signal shaping.

In Figure 2.9, we show the autocorrelation function $\varphi_{gg}(t)$ and the periodic continuation corresponding to the RRC shape $g_{\mathrm{T}}(t) = T \cdot h_{\mathrm{RRC}}(t)$ with roll-off factor $\rho = 0.2$, cf. (2.78), (2.79). This property will become relevant in the deviation of the discrete-time end-to-end channel model in the subsequent chapters.

2.2.2 2×2 MIMO AWGN Channel

If only linear effects are considered during fiber-optic transmission, cf. Section 3.3.3, it is sufficient to describe the channel via a complex-valued 2×2 transfer matrix [Kar14]. This system description results in the so-called *Jones calculus* which allows to express any *physical*

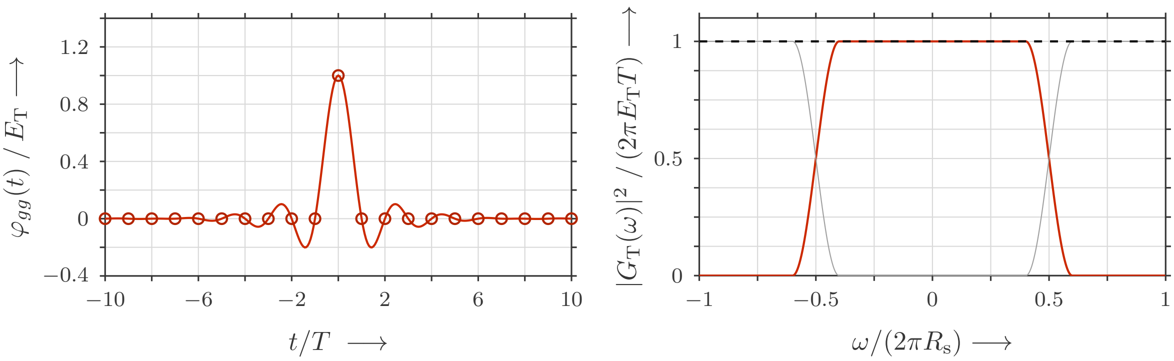


Figure 2.9: The autocorrelation function $\varphi_{gg}(t)$ (left) and the squared magnitude $|G_{\mathrm{T}}(\omega)|^2$ (right) of the basic pulse with RRC shape and roll-off factor $\rho = 0.2$. The autocorrelation function meets the first Nyquist criterion (cf. bullet markers), and vice-versa, the *periodic continuation*, i.e., the sum of all shifted (by $2\pi R_{\mathrm{s}}$) squared magnitudes $|G_{\mathrm{T}}(\omega)|^2$ (fundamental spectrum in red and shifted ones in gray), adds up to a constant (dashed black line).

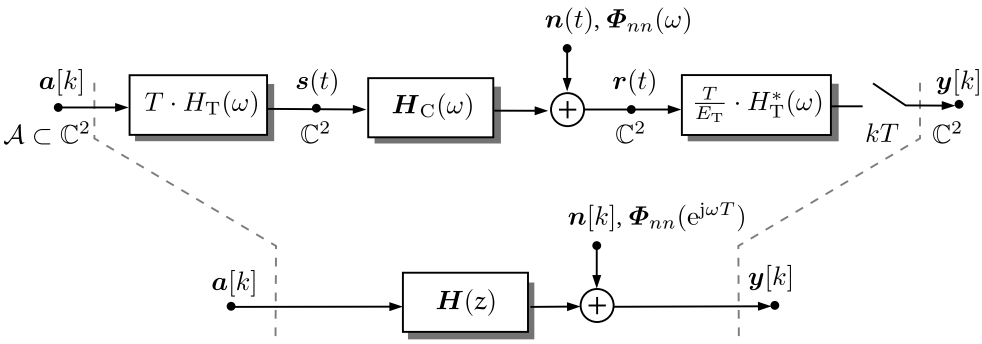


Figure 2.10: Block diagram of the complex-valued 2×2 inter-symbol interference transmission model and the derived discrete-time end-to-end channel, adapted from [Fis02, Fig. 2.3].

transformation of the electromagnetic field. The more general system description using the real-valued 4×4 signal space is a useful extension to the Jones calculus, e.g., for DSP, but is in principle not required to model fiber-optic transmission. A short introduction to the Jones formalism is given in the Appendix A.1.

In Figure 2.10, we show the block diagram of the complex-valued 2×2 linear, *dispersive* channel with AWGN at the channel output. The two complex components of the transmit signal $\boldsymbol{s}(t)$ are transmitted over the channel characterized by its complex-valued 2×2 channel impulse response and transfer function, interrelated by a Fourier transform according to

$$\boldsymbol{h}_{\mathrm{C}}(t) = \begin{bmatrix} h_{\mathrm{C,xx}}(t) & h_{\mathrm{C,xy}}(t) \\ h_{\mathrm{C,yx}}(t) & h_{\mathrm{C,yy}}(t) \end{bmatrix} \in \mathbb{C}^{2\times 2} \tag{2.84}$$

$$\circ\!\!-\!\!\bullet$$

$$\boldsymbol{H}_{\mathrm{C}}(\omega) = \begin{bmatrix} H_{\mathrm{C,xx}}(\omega) & H_{\mathrm{C,xy}}(\omega) \\ H_{\mathrm{C,yx}}(\omega) & H_{\mathrm{C,yy}}(\omega) \end{bmatrix} \in \mathbb{C}^{2\times 2}. \tag{2.85}$$

We neglect any effects of polarization dependent loss (PDL), and assume that the complex-valued channel matrix is *unitary* within the transmission band of the signal, i.e., the channel transfer function must satisfy [Kar14]

$$\boldsymbol{H}_{\mathrm{C}}^{-1}(\omega) \overset{!}{=} \boldsymbol{H}_{\mathrm{C}}^{\mathsf{H}}(\omega) \iff |\det\left(\boldsymbol{H}_{\mathrm{C}}(\omega)\right)| \overset{!}{=} 1, \quad \forall\, \omega \in \mathcal{B}\,. \tag{2.86}$$

We can understand the distortions introduced by the channel matrix as dispersive causing, e.g., inter-symbol interference (ISI), and as complex-valued rotations in $\mathbb{C}^2$, later termed *polarization rotations*. One instance of a channel matrix is given by *chromatic dispersion*, cf. Section 3.3.2, which effects both polarizations equally ($h_{\mathrm{C,xx}}(t) = h_{\mathrm{C,yy}}(t)$) and induces no polarization cross-talk ($h_{\mathrm{C,xy}}(t) = h_{\mathrm{C,yx}}(t) = 0$), i.e., the channel matrix reduces to a scalar. Both dispersion- and polarization-related effects are typically compensated using receiver-side equalization methods, see [Sav10] for an overview.

The transmit signal is inflicted with stationary, complex-valued, Gaussian noise in each of the two components $\boldsymbol{n}(t) = [n_1(t), n_2(t)]^{\mathsf{T}}$ at the channel output. We define the noise *correlation matrix* $\boldsymbol{\varphi}_{nn}(\tau)$ as

$$\boldsymbol{\varphi}_{nn}(\tau) = \mathrm{E}\{\, \boldsymbol{n}(t+\tau)\; \boldsymbol{n}^{\mathsf{H}}(t)\, \} \overset{\mathrm{AWGN}}{=} N_0\, \delta(\tau)\, \mathbf{I} \tag{2.87}$$

$$\circ\!\!-\!\!\bullet$$

$$\boldsymbol{\Phi}_{nn}(\omega) = \mathcal{F}_{\tau\leftrightarrow\omega}\{\, \mathrm{E}\{\, \boldsymbol{n}(t+\tau)\; \boldsymbol{n}^{\mathsf{H}}(t)\, \}\, \} \overset{\mathrm{AWGN}}{=} N_0 \mathbf{I} = \text{const.}, \quad \forall\, \omega\,, \tag{2.88}$$

which is interrelated to the average power spectral density (PSD) of the noise $\boldsymbol{\Phi}_{nn}(\omega)$ via the Fourier transform. We assume that the noise is independent of the transmit signal and *white* within and between the two complex components. Using the AWGN assumption we find that the noise PSD is *constant* in frequency. This results in a constant value N_0 on the main diagonal (i.e., same noise power in both polarizations) of the *PSD matrix* $\boldsymbol{\Phi}_{nn}(\omega)$ and a zero value for the x/y cross-terms.

In analogy to the conventional complex-valued AWGN system model, we denote the noise PSD per complex dimension by the constant N_0 and, conversely, the noise PSD per real dimension (i.e., per quadrature) is $N_0/2$. The *total* noise PSD in two components (i.e., *four* real dimensions) is equal to the *trace* of $\boldsymbol{\Phi}_{nn}(\omega)$ and equals $2N_0$ using the AWGN assumption.

The receive signal $\boldsymbol{r}(t)$ can hence be written as

$$\boldsymbol{r}(t) = \boldsymbol{h}_{\mathrm{C}}(t) * \boldsymbol{s}(t) + \boldsymbol{n}(t) \tag{2.89}$$

$$\circ\!\!-\!\!\bullet$$

$$\boldsymbol{R}(\omega) = \boldsymbol{H}_{\mathrm{C}}(\omega)\; \boldsymbol{S}(\omega) + \boldsymbol{N}(\omega)\,. \tag{2.90}$$

The objective of the receiver is then to detect the received signal and recover the transmitted information sequence $\langle\, \mathfrak{q}\, \rangle$.

2.2.3 Receiver Front-end

For the moment, we will neglect the effects of signal dispersion and polarization rotation, introduced by the channel matrix, and only consider the ISI–free AWGN channel. We assume that the transmit pulse-shape $g_{\mathrm{T}}(t)$ has $\sqrt{\text{Nyquist}}$ property. Then, the receiver filter that maximizes the signal-to-noise ratio (SNR) after detection is the *matched filter* w.r.t. the transmit pulse given as [Fis02]

$$h_{\mathrm{R}}^{(\mathsf{MF})}(t) \overset{\text{def}}{=} \frac{T}{E_{\mathrm{T}}} h_{\mathrm{T}}^{*}(-t) \tag{2.91}$$

$$\circ\!\!-\!\!\bullet$$

$$H_{\mathrm{R}}^{(\mathsf{MF})}(\omega) = \frac{T}{E_{\mathrm{T}}} H_{\mathrm{T}}^{*}(\omega)\,. \tag{2.92}$$

The leading factor $\frac{T}{E_{\mathrm{T}}}$ is a scaling factor to re-normalize the received sequence after sampling to the variance of the transmit sequence σ_a^2. Assuming an ISI–free channel, the transmit/receive filter cascade has an overall Nyquist impulse response.

Using the matched-filter receiver front-end and taking the (unitary) channel matrix $\boldsymbol{H}(\omega)$ into account, the linear end-to-end channel cascade is defined as

$$\boldsymbol{H}(\omega) \overset{\text{def}}{=} T \cdot H_{\mathrm{T}}(\omega)\, \boldsymbol{H}_{\mathrm{C}}(\omega)\, H_{\mathrm{R}}^{(\mathsf{MF})}(\omega)\,. \tag{2.93}$$

We arrive at the discrete-time end-to-end transfer function between the transmit sequence and receive sequence written as [Fis02, (2.1.7a)]

$$\boldsymbol{H}(\mathrm{e}^{\mathrm{j}\omega T}) = \textsc{Alias}_{\omega_{\mathrm{Nyq}}}\{\, \boldsymbol{H}(\omega)\,\} = \frac{1}{T} \sum_{m \in \mathbb{Z}} \boldsymbol{H}(\omega - \frac{2\pi m}{T}) \tag{2.94}$$

$$= \frac{1}{T} \sum_{m \in \mathbb{Z}} T \cdot H_{\mathrm{T}}\left(\omega - \frac{2\pi m}{T}\right) \boldsymbol{H}_{\mathrm{C}}\left(\omega - \frac{2\pi m}{T}\right) H_{\mathrm{R}}^{(\mathsf{MF})}\left(\omega - \frac{2\pi m}{T}\right) \tag{2.95}$$

$$= \frac{T}{E_{\mathrm{T}}} \sum_{m \in \mathbb{Z}} H_{\mathrm{T}}\left(\omega - \frac{2\pi m}{T}\right) \boldsymbol{H}_{\mathrm{C}}\left(\omega - \frac{2\pi m}{T}\right) H_{\mathrm{T}}^{*}\left(\omega - \frac{2\pi m}{T}\right), \tag{2.96}$$

where $\boldsymbol{H}(\mathrm{e}^{\mathrm{j}\omega T}) = \boldsymbol{H}(z)|_{z=\mathrm{e}^{\mathrm{j}\omega T}}$ is the periodic continuation of the end-to-end channel cascade.

In Figure 2.10, the block diagram of the discrete-time end-to-end equivalent is shown, including the channel matrix $\boldsymbol{H}(z)$ and the discrete-time noise equivalent $\boldsymbol{n}[k]$. The noise PSD matrix of the discrete-time sequence $\langle\, \boldsymbol{n}[k]\,\rangle$ is [Fis02, (2.1.7b)]

$$\boldsymbol{\Phi}_{nn}(\mathrm{e}^{\mathrm{j}\omega T}) = \frac{1}{T} \sum_{m \in \mathbb{Z}} \boldsymbol{\Phi}_{nn}\left(\omega - \frac{2\pi m}{T}\right) \left| H_{\mathrm{R}}^{(\mathsf{MF})}\left(\omega - \frac{2\pi m}{T}\right)\right|^2 \tag{2.97}$$

$$= \frac{N_0}{T} \mathbf{I} \sum_{m \in \mathbb{Z}} \left| H_{\mathrm{R}}^{(\mathsf{MF})}\left(\omega - \frac{2\pi m}{T}\right)\right|^2 \tag{2.98}$$

$$= \frac{N_0}{E_{\mathrm{T}}} \mathbf{I}\,, \tag{2.99}$$

where we used that the noise PSD is *white*, i.e., $\boldsymbol{\Phi}(\omega) = N_0 \mathbf{I}$, and that the receive filter (matched to the transmit basic pulse) also fulfills the $\sqrt{\text{Nyquist}}$ property in (2.78).

Under the white noise assumption, the (dimensionless) noise *variance* of the discrete-time Gaussian process in *four* dimensions then reads [Fis02, 2.2.6]

$$\sigma_n^2 = \frac{T}{2\pi} \int_{\mathbb{T}} \operatorname{trace}\left(\boldsymbol{\Phi}_{nn}(\mathrm{e}^{\mathrm{j}\omega T})\right) \mathrm{d}\omega = \operatorname{trace}\left(\frac{N_0}{E_{\mathrm{T}}} \mathbf{I}\right) = 2 \frac{N_0}{E_{\mathrm{T}}}, \tag{2.100}$$

where—by assumption—the variance per *real dimension* is $\sigma_{n,\mathrm{d}}^2 = \sigma_n^2/4$.

We can now give a formal definition of the SNR in 4D with

$$\mathsf{SNR} \overset{\text{def}}{=} \frac{E_\mathrm{s}}{2N_0} = \frac{\sigma_a^2}{\sigma_n^2}, \tag{2.101}$$

as the ratio between the *signal energy* in 4D and $2N_0$, which translates to the ratio between the 4D variance of the transmit and the noise sequence. For completeness, we also give the definition of the SNR *per real dimension* as

$$\mathsf{SNR}_\mathrm{d} \overset{\text{def}}{=} \frac{E_{\mathrm{s,d}}}{N_0/2} = \frac{\sigma_{a,\mathrm{d}}^2}{\sigma_{n,\mathrm{d}}^2}, \tag{2.102}$$

which is the relevant metric for *real-valued* transmission. Here, we used the corresponding quantities *per real dimension*, i.e., symbol energy per real-dimension $E_{\mathrm{s,d}}$ and variance per real-dimension.

The *Shannon capacity* (or Shannon limit) is the upper bound of the transmission rate over the AWGN channel at which error-free communication is possible—achieved for Gaussian-distributed input.

The Shannon capacity in *bits per real dimension* ($\mathrm{b/1D}$) is given as [Sha48]

$$C_{[\mathrm{b/1D}]} = \frac{1}{2} \cdot \log_2\left(1 + \mathsf{SNR}_\mathrm{d}\right) = \frac{1}{2} \cdot \log_2\left(1 + \frac{E_{\mathrm{s,d}}}{N_0/2}\right) = \frac{1}{2} \cdot \log_2\left(1 + \frac{\sigma_{a,\mathrm{d}}^2}{\sigma_{n,\mathrm{d}}^2}\right). \tag{2.103}$$

In analogy, we define the capacity over the complex-valued 2×2 AWGN channel (in *bits per four dimensions*) as

$$C_{[\mathrm{b/4D}]} = 2 \cdot \log_2\left(1 + \mathsf{SNR}\right) = 2 \cdot \log_2\left(1 + \frac{E_\mathrm{s}}{2N_0}\right) = 2 \cdot \log_2\left(1 + \frac{\sigma_a^2}{\sigma_n^2}\right). \tag{2.104}$$

In Figure 2.11, the Shannon capacity per four dimensions is shown together with the *constellation-constrained* capacity for PDM $\{4, 16, 64\}$-ary QAM, i.e., the upper limit of *achievable information rate* given the respective QAM format at the channel input.

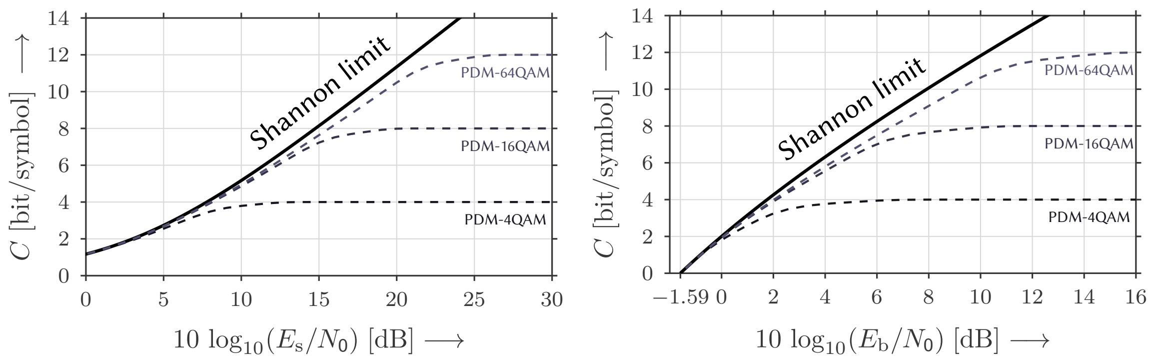


Figure 2.11: Shannon capacity (solid) and constellation-constrained capacities (dashed) of PDM $\{4, 16, 64\}$-ary QAM over the complex-valued 2×2 AWGN channel. The capacity C is given in *bit per (4D) symbol* and the SNR in dB is expressed via the ratio of the (4D) symbol energy E_s over the (one-sided) noise power spectral density N_0 (left). The same quantities are also shown over the ratio between the *energy per information bit* E_b and N_0 in dB for uncoded transmission.

3. Principles of Fiber-Optic Transmission

This chapter concerns with the principal effects of fiber-optic transmission.

It first reviews the concept of the *optical field envelope* in Section 3.1. The optical field envelope is the equivalent in optical communication to the complex baseband signal. The main approximations and assumptions to arrive at the baseband signal and system description are briefly reviewed.

In Section 3.2, we extend the signal model to frequency-multiplexed signals which is historically termed wavelength-division multiplexing (WDM) in the optical communications community.

In the last Section, the signal evolution equation, i.e., the *Manakov equation* and some fundamental transmission effects are reviewed. This includes *chromatic dispersion*, signal *attenuation* and (common) *amplification* schemes, and the *nonlinear Kerr effect.*

This chapter forms the basis for the topics dealing with the perturbation-based channel models developed in the following two chapters.

3.1 The Optical Field Envelope

The two-dimensional complex vector $\boldsymbol{u}(z,t) = [u_{\mathsf{x}}(z,t), u_{\mathsf{y}}(z,t)]^{\mathsf{T}} \in \mathbb{C}^2$ in *Jones space* [Jon41] is associated with the transverse electric field component of a plane electro-magnetic wave propagating in z-direction. The two components of the *Jones vector* $u_{\mathsf{x}}(z,t)$ and $u_{\mathsf{y}}(z,t)$ represent the x- and y-components, i.e., the two *polarizations* of the electric field. In practice, the electric field is modeled as a continuous-time stochastic process, where $\boldsymbol{u}(z,t)$ is a particular realization of that process with dependency on the time variable t and spatial direction z.

Similar to radio-frequency transmission, the quantity $\boldsymbol{u}(z,t) \in \mathbb{C}^2$ is in fact the *equivalent complex baseband* (ECB) representation of the real-valued electric field. In the optics community, $\boldsymbol{u}(z,t)$ is denoted *optical field envelope* and its relation to the real optical field in the passband domain is given by

$$\boldsymbol{u}_{\mathrm{o}}(z,t) = \sqrt{2}\,\mathrm{Re}\{\boldsymbol{u}(z,t)\,\mathrm{e}^{\mathrm{j}\omega_0 t - \mathrm{j}\beta(z,\omega_0)z}\} \in \mathbb{R}^2\,, \tag{3.1}$$

with the center frequency (of the signaling regime of interest) $\omega_0 = 2\pi f_0$. The center frequency ω_0 typically coincides with the carrier frequency of the probe signal (i.e., the signal or channel of interest) which is in the order of $184\,\mathrm{THz}$ to $237\,\mathrm{THz}$ in optical telecommunication.

Example 3.1: RRC baseband and passband pulse subject to chromatic dispersion

A preliminary example of the optical field $u_{\mathrm{x,o}}(z,t) \in \mathbb{R}$ and the optical field envelope $u_{\mathrm{x}}(z,t) \in \mathbb{C}$ in one polarization subject to chromatic dispersion is sketched in Figure 3.1.

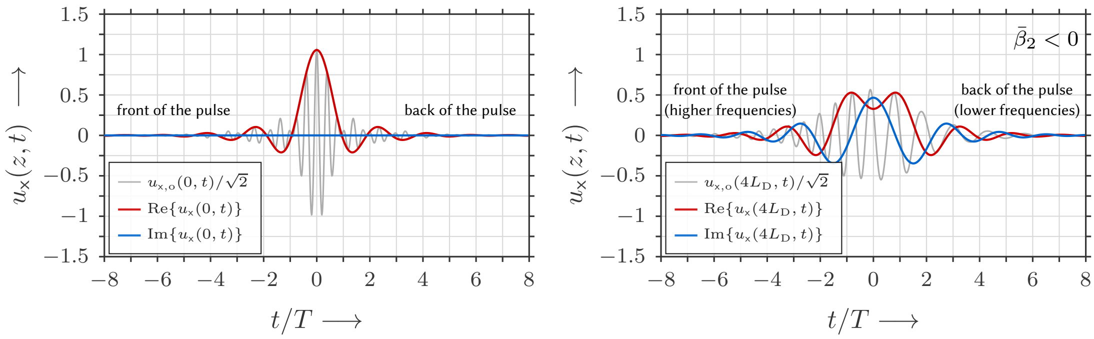


Figure 3.1: Illustration of the *optical field envelope* (in one polarization) $u_{\mathrm{x}}(z,t)$ and the real-valued *optical field* $u_{\mathrm{x,o}}(z,t)$ at the carrier frequency $\omega_0 \gg 2\pi B$ (in the drawing, the oscillation of the carrier frequency ω_0 is heavily understated, and orders of magnitude higher in practice). The envelope of a single RRC pulse with roll-off factor $\rho = 0.2$ is shown at $z = 0$ at the input of a transmission link (left). After propagation (right), the initial pulse is *dispersed*, here shown at $z = 4L_{\mathrm{D}}$ where L_{D} is the so-called *dispersion length*, see Section 3.3.2. Only second-order dispersion assumed, group delay canceled by the retarded time-frame, and constant phase rotation is $\beta(4L_{\mathrm{D}}, \omega_0) = 0$.

The initial baseband pulse at $z = 0$ is real-valued and has an RRC shape with roll-off factor $\rho = 0.2$. The baseband signal is modulated to the (high-frequency) optical passband signal at carrier frequency $\omega_0 \gg 2\pi B$, with the signal bandwidth B. After transmission at $z = 4L_{\mathrm{D}}$ (a formal definition of the so-called dispersion length L_{D} will be given in Section 3.3.2), the optical pulse is dispersed, i.e., the pulse is temporally broadened w.r.t. its initial shape. In the *anomalous dispersion* regime (i.e., at the typical telecommunication wavelength), high-frequency components of an optical pulse travel faster (to the temporal front of the pulse) than low-frequency components of the same pulse. This can be seen in Figure 3.1 (right) recognizing that the passband signal (gray trace) has higher frequency components in the temporal front, and lower frequency components in the temporal back.

The main difference to the classical ECB representation, cf., e.g., [PS08, Fis02], is the additional term appearing in the complex exponential which depends on the so-termed (space- and frequency-dependent) *common propagation constant* $\beta(z,\omega)$, where the term "common" relates to the x- and y-component of the electric field. Just like $\omega_0/(2\pi)$ measures the number of oscillation of the electric field per unit time, $\beta(z,\omega_0)/(2\pi)$ measures the number of field oscillations per unit length in z-direction. By only considering the optical field envelope, i.e., by using the (generalized) ECB representation, this (trivial) phase rotation at arbitrary (fixed) position z_0 is removed.

The frequency-dependency of the propagation constant $\beta(z,\omega)$ is the source for the *disper-*

sive nature of light—the predominant, linear distortion in optical communication. An in-depth discussion on chromatic dispersion will be provided in Section 3.3.2.

In writing (3.1), we already made use of a number of common assumptions which go beyond the conventional ECB representation. The required transformations on the optical signal and the corresponding systems to model optical field propagation over a single-mode fiber can be found in many popular textbooks on that topic (see, e.g., [Agr02,Sei09,Agr10,Kam13,KD14]). We briefly summarize these assumptions in the following paragraphs.

Plane Wave Approximation The *plane wave approximation* allows us to consider only the longitudinal z-dependence of the optical field (no transversal field components into the x- and y-direction, cf. [Men89]). The z-direction is the direction of propagation along the fiber, depicted as a straight line/link from the transmitter at $z = 0$ to the receiver at $z = L$.

Slowly Varying Amplitude Approximation The *slowly varying amplitude approximation*, cf. [Agr10, Eq. (2.4.5)], assumes that the envelope $\boldsymbol{u}(z,t)$ changes only slowly *in time* compared to the order of the carrier period $1/f_0$. In other words, the spectrum of the signal is narrow-banded (compared to the carrier frequency) and we have $B_{\text{WDM}} \ll f_0$, where B_{WDM} is the full spectral bandwidth of the signal under consideration (including co-propagating wavelength channels, see Section 3.2 on WDM). E.g., at the typical communication wavelength $\lambda_0 = \mathrm{c}/f_0$ of $1550\,\mathrm{nm}$, the ratio between the bandwidth $B = R_{\mathrm{s}}(1+\rho)$ of a single wavelength channel and the carrier frequency $f_0 = 193.4\,\mathrm{THz}$ is in the order of $1/2500$ for $R_{\mathrm{s}} = 64\,\mathrm{GBd}$ and $\rho = 0.2$. Hence, even when hundreds of such wavelength channels are included in the same ECB signal, the approximation is still valid.

Quasi-monochromatic Approximation The *quasi-monochromatic approximation*, cf. [Agr10, Eq. (2.4.4)], extends the former to the space- and frequency-dependent common propagation constant $\beta(z,\omega)$, which is expanded into a Taylor series in the vicinity of the center frequency ω_0. The real optical field $\boldsymbol{u}_{\mathrm{o}}(z,t)$ can then be separated into its (temporal and spatial) carrier $\mathrm{e}^{\mathrm{j}\omega_0 t - \mathrm{j}\beta(z,\omega_0)z}$ and the envelope $\boldsymbol{u}(z,t)$. The evolution of the optical field is solely described by considering its envelope. Then, the evolution equation in the baseband description (i.e., the system description which will be used in this work) is solely determined by the Taylor coefficients of $\beta(z,\omega)$ developed at ω_0.

The Taylor series expansion of the propagation constant $\beta(z,\omega)$ w.r.t. ω (in the passband domain) reads

$$\beta(z,\omega) = \beta_0(z) + \beta_1(z)(\omega-\omega_0) + \frac{1}{2}\beta_2(z)(\omega-\omega_0)^2 + \frac{1}{6}\beta_3(z)(\omega-\omega_0)^3 + \dots\,, \tag{3.2}$$

with the partial derivatives represented by the coefficients [Agr10, Eq. (2.4.4)]

$$\beta_n(z) \overset{\text{def}}{=} \left.\frac{\partial^n \beta(z,\omega)}{\partial \omega^n}\right|_{\omega=\omega_0}, \quad n \in \mathbb{N}\,. \tag{3.3}$$

We also introduce the following notation to denote *path-average* Taylor coefficients

$$\bar{\beta}_n \overset{\text{def}}{=} \frac{1}{L}\int_0^L \beta_n(z)\,\mathrm{d}z\,. \tag{3.4}$$

which will turn out to be practical when describing path-average quantities in the following sections.

The zero$^{\text{th}}$-order coefficient is given by

$$\beta_0(z) = \beta(z, \omega_0)\,, \tag{3.5}$$

which measures the number of oscillation per unit length into the longitudinal direction. This (frequency-independent) oscillation in z-direction by $\mathrm{e}^{-\mathrm{j}\beta(z,\omega_0)z}$ is removed from the optical field envelope $\boldsymbol{u}(z,t)$ (i.e., the optical baseband signal) following the ansatz in (3.1).

The first-order Taylor coefficient reads

$$\beta_1(z) = \left.\frac{\partial\beta(z,\omega)}{\partial\omega}\right|_{\omega=\omega_0}\,, \tag{3.6}$$

and has units $\mathrm{s/m}$. The (path-average) first-order coefficient relates inversely to the (path-average) *group velocity* at the center frequency ω_0 given by [Agr10, Eq. (2.3.1)]

$$\mathrm{v_g}(\omega_0) \stackrel{\text{def}}{=} \frac{1}{\bar{\beta}_1}\,. \tag{3.7}$$

As a result, a spectral component at ω_0 propagates from the fiber input at $z = 0$ to the output at $z = L$ within $L/\mathrm{v_g}(\omega_0)$ *seconds*.

The second-order Taylor coefficient is

$$\beta_2(z) = \left.\frac{\partial^2\beta(z,\omega)}{\partial\omega^2}\right|_{\omega=\omega_0}\,, \tag{3.8}$$

and quantifies the dependency of the group-velocity on the frequency deviation from ω_0. This dependency gives rise to *chromatic dispersion*, and the Taylor coefficient is hence termed group velocity dispersion (GVD) parameter. The (*path-average*) GVD parameter $\bar{\beta}_2$ has a zero-crossing for typical standard single-mode fibers (SSMFs) at the *zero dispersion wavelength* $\lambda_{\mathrm{ZD}} \approx 1276\,\mathrm{nm}$. The transmission regime for shorter wavelengths is termed *normal* dispersion regime with $\bar{\beta}_2 > 0$, and the transmission regime for longer wavelengths (i.e., also at typical communication wavelengths) is termed *anomalous* dispersion regime with $\bar{\beta}_2 < 0$, cf. [Agr10, P. 41].

The GVD parameter $\beta_2(z)$ has units of $\mathrm{s^2/m}$ or $\mathrm{s/Hz/m}$, however, for historical reasons chromatic dispersion is often quantified in terms of wavelength separation instead of frequency separation from ω_0. The corresponding (path-average) dispersion parameter in terms of wavelength is [Agr10, Eq. (2.3.5)]

$$\bar{D}_{\mathrm{GVD}} = -\frac{2\pi\mathrm{c}}{\lambda_0^2}\bar{\beta}_2\,. \tag{3.9}$$

A typical value for the GVD parameter is $\bar{\beta}_2 = -21.4\ \mathrm{ps^2/km}$ at $\lambda_0 = 1550\,\mathrm{nm}$, which in turn corresponds to $\bar{D}_{\mathrm{GVD}} = 16.8\ \mathrm{ps/nm/km}$.

The third-order Taylor coefficient reads

$$\beta_3(z) = \left.\frac{\partial^3 \beta(z,\omega)}{\partial \omega^3}\right|_{\omega=\omega_0} , \tag{3.10}$$

and is known as the *dispersion-slope* parameter. Higher order terms of the Taylor expansion are usually negligible and dispersion-slope only becomes relevant when $\beta_2(z)$ approaches zero, i.e., close to the zero-dispersion wavelength. The implications of dispersion in fiber transmission systems will be discussed in detail in Section 3.3.

Example 3.2: Taylor series expansion of $\beta(z,\omega)$

Transmission in the so-called *C-band* (conventional operation regime of erbium-doped fiber amplifiers (EDFAs)) spans a wavelength range from 1530 to $1565\ \mathrm{nm}$, i.e., a transmission window of $35\ \mathrm{nm}$, which is equivalent to a total bandwidth of $4.375\ \mathrm{THz}$. This supports a total of $N_{\mathrm{ch}} = 87$ wavelength channels each having a $50\ \mathrm{GHz}$ slot. The center frequency is $f_0 \approx 193.4\ \mathrm{THz}$ ($\lambda_0 = 1550\ \mathrm{nm}$) and the ratio between the total bandwidth and the carrier frequency is approximately $1/50$, i.e., the signal is still narrow-banded compared to its carrier frequency. In that frequency region, $\beta(z,\omega)$ can be well approximated with only two Taylor coefficients $\beta_1(z)$ and $\beta_2(z)$. The third Taylor coefficient $\beta_3(z)$ becomes only relevant if the center wavelength $\lambda_0 = \mathrm{c}/\omega_0$ is close to the *zero dispersion wavelength* $\lambda_{\mathrm{ZD}} \approx 1276\ \mathrm{nm}$ where $\beta_2(z) \to 0$.

Retarded Time Frame To make the mathematical treatment of the z- and t-dependent optical field envelope $\boldsymbol{u}(z,t)$ in the upcoming sections more tractable, the accumulated (path-average) timing delay $z/\nu_{\mathrm{g}}(\omega_0)$ of the optical field at the center frequency ω_0 is canceled out by considering signal evolution in the so-termed *retarded time frame*.

The retarded time frame is defined as [Agr10, Eq. (2.4.8)]

$$t \stackrel{\text{def}}{=} t' - z/\nu_{\mathrm{g}}(\omega_0)\,, \tag{3.11}$$

where t' is the *original*, un-retarded time base of the real optical field. The retarded time frame can be imagined as a reference time frame that moves at the path-average group velocity $\nu_{\mathrm{g}}(\omega_0)$ of the probe channel with increasing z (we assume that the probe channel is centered at ω_0). At fixed z, e.g., at the receiver with $z = L$, the end-to-end retardation of the probe signal can be modeled (if required) as a linear contribution of the overall system's phase response.

As a consequence, the average group delay of the probe signal is removed from the propagation equation. This has, e.g., already been used in Figure 3.1 where no group delay is present after propagation, i.e., the dispersed pulse at $z = 4L_{\mathrm{D}}$ is still centered around $t/T = 0$. Similarly, the (linear) impulse response of the optical end-to-end system will be centered at $t/T = 0$ since the (average) timing delay has been removed from the propagation equation, see Section 3.3.

All wavelength channels other than the probe channel experience a residual timing delay after transmission relative to $t/T = 0$ due to chromatic dispersion, see Section 3.3.2.

Example 3.3: Group delay

A pulse centered at $\omega_0 = 2\pi \cdot 193.4$ THz (i.e., $\lambda_0 \approx 1550$ nm) propagating over a SSMF (e.g., according to the ITU recommendation G.652) with refractive index of 1.4682 at ω_0 has a (path-average) group velocity of

$$\nu_\mathrm{g}(\omega_0) = \frac{1}{\bar{\beta}_1} \approx \frac{\mathrm{c}}{1.4682} \approx 204.190\ \mathrm{m/\mu s}\,,$$

and is delayed by $\bar{\beta}_1 L \approx 489.73\ \mu\mathrm{s}$ for $L = 100\ \mathrm{km}$. This average group delay at ω_0 is canceled in the signal's baseband description by considering the retarded time frame.
If a second pulse centered at a different frequency $\omega_0 + \Delta\omega$ is launched over the same fiber, it is delayed by a different amount if $\nu_\mathrm{g}(\omega_0) \neq \nu_\mathrm{g}(\omega_0 + \Delta\omega)$, i.e., if any of the higher order dispersion coefficients $\beta_n(z)$ is different form zero. This frequency-dependent delay is modeled in baseband by considering higher-order dispersion coefficients.
E.g., in a SSMF according to G.652, the refractive index at 228 THz (i.e., 1310 nm) is 1.4677. A pulse centered at this frequency is instead delayed by $489.57\ \mu\mathrm{s}$. In the ECB model using the retarded time frame, only the relative delay between the two pulses is considered. E.g., the second pulse at the relatively higher frequency will be delayed w.r.t. the first pulse at ω_0 by $\tau \approx -0.16\ \mu\mathrm{s}$.

Local Birefringence Coordinate Transformation Transmission of a dual-polarized signal over a fiber is subject to *birefringence*, i.e., the local refractive index of the fiber core may vary for the x- and the y-polarization. This is caused by the anisotropy of the core material or, e.g., mechanical stress applied to the fiber.

To ease the analysis, we treat the optical field envelope $\boldsymbol{u}(z,t)$ as a transformed signal representation where the influence of the *local*, frequency-*independent* birefringence is canceled from the propagation equation using a suited (z-dependent) coordinate transformation, see, e.g., [MM06, Sec. B]. In doing so, only differential polarization rotations (what will be introduced as first- and higher-order *polarization-mode dispersion* in the next section) are still present.

The details of this coordinate transformation are discussed in the Appendix A.1.

3.2 Wavelength-Division Multiplexing

One of the main constraints of fiber-optical transmission systems is the bandwidth of electronic devices which is orders of magnitude smaller than the available bandwidth of optical fibers. It is hence routine to use wavelength-division multiplexing (WDM), where a number of so-called *wavelength channels* are transmitted simultaneously over the same fiber. Each wavelength signal is modulated on an individual laser operated at a certain wavelength (or respectively at a certain frequency) such that neighboring signals do not share the same frequency band when transmitted jointly over the same fiber medium.

We assume, that each wavelength channel comprises essentially the same transmit and receive frond-end as discussed in Section 2.2 on the linear, point-to-point system model. We continue to use the same nomenclature also for the following considerations, but add the optical domain as part of the transmission channel.

The nonlinear property of the fiber-optical transmission medium is the source of nonlinear interference within and between different wavelength channels, so-called *signal-to-signal* nonlinear interference. To ease the analysis, the accumulated nonlinear distortion on a single selected wavelength channel in the neighborhood of other wavelength channels is considered. The channel under consideration is called *probe* channel, while co-propagating wavelength channels are called *interfering* channel.

When modeling transmission of a WDM signal comprising multiple wavelength channels, both signal and system description is done using the ECB domain introduced in the previous section. The individual wavelength ECB signals are conceptionally combined into a single, comb-shaped communication signal which is then used to model the joint transmission of all wavelength signals at once.

Figure 3.2 shows the block diagram of a coherent optical transmission system exemplifying the digital, analog, and optical domains of a probe wavelength channel in the neighborhood of other wavelength channels. Both the digital and analog domain remain the same as compared

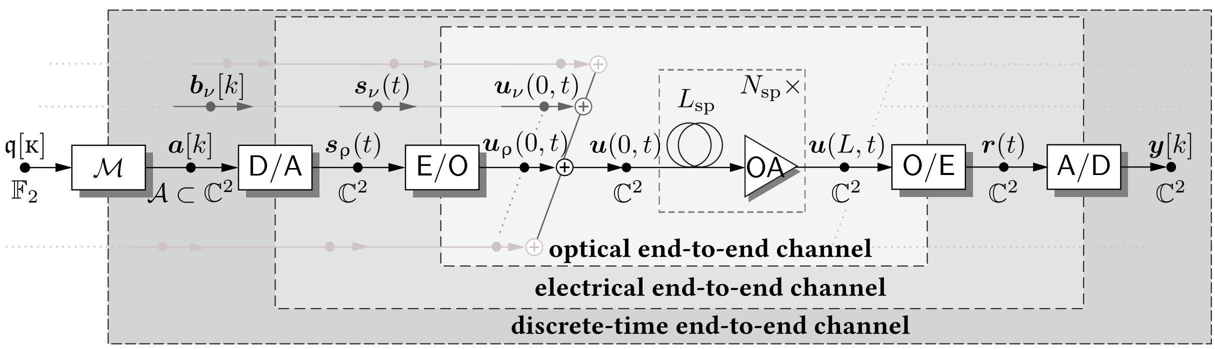


Figure 3.2: Equivalent complex baseband model of a point-to-point fiber-optical transmission system [FFF20]. The individual wavelength channels have a transmit and receive front-end similar to the previous chapter. At the beginning of the optical transmission link, the optical probe signal $\boldsymbol{u}_\rho(0,t)$ is combined with the co-propagating wavelength signals $\boldsymbol{u}_\nu(0,t)$ to be transmitted jointly over the transmission link. The link consists of N_{sp} spans, each a cascade of fiber of length L_{sp} and optical amplifier (OA). On the receive-side at $z = L$, the probe signal is selected via a suited channel selection filter, and post-processed in the analog and digital domain.

to Section 2.2. We now focus on the optical domain.

The probe signal in the optical domain $\boldsymbol{u}_\rho(z,t)$ is denoted by a subscript

$$\rho \in \{1, 2, \ldots, N_\text{ch}\}, \tag{3.12}$$

whereas interferer signals $\boldsymbol{u}_\nu(z,t)$ are labeled by the channel index ν with

$$\nu \in \{1, 2, \ldots, N_\text{ch} \mid \nu \neq \rho\}, \tag{3.13}$$

where N_ch is the total number of wavelength channels considered.

The electrical-to-optical (E/O) conversion is assumed to be ideal. This corresponds to a lossless and frequency-flat dual-polarization (DP) inphase-quadrature (IQ) converter such that the analog baseband signal is ideally mixed to its respective carrier frequency. Again, the linear characteristics of the converter can be modeled as part of the analog frond-end of the individual wavelength channels, if necessary.

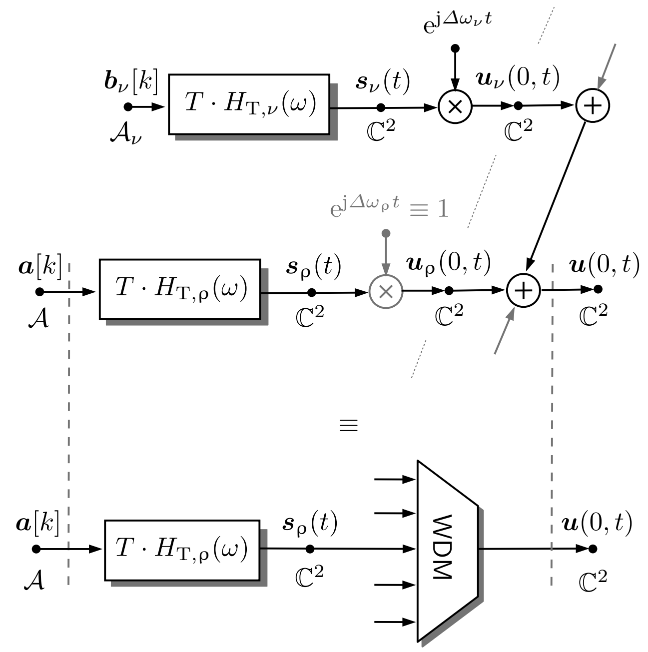


Figure 3.3: Modulation of the probe's electrical transmit signal $\boldsymbol{s}_\rho(t)$ to the optical field envelope $\boldsymbol{u}_\rho(0,t)$ via ideal (i.e., lossless and frequency-flat) DP-IQ conversion and mixing to the target carrier frequency. The probe's optical transmit signal is ideally combined with the modulated optical signals of the remaining $N_\text{ch}-1$ wavelength channels to obtain the WDM signal $\boldsymbol{u}(0,t)$. The joint operation is denoted as WDM in the equivalent block diagram.

The transmitter front-end of the WDM communication signal is shown in Figure 3.3. The two elements of the analog transmit signal $\boldsymbol{s}_\nu(t)$ are converted to the modulated *optical* signals in the x- and y-polarization. The *optical field envelope* $\boldsymbol{u}_\nu(z,t)$ of each wavelength channel is

$$\boldsymbol{u}_\nu(0,t) = \boldsymbol{s}_\nu(t)\exp(\mathrm{j}\Delta\omega_\nu t), \tag{3.14}$$

modulated at $\Delta\omega_\nu$ relative to the center frequency such that the (passband) carrier frequency is $\omega_\nu = \omega_0 + \Delta\omega_\nu$. The carrier frequency of the probe channel ω_ρ typically coincides with ω_0

such that $\Delta\omega_\rho = 0$ and $\boldsymbol{u}_\rho(0,t) = \boldsymbol{s}_\rho(t)$. This also implies that the group delay of the probe channel is already canceled from the propagation equation.

In this simplified view, all co-propagating wavelength channels are co-polarized and share the same common phase at the input of the link. An additional phase term and birefringent element can be used in (3.14) to invoke a random initial phase and polarization state for each wavelength channel. Beyond that, laser phase noise (PN), both on the transmit and receive end, is not considered at this point to focus only on deterministic effects; that is, nonlinear signal-to-signal interference in the following chapter.

The N_ch wavelength signals $\boldsymbol{u}_\nu(0,t)$ at $z=0$ are combined by an ideal optical multiplexer to a single WDM signal. The optical field envelope of the signal comb is given by

$$\boldsymbol{u}(0,t) = \sum_{\nu=1}^{N_\mathrm{ch}} \boldsymbol{u}_\nu(0,t) = \sum_{\nu=1}^{N_\mathrm{ch}} \boldsymbol{s}_\nu(t)\exp(\mathrm{j}\Delta\omega_\nu t) \tag{3.15}$$

$$\circ\!\!-\!\!\bullet$$

$$\boldsymbol{U}(0,\omega) = \sum_{\nu=1}^{N_\mathrm{ch}} \boldsymbol{U}_\nu(0,\omega) = \sum_{\nu=1}^{N_\mathrm{ch}} \boldsymbol{S}_\nu(\omega-\Delta\omega_\nu)\,, \tag{3.16}$$

with the Fourier pairs $\boldsymbol{s}_\nu(t) \circ\!\!-\!\!\bullet\, \boldsymbol{S}_\nu(\omega)$ and $\boldsymbol{u}(0,t) \circ\!\!-\!\!\bullet\, \boldsymbol{U}(0,\omega)$. Since $\Delta\omega_\rho = 0$, we can write the optical field also as

$$\boldsymbol{u}(0,t) = \mathrm{WDM}\{\,\boldsymbol{s}_\nu(t)\,\} \overset{\mathrm{def}}{=} \boldsymbol{s}_\rho(t) + \sum_{\nu\neq\rho} \boldsymbol{s}_\nu(t)\exp(\mathrm{j}\Delta\omega_\nu t)\,. \tag{3.17}$$

In analogy to the spectral support of a single wavelength channel in its *baseband* $\mathcal{B}_\nu$, see (2.81), we define the WDM transmission band as

$$\mathcal{B}_\mathrm{WDM} \overset{\mathrm{def}}{=} \{\,\omega \,\big|\, H_{\mathrm{T},\nu}(\omega-\Delta\omega_\nu) \neq 0,\ \forall\,\nu\,\}\,. \tag{3.18}$$

The total bandwidth of the WDM signal B_WDM is defined as the width of the closed interval

$$[\omega_\mathrm{WDM}^\mathrm{min}, \omega_\mathrm{WDM}^\mathrm{max}] = \{\,\omega \in \mathcal{B}_\mathrm{WDM} \,\big|\, \omega_\mathrm{WDM}^\mathrm{min} \leq \omega \leq \omega_\mathrm{WDM}^\mathrm{max}\,\}\,, \tag{3.19}$$

i.e., we have

$$B_\mathrm{WDM} \overset{\mathrm{def}}{=} \omega_\mathrm{WDM}^\mathrm{max} - \omega_\mathrm{WDM}^\mathrm{min}\,. \tag{3.20}$$

3.3 The Fundamental Evolution Equation

The propagation of the optical baseband signal $\boldsymbol{u}(z,t)$ is governed by the *Manakov equation.* It is a coupled set of partial differential equations in time domain which combines the dominating effects governing the propagation of the optical signal—chromatic dispersion, signal power (or equivalently field amplitude) variation, and the nonlinear Kerr interaction.

The (noiseless) Manakov equation reads [MM06, Eq. (64)]

$$\frac{\partial}{\partial z}\boldsymbol{u}(z,t) = \underbrace{\mathrm{j}\frac{\beta_2(z)}{2}\frac{\partial^2}{\partial t^2}\boldsymbol{u}(z,t)}_{\textit{chromatic dispersion}} + \underbrace{\frac{g(z)-\alpha(z)}{2}\boldsymbol{u}(z,t)}_{\textit{signal gain/loss}} \underbrace{- \mathrm{j}\gamma(z)\frac{8}{9}\,\|\boldsymbol{u}(z,t)\|^2\,\boldsymbol{u}(z,t)}_{\textit{optical Kerr effect}}\,. \tag{3.21}$$

The Manakov equation is formulated in the ECB as introduced in the previous section, i.e., the coefficients $\beta_0(z)$ and the $\beta_1(z)$, as well as the zero-order birefringence $\Delta\beta_0(z)$ have already been removed from the evolution of the optical field envelope. We also neglect the dispersion slope $\beta_3(z)$, and drop the PMD term $\Delta\beta_1(z)$ and higher-order terms. To that end, the nonlinearity coefficient $\gamma(z)$ is weighted by the factor $8/9$ to account for the reduced effective nonlinear strength due to birefringence and PMD-induced polarization mixing, see e.g., the discussion in [MM06, Sec. B]. We also neglect the time- (and frequency-) dependency of the attenuation, gain, and nonlinearity coefficient. At last, a (distributed, i.e., z-dependent) noise term $\boldsymbol{n}(z,t)$ is also not considered in the Manakov equation as we will focus on *signal-to-signal* nonlinear interference.

As before, we define the path-average coefficients as

$$\bar{\alpha} \overset{\text{def}}{=} \frac{1}{L}\int_0^L \alpha(\zeta)\,\mathrm{d}\zeta\,, \qquad \bar{\gamma} \overset{\text{def}}{=} \frac{1}{L}\int_0^L \gamma(\zeta)\,\mathrm{d}\zeta\,. \qquad \bar{\beta}_2 \overset{\text{def}}{=} \frac{1}{L}\int_0^L \beta_2(\zeta)\,\mathrm{d}\zeta\,. \tag{3.22}$$

In the following, we will discuss the individual terms and the involved quantities of the Manakov equation in more detail.

3.3.1 Power Profile and Amplification Noise

Despite the low attenuation of optical fibers, optical amplification (OA) is required to facilitate reliable communication over hundreds of kilometers. E.g., a typical value for the path-average attenuation coefficient[1] is $10\log_{10}(\mathrm{e}^{\bar{\alpha}}) = 0.2\ \mathrm{dB/km}$, i.e., the signal power P is reduced over one span of length $L_{\mathrm{sp}} = 80\,\mathrm{km}$ by the amount of $10\log_{10}(\mathrm{e}^{-\bar{\alpha}L_{\mathrm{sp}}}) = 16\,\mathrm{dB}$. In deployed metro or long-haul transmission systems, the optical signal is typically amplified every 50–$100\,\mathrm{km}$.

At the same time, signal amplification adds optical amplification noise to the signal which is the basic limit on the performance of optical transmission systems (in the *linear* transmission regime). Apparently, in the context of *nonlinear* signal propagation the power profile of the transmission link plays a key role, because nonlinear interactions occur essentially in the fiber segment where the signal power is *high.*

[1]Fiber attenuation is typically specified in terms of a (relative) power loss (in logarithmic scale) per distance, i.e., here in units of $\mathrm{dB/km}$. In the literature, e.g., in [Pog12], the attenuation coefficient $\alpha(z)$ is sometimes also defined as *field* attenuation coefficient where the factor $1/2$ in the gain/loss term of (3.21) is not present.

We assume that thermal and shot noise is negligible compared to amplified spontaneous emission (ASE) noise [Agr02, Ch. (6.1.3)], which is a reasonable assumption if we consider, e.g., an optically pre-amplified receiver (i.e., an optical receiver front-end which is preceded by an optical amplifier to boost the optical signal power before coherent reception). It is important to note, that due to coherent reception the statistics of noise phenomena in the optical domain are fully preserved, meaning a noise process with a circularly symmetric complex Gaussian distribution in the optical domain maintains its characteristics in the analog electrical domain.

Neglecting all terms in the Manakov equation from (3.21) except for the amplification and attenuation (i.e., gain and loss) term, we yield

$$\frac{\partial}{\partial z}\boldsymbol{u}(z,t) = \frac{g(z)-\alpha(z)}{2}\boldsymbol{u}(z,t)\,. \tag{3.23}$$

To describe the power evolution of the signal, we introduce the *normalized* power profile $\mathcal{P}(z)$ as a function that satisfies the equation [JK13, Eq. (7)] [Agr10, Eq. (7.1.7)]

$$\frac{\mathrm{d}\mathcal{P}(z)}{\mathrm{d}z} = \left(g(z)-\alpha(z)\right)\mathcal{P}(z)\,, \tag{3.24}$$

with boundary condition $\mathcal{P}(0) = \mathcal{P}(L) = 1$, i.e., the last optical amplifier resets the signal power to the transmit power[2]. The z-dependence of $\alpha(z)$ allows for varying attenuation coefficients over, e.g., different fiber segments or spans. We may also define the logarithmic gain/loss profile as

$$\mathcal{G}(z) \overset{\text{def}}{=} \ln\left(\mathcal{P}(z)\right) = \int_0^z \left(g(\zeta)-\alpha(\zeta)\right)\,\mathrm{d}\zeta\,. \tag{3.25}$$

The last expression in (3.25) is obtained by solving (3.24) for $\mathcal{P}(z) = \mathrm{e}^{\mathcal{G}(z)}$. The boundary conditions on $\mathcal{P}(z)$ immediately give the boundary condition $\mathcal{G}(0) = \mathcal{G}(L) = 0$.

The PDE in (3.23) is solved by

$$\boldsymbol{U}(z,\omega) = \boldsymbol{U}(0,\omega)\exp\left(\frac{\mathcal{G}(z)}{2}\right)\,, \tag{3.26}$$

i.e., the optical field at position z is simply obtained by an appropriate scaling of the input optical signal.

The *effective length* L_{eff} of a transmission link of length L evaluates to

$$L_{\text{eff}} \overset{\text{def}}{=} \int_0^L \mathcal{P}(\zeta)\,\mathrm{d}\zeta\,, \tag{3.27}$$

where L_{eff} equals the length of a fictitious *lossless* fiber with the same integrated power (i.e., causing the same nonlinear impact) as the total fiber link with power/gain profile $\mathcal{P}(z)$. We also define the effective length of a single span as

$$L_{\text{eff,sp}} \overset{\text{def}}{=} \int_0^{L_{\text{sp}}} \mathrm{e}^{-\bar{\alpha}\zeta}\,\mathrm{d}\zeta \tag{3.28}$$

$$= \frac{1-\mathrm{e}^{-\bar{\alpha}L_{\text{sp}}}}{\bar{\alpha}}\,, \tag{3.29}$$

[2] This is a convenient, but not required assumption for the formulation of end-to-end channel models.

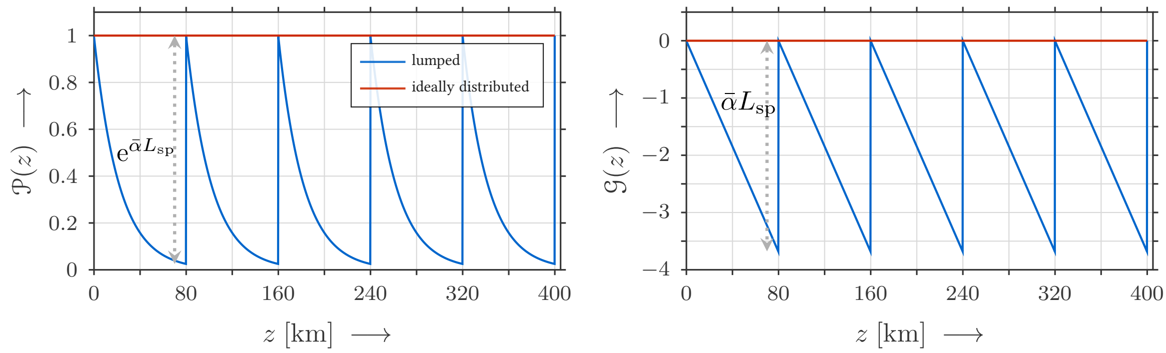


Figure 3.4: The gain/loss profile $\mathcal{P}(z)$ (left) and the logarithmic gain/loss profile $\mathcal{G}(z)$ (right) of a lumped and distributed optical amplification scheme for a 5×80 km link with path-invariant attenuation coefficient $10\log_{10}(e^{\bar{\alpha}}) = 0.2$ dB/km, and neglecting signal gain depletion.

with path-average attenuation coefficient $\bar{\alpha}$. For a link with homogeneous spans and end-of-span amplification we have $L_{\text{eff}} = N_{\text{sp}} L_{\text{eff,sp}}$.

For a fictitious single-span link of *infinite* length and path-invariant attenuation, we find that the effective length approaches its *asymptotic length* defined as

$$L_{\text{eff,a}} \stackrel{\text{def}}{=} \lim_{L_{\text{sp}}\to\infty} L_{\text{eff,sp}} = \frac{1}{\bar{\alpha}}\,. \tag{3.30}$$

For a lossless fiber span or for a very short fiber segment, we find that the effective length is (by definition) equivalent to the span length, i.e.,

$$\lim_{\bar{\alpha}\to 0} L_{\text{eff,sp}} = \lim_{L_{\text{sp}}\to 0} L_{\text{eff,sp}} = L_{\text{sp}}\,. \tag{3.31}$$

The quantities above will become relevant in the following Chapters when assesing the strength and temporal scale of some (first-order) perturbative terms.

Lumped Amplification In the absence of any distributed amplification the normalized power profile $\mathcal{P}(z)$ decreases exponentially between the lumped amplifier stages, see (3.26). A lumped fiber amplifier gain can be modeled by a δ-function in the local gain coefficient $g(z)$ at the position of each amplification stage along the link. Figure 3.4 shows the lumped amplification scheme for a homogeneous 5×80 km link with end-of-span amplification.

Fiber amplifiers can be driven in either *constant output power* or *constant gain* mode. In the constant power mode, the effective signal power is decreased gradually over a cascade of amplifiers since at every amplifier stage ASE noise is added to the signal. The additional noise itself is amplified at the subsequent amplifier stage. If the output power at each lumped fiber amplifier is kept constant, then the signal experiences a reduced gain due to the amount of ASE added in the previous stage. This phenomenon is called *signal-gain depletion*, cf. [GD91] [Gha17, Sec. II B.]. On the other hand, in constant gain mode, the overall power, i.e., the joint signal and noise power, in the system increases over the number of amplifier stages. We will neglect the effects of signal-gain depletion and only consider so-termed *transparent* spans where noise does not reduce the effective signal gain. Alternatively, the effective loss of signal power can be incorporated into the model by introducing a so-called *droop* exponent into the power profile $\mathcal{P}(z)$, cf. [BAL+20].

Assuming end-of-span amplification and a path-invariant attenuation coefficient with average value $\bar{\alpha}$, the signal attenuation is fully compensated by a (power) gain $\mathrm{e}^{\bar{\alpha}L_{\mathrm{sp}}}$ at the end of each span resulting in a gain profile $g(z)$ with

$$g(z) = \sum_{i=1}^{N_{\mathrm{sp}}} G_i\, \delta(z - z_i) \tag{3.32}$$

$$= \bar{\alpha} L_{\mathrm{sp}} \sum_{i=1}^{N_{\mathrm{sp}}} \delta(z - iL_{\mathrm{sp}})\,, \tag{3.33}$$

where (3.33) follows for homogeneous spans with equally-spaced amplifier stages at $z_i = iL_{\mathrm{sp}}$ and equal (logarithmic) power gain $G_i = \bar{\alpha}L_{\mathrm{sp}}$ with $i = 1, 2, \ldots, N_{\mathrm{sp}}$.

Amplification Noise Optical amplification inflicts the signal with ASE noise at each point along the link where an optical amplifier is placed. This can be represented in the Manakov equation by a (local) noise term $\boldsymbol{n}(z,t) = [n_{\mathsf{x}}(z,t), n_{\mathsf{y}}(z,t)]^{\mathsf{T}}$, where the noise is assumed, in accordance with the previous chapter, to be a zero-mean and complex circular Gaussian[3] in each polarization and spatially independent for each amplifier position.

Beside *signal-to-signal* nonlinear interference, *signal-to-noise* nonlinear interference is particularly relevant for systems with in-line dispersion compensation, but is beyond the scope of this work. We will consider only noiseless transmission of the signal and emulate, if necessary, the in-line ASE noise due to fiber amplifiers by a lumped noise source located at the end of the transmission link at $z = L$. The noise power of the equivalent noise source at the receiver is equal to the power of the accumulated noise of the individual sources. In doing so, signal-to-noise nonlinear interference is not considered.

The local noise autocorrelation function (ACF) of the lumped fiber amplifier at position z_i is defined as [Agr10, Eq. (7.1.5)]

$$\boldsymbol{\varphi}_{nn}(z_i, \tau) \overset{\text{def}}{=} \mathrm{E}\{\, \boldsymbol{n}(z_i, t+\tau)\, \boldsymbol{n}^{\mathsf{H}}(z_i, t)\,\} \tag{3.34}$$

$$= n_{\mathrm{sp},i}\, (\mathrm{e}^{G_i} - 1)\hbar\omega_0\, \delta(\tau)\, \mathbf{I}\,, \tag{3.35}$$

with the timing offset $\tau \in \mathbb{R}$ in *seconds*. Here, (3.35) follows for white noise with independent components in the x- and y-polarization centered at the passband frequency ω_0. The Planck constant is defined as $\hbar \overset{\text{def}}{=} \frac{\mathrm{h}}{2\pi}$ and $n_{\mathrm{sp},i}$ denotes the *spontaneous emission factor* [Agr10, Sec. 7.2.3] of the i^{th} optical amplifier.

Similar to (2.88), the Fourier transform of the noise ACF w.r.t. τ gives the local noise PSD

$$\boldsymbol{\Phi}_{nn}(z_i, \omega) = n_{\mathrm{sp},i}\, (\mathrm{e}^{G_i} - 1)\hbar\omega_0\, \mathbf{I} = \text{const.}, \qquad \forall\omega\,. \tag{3.36}$$

The noise PSD is (by assumption) constant, i.e., white over all frequencies (in the vicinity of ω_0, where the Taylor expansion and hence the Manakov equation is valid) and equally

[3] Note, that the transformation to the retarded time frame by $t = t' - \bar{\beta}_1 z$, or unitary transformation to the local birefringence coordinate system do not affect the Gaussianity or whiteness of the process and the following considerations still hold.

distributed over both polarizations. This noise model is justified because we assume that the OA bandwidth is much larger than the bandwidth of a single wavelength channel and also beyond the processing capabilities of any practical receiver. Note, however, that the noise PSD depends on the center frequency ω_0 which coincides with the carrier frequency of the ECB model, and is strictly speaking not white over the full WDM bandwidth.

The noise PSD accumulated over a cascade of independent, lumped amplifier stages is given by

$$\begin{aligned}\boldsymbol{\Phi}_{nn}(\omega) &= \sum_{i=1}^{N_{\mathrm{sp}}} \boldsymbol{\Phi}_{nn}(z_i, \omega) && (3.37)\\ &= N_{\mathrm{sp}} n_{\mathrm{sp}}\, (\mathrm{e}^{G} - 1)\hbar\omega_0\, \mathbf{I} && (3.38)\\ &= N_{\mathrm{ASE}}\, \mathbf{I} = \mathrm{const.}, \quad \forall \omega\,, && (3.39)\end{aligned}$$

where again (3.39) follows for homogeneous spans with equal logarithmic gain $G = G_i$ and spontaneous emission factor $n_{\mathrm{sp}} = n_{\mathrm{sp},i}\ \forall i$. Here, we also implicitly defined the PSD constant N_{ASE} for ASE noise (per polarization).

Given the previous assumptions to arrive at (3.39), the optical signal-to-noise ratio (OSNR) of the probe signal centered at the carrier $\omega_\rho = \omega_0$ is defined as [Kam13, Eq. (2.11)]

$$\mathsf{OSNR} \overset{\mathrm{def}}{=} \frac{P_\rho}{2N_{\mathrm{ASE}}\, B_{\mathrm{ref}}} = \frac{R_{\mathrm{s}}}{B_{\mathrm{ref}}} \mathsf{SNR}\,, \tag{3.40}$$

which reflects the ratio between the probe's signal power P_ρ and the ASE noise power in *both* polarizations within the resolution bandwidth (RBW). The latter is defined by convention as

$$\begin{aligned}B_{\mathrm{ref}} &\overset{\mathrm{def}}{=} \frac{\omega_0^2}{(2\pi)^2\, \mathrm{c}}\, (0.1\,\mathrm{nm}) = \frac{\mathrm{c}}{\lambda_0^2}\, (0.1\,\mathrm{nm}) && (3.41)\\ &\approx 12.5\ \mathrm{GHz}\Big|_{\lambda_0 = 1550\,\mathrm{nm}}\,, && (3.42)\end{aligned}$$

such that the ASE noise is always measured in a spectral bandwidth equivalent to $0.1\,\mathrm{nm}$.

The relation between OSNR and SNR in (3.40) follows for a matched filter receiver, as discussed in Section 2.2.3, with $N_0 = N_{\mathrm{ASE}}$.

In Example 3.4, a simple approximation to estimate the received OSNR for a homogeneous link is given.

Example 3.4: OSNR approximation

At the telecommunication wavelength of $\lambda_0 = 1550$ nm, we can approximate the OSNR given in dB at the end of a homogeneous transmission link as [ZNK97, (2.8)]

$$10\log_{10}(\mathsf{OSNR}) \approx 58 + 10\log_{10}(P_\rho) - 10\log_{10}(\mathit{\Gamma}_{\mathrm{EDFA}}) - 10\log_{10}(\mathrm{e}^{\bar{\alpha}L_{\mathrm{sp}}}) - 10\log_{10}(N_{\mathrm{sp}})\,,$$

where we used that the *shot noise limit* at $\omega_0 = 2\pi\mathrm{c}/\lambda_0$ within the RBW of 12.5 GHz is

$$10\log_{10}(\hbar\omega_0 B_{\mathrm{ref}}) \approx -58\ \mathrm{dBm}\big|_{\lambda_0 = 1550\,\mathrm{nm}}\,.$$

We also assume that the amplifier gain $\mathrm{e}^G = \mathrm{e}^{\bar{\alpha}L_{\mathrm{sp}}} \gg 1$ and approximate the noise figure of an EDFA with

$$\mathit{\Gamma}_{\mathrm{EDFA}} \approx 2n_{\mathrm{sp}} \geq 2\,,$$

i.e., the amplifier noise figure is at least $10\log_{10}(2) = 3$ dB.
For an optical input power of $10\log_{10}(P_\rho) = 0\,\mathrm{dBm}$, an EDFA noise figure of $10\log_{10}(\mathit{\Gamma}_{\mathrm{EDFA}}) = 6$ dB, a total loss between two amplifier stages of $10\log_{10}(\mathrm{e}^{\bar{\alpha}L_{\mathrm{sp}}}) = 20$ dB and $N_{\mathrm{sp}} = 10$ spans, one arrives at $10\log_{10}(\mathsf{OSNR}) = 22$ dB.

3.3.2 Chromatic Dispersion Profile

To derive the transfer function of chromatic dispersion (CD) we drop all terms in the Manakov equation (3.21) that are not related to CD to obtain

$$\frac{\partial}{\partial z}\boldsymbol{u}(z,t) = \mathrm{j}\frac{\beta_2(z)}{2}\frac{\partial^2}{\partial t^2}\boldsymbol{u}(z,t) \tag{3.43}$$

$$\circ\!\!-\!\!\bullet$$

$$\frac{\partial}{\partial z}\boldsymbol{U}(z,\omega) = -\mathrm{j}\frac{\beta_2(z)}{2}\omega^2\boldsymbol{U}(z,\omega)\,, \tag{3.44}$$

with the correspondence $\frac{\partial^n}{\partial t^n} \circ\!\!-\!\!\bullet\ (\mathrm{j}\omega)^n$.

The result in (3.44) is a linear, homogeneous, first-order differential equation and can be solved using the method of *integrating factor* [RE10, P. 449]. It yields

$$\boldsymbol{U}(z,\omega) = \boldsymbol{U}(0,\omega)\exp\left(-\mathrm{j}\omega^2\frac{\mathcal{B}(z)}{2}\right)\,, \tag{3.45}$$

where the accumulated dispersion $\mathcal{B}(z)$ is a function that satisfies [JK13, Eq. (8)]

$$\frac{\mathrm{d}\mathcal{B}(z)}{\mathrm{d}z} = \beta_2(z)\,. \tag{3.46}$$

Here, $\mathcal{B}(z)$ can be used to express a z-dependency in the dispersion profile, e.g., lumped in-line dispersion compensation or distinct fiber properties across a multi-span link. We obtain

$$\mathcal{B}(z) = \int_0^z \beta_2(\zeta)\mathrm{d}\zeta + \mathcal{B}_0\,, \tag{3.47}$$

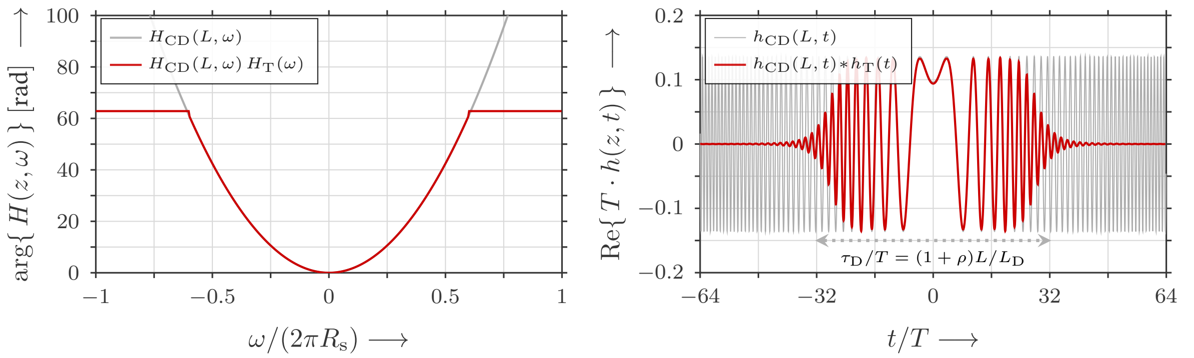


Figure 3.5: Chromatic dispersion transfer function in the equivalent complex baseband (left) and impulse response (right) for $R_{\mathrm{s}} = 64$ GBd, RRC pulse shape with roll-off $\rho = 0.1$, $\bar{\beta}_2 = -21\ \mathrm{ps}^2/\mathrm{km}$ and $L = 100$ km. Due to the phase unwrap of the transfer function (right) the phase remains constant for frequencies outside the spectral support of the communication signal.

where $\mathcal{B}_0 \overset{\mathrm{def}}{=} \mathcal{B}(0)$ is the amount of *pre-dispersion* (in units of squared *seconds*, typically given in ps^2) at the beginning of the transmission line. Electronic pre-dispersion, as a means of pre-compensation, is a common method to mitigate distortions due to nonlinear interactions.

If we assume that the GVD parameter is invariant under z with path-average value $\bar{\beta}_2$, and we discard pre-dispersion, i.e., $\mathcal{B}_0 = 0$, than we have the simple relation $\mathcal{B}(z) = \bar{\beta}_2 z$.

We define the transfer function and impulse response of CD as

$$H_{\mathrm{CD}}(z,\omega) \overset{\mathrm{def}}{=} \exp\left(-\mathrm{j}\omega^2 \frac{\mathcal{B}(z)}{2}\right) \tag{3.48}$$

$$\bullet\!\!-\!\!\circ$$

$$h_{\mathrm{CD}}(z,t) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{\mathrm{j}\mathcal{B}(z)}} \exp\left(\mathrm{j}t^2 \frac{1}{2\mathcal{B}(z)}\right). \tag{3.49}$$

The transfer function $H_{\mathrm{CD}}(z,\omega)$ is an all-pass, i.e., $|H_{\mathrm{CD}}(z,\omega)| = 1$.

The phase response, $\arg\{H_{\mathrm{CD}}(z,\omega)\}$, is shown in Figure 3.5 for the system parameters $\bar{\beta}_2 = -21\ \mathrm{ps}^2/\mathrm{km}$, $R_{\mathrm{s}} = 64$ GBd and $L = 100$ km. From the unwrapped phase response (jumps between consecutive phase angles greater than $\pm\pi$ are corrected by adding appropriate multiples of $\pm 2\pi$) the quadratic dependency of $\arg\{H_{\mathrm{CD}}(z,\omega)\}$ w.r.t. the angular frequency ω is observed. The cascade of $H_{\mathrm{CD}}(z,\omega)$ and the transmit pulse shape $H_{\mathrm{T}}(\omega) = H_{\mathrm{RRC}}(\omega)$ is also shown which is the relevant part of the response for intra-channel effects of CD on the probe channel in the baseband description.

The impulse response $h_{\mathrm{CD}}(z,t)$ is complex valued, since $H_{\mathrm{CD}}(z,\omega)$ does not satisfy the symmetry condition of an odd phase. Moreover, from (3.49) it is apparent that $h_{\mathrm{CD}}(z,t)$ has an infinite impulse response. The real part $\mathrm{Re}\{h_{\mathrm{CD}}(z,t)\}$ is shown in Figure 3.5 (right). The joint impulse response of the cascade $h_{\mathrm{CD}}(z,t) * h_{\mathrm{RRC}}(t)$ is also shown. The width of the joint impulse response for this specific example spans approximately 64 symbol periods. In the next paragraph, we will see that the length of the (intra-channel) CD impulse response (and equivalently the CD compensation filter given a straightforward DSP implementation) scales linearly in the transmission distance z and quadratical in the symbol rate R_{s}.

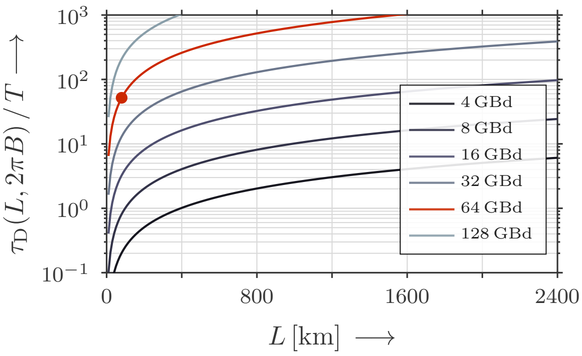


Figure 3.6: Temporal width $\tau_{\rm D}(L, 2\pi B)/T$ in logarithmic scale of the chromatic dispersion impulse response $h_{\rm CD}(z,t)$ over the transmission distance L (for $\bar{\beta}_2 = -21\ \text{ps}^2/\text{km}$ and RRC pulse-shape with $B = (1+\rho)R_{\rm s}$ and $\rho = 0.2$). The red marker ● coincides with the example in Figure 3.5.

Intra-channel Dispersion Response By introducing the retarded time in (3.11), the group delay at the center frequency ω_0, i.e., the $\beta_1(z)$-term in the Manakov equation is removed. However, all other frequencies experience a residual group delay relative to the reference frequency (or in baseband, relative to the zero-frequency) due to GVD. In this paragraph, we investigate the effect of GVD on a single wavelength channel.

To quantify the relative delay between two spectral components at ω_0 and $\omega_0 + \upsilon$ accumulated over a transmission length L, we can use the Taylor expansion of the path-average propagation constant $\bar{\beta}(\omega) = \frac{1}{L}\int_0^L \beta(\zeta,\omega)\mathrm{d}\zeta$. Since we focus on the intra-channel effect of CD, we assume that the frequency deviation υ is within the spectral support of the probe channel, i.e., $\upsilon \in \mathcal{B}$. By definition we find the magnitude of the accumulated differential propagation delay [KD14, Eq. (2.193)]

$$\tau_{\rm D}(z,\upsilon) \overset{\text{def}}{=} \text{abs}\left(\frac{z}{\mathrm{v_g}(\omega_0+\upsilon)} - \frac{z}{\mathrm{v_g}(\omega_0)}\right) \tag{3.50}$$

$$= \text{abs}\left(z\frac{\partial\bar{\beta}(\omega)}{\partial\omega}\bigg|_{\omega=\omega_0+\upsilon} - z\frac{\partial\bar{\beta}(\omega)}{\partial\omega}\bigg|_{\omega=\omega_0}\right) \tag{3.51}$$

$$= \text{abs}\left(z\,\upsilon\,\frac{\partial^2\bar{\beta}(\omega)}{\partial\omega^2}\bigg|_{\omega=\omega_0}\right) = z|\upsilon\bar{\beta}_2|\,. \tag{3.52}$$

Note, the reference frequency ω_0 equals zero in the ECB representation and coincides with the center frequency of the probe channel ω_ρ. The accumulated temporal walk-off between two spectral components at ω_0 and $\omega_0+\upsilon$ in terms of the *local* GVD coefficient is given by

$$\tau_{\rm D}(z,\upsilon) = \left|\upsilon\int_0^z \beta_2(\zeta)\,\mathrm{d}\zeta\right|\,. \tag{3.53}$$

We can now approximate the temporal length of the intra-channel CD impulse response at $z = L$ and $\upsilon = 2\pi B$, i.e., the delay spread of the frequency components within the probe's spectral support B at the receiver by [IK07, Eq. (30)] [Spi10, Eq. (9)]

$$\tau_{\rm D}(L, 2\pi B) = 2\pi|\bar{\beta}_2|BL = 2\pi|\bar{\beta}_2|R_{\rm s}(1+\rho)L\,, \tag{3.54}$$

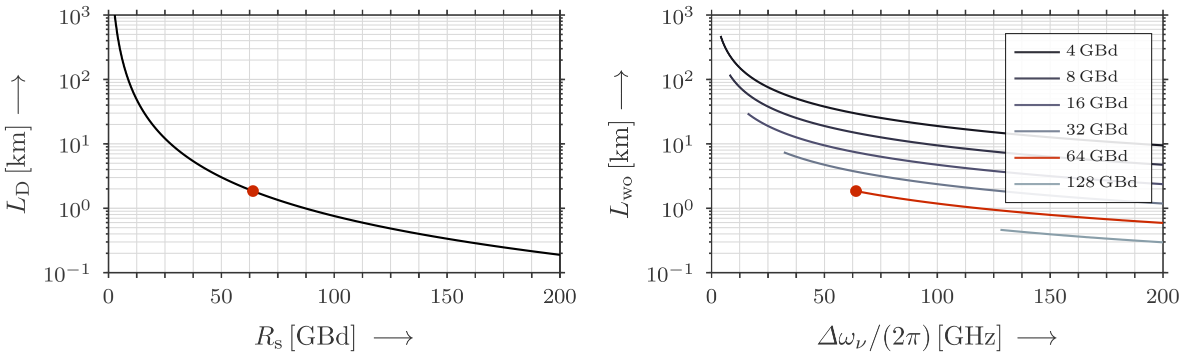


Figure 3.7: Dispersion length L_D over symbol rate R_s (left) and walk-off length $L_{wo,\nu}$ over channel separation $\Delta\omega_\nu/(2\pi)$ (right) in logarithmic scale for $\bar{\beta}_2 = -21\ \mathrm{ps}^2/\mathrm{km}$. The dispersion length L_D for a single wavelength channel with $R_s = 64\,\mathrm{GBd}$ is about $1.85\,\mathrm{km}$ (indicated by •), which is equal to the walk-off length L_{wo} of two $64\,\mathrm{GBd}$ wavelength channels with RRC pulse-shape and $\rho = 0$, and Nyquist channel separation of $\Delta\omega/(2\pi) = 64\,\mathrm{GHz}$ (starting point of the red curve, right).

where the maximum frequency separation within the probe channel corresponds to the spectral bandwidth of the probe $B = R_s(1+\rho)$. Normalized to the symbol period, the temporal width of the impulse response scales with [Spi10, Eq. (9)]

$$\frac{\tau_D(L, 2\pi B)}{T} = 2\pi|\bar{\beta}_2|R_s^2(1+\rho)L\,, \tag{3.55}$$

which is also the approximate number of overlapping (intra-channel) basic pulses. In Figure 3.5, the normalized temporal extend $\tau_D(L, 2\pi B)/T$ of about 64 symbol durations is shown by the dashed arrow. Figure 3.7 shows the dependency of $\tau_D(L, 2\pi B)/T$ on the symbol rate R_s and the transmission distance L in logarithmic scale.

We introduce the path-average *dispersion length*

$$L_D \stackrel{\text{def}}{=} \frac{1}{2\pi|\bar{\beta}_2|R_s^2}\,, \tag{3.56}$$

which denotes the distance after which two spectral components spaced $B = R_s$ *Hertz* apart, experience a differential group delay of $T = 1/R_s$ due to CD. Note, that in contrast to many textbooks (e.g., [Agr10, Eq. (9.1.10)] or [KL02, Eq. (6.18)]) we include the scaling factor $1/(2\pi)$. This allows us to re-write the length of the impulse response from (3.55) as

$$\frac{\tau_D(L, 2\pi B)}{T} = (1+\rho)\frac{L}{L_D}\,. \tag{3.57}$$

The dispersion length L_D is also often used as a normalization constant to obtain a Manakov equation with dimensionless variables, cf. [Agr10, Eq. (9.1.11)]. In Figure 3.7, the dependency of the dispersion length w.r.t. the symbol rate R_s is shown for $\bar{\beta}_2 = -21\ \mathrm{ps}^2/\mathrm{km}$.

Example 3.5: Intra-channel dispersion after multiples of L_D

Figure 3.8 and 3.9 show how the probe's optical transmit signal is dispersed after propagating multiples of the dispersion length L_D.
In this simplified setting, we only consider the x-polarization modulated with symbols from the set $\mathcal{A} = \{0, 1\}$, i.e., *on-off keying*. The basic pulse $h_{\mathrm{T},\rho}(t)$ has an RRC shape with roll-off factor $\rho = 0.2$.
The optical baseband transmit signal of the probe at $z = 0$ is given by

$$u_\mathsf{x}(0,t) = s_\mathsf{x}(t) = T \cdot \sum_{k\in\mathbb{Z}} a_\mathsf{x}[k]\, h_{\mathrm{T},\rho}(t-kT)\,.$$

The optical signal subject to (3.44) is dispersed after propagation, mathematically expressed by the convolution of the transmit signal with the impulse response $h_\mathrm{CD}(z,t)$, i.e.,

$$\begin{aligned} u_\mathsf{x}(z,t) &= h_\mathrm{CD}(z,t) * u_\mathsf{x}(0,t) \\ &= \int_\mathbb{R} h_\mathrm{CD}(z,\tau)\, u_\mathsf{x}(0,t-\tau)\,\mathrm{d}\tau \\ &= T \cdot \sum_{k\in\mathbb{Z}} a_\mathsf{x}[k] \int_\mathbb{R} h_\mathrm{CD}(z,\tau) h_{\mathrm{T},\rho}(t-\tau-kT)\,\mathrm{d}\tau \\ &= \sum_{k\in\mathbb{Z}} a_\mathsf{x}[k]\, \underbrace{T \cdot h_\mathrm{CD}(z,t-kT) * h_{\mathrm{T},\rho}(t-kT)}_{\tilde{g}_{\mathrm{T},\rho}(z,t-kT)}\,, \end{aligned}$$

where the order of summation and integration can be exchanged to obtain the now z-dependent, *dispersed* transmit pulse $\tilde{g}_{\mathrm{T},\rho}(z,t)$. The optical signal $u_\mathsf{x}(z,t)$ can be interpreted as a weighted sum of the transmit symbols $a_\mathsf{x}[k]$ and shifted (by integer multiples of T) dispersed basic pulse $\tilde{g}_{\mathrm{T},\rho}(z,t)$ depending on the transmission distance z. This view will become relevant in understanding the so-termed *pulse collision picture* in the following chapter.
The dispersed signal is shown at $z = \{2, 3, 4, 16\}\, L_\mathrm{D}$. The width of the effective transmit pulse scales approximately with z/L_D.

Inter-channel Walk-off Similar to in-band (i.e., within the bandwidth of the probe signal) dispersive effects due to CD in the previous paragraph, the out-of-band effect is studied, that is, how CD affects propagation of multiple wavelength channels (see also Example 3.3). Similar to (3.53), we define the accumulated *temporal walk-off* between the probe channel and a co-propagating wavelength channel as

$$\tau_\mathrm{wo}(z, \Delta\omega_\nu) \overset{\mathrm{def}}{=} \Delta\omega_\nu \mathcal{B}(z)\,, \tag{3.58}$$

where $\Delta\omega_\nu = \omega_\nu - \omega_\rho$ is equivalent to the channel spacing between the wavelength channel and the probe. From (3.58) it is obvious that $\tau_\mathrm{wo}(z, \Delta\omega_\nu)$ can take positive or negative values, e.g., in case of anomalous dispersion where $\bar{\beta}_2 < 0$ and assuming a frequency separation between a wavelength channel and the probe $\Delta\omega_\nu > 0$ (to higher frequencies), we have a relative group delay $\tau_\mathrm{wo} < 0$. We conclude that in the anomalous dispersion regime, a wavelength channel with positive frequency offset (or smaller wavelength) w.r.t. the probe channel has a *reduced* latency, and vice-versa a wavelength channel with negative frequency offset has an *increased* latency w.r.t. the probe.

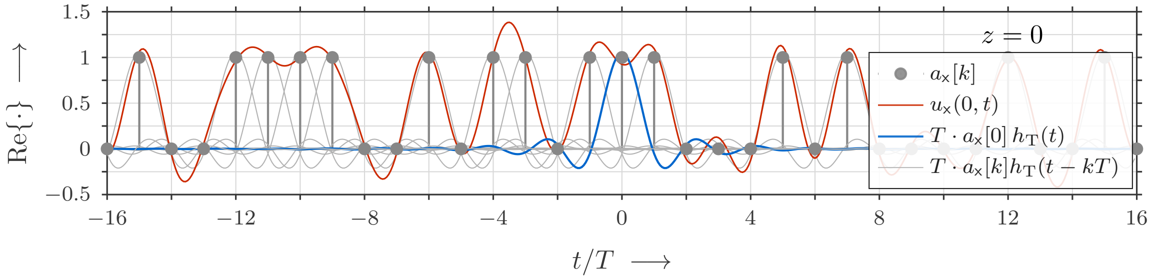


Figure 3.8: Transmit-side optical field envelope $u_\mathsf{x}(0,t)$ of a single channel (and single polarization) with modulation $a_\mathsf{x} \in \{0, 1\}$ and basic pulse $h_\mathrm{T}(t)$ using a root-raised cosine (RRC) with roll-off $\rho = 0.2$.

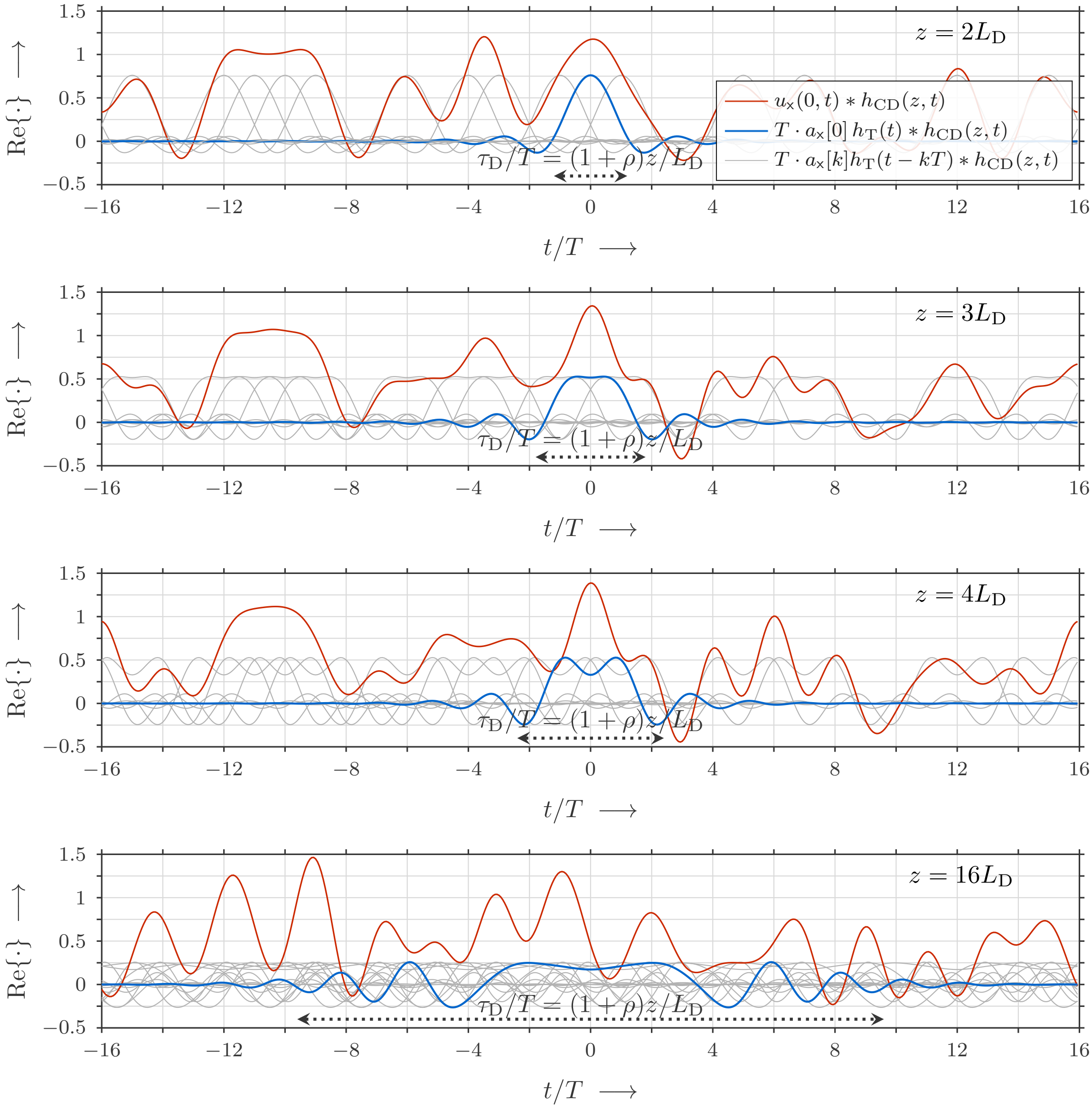


Figure 3.9: *Dispersed* optical field envelope $u_\mathsf{x}(z,t)$ and dispersed basic pulses after $z = \{2, 3, 4, 16\}\, L_\mathrm{D}$. The ratio z/L_D is a measure of the *temporal* width of $h_\mathrm{T}(t) * h_\mathrm{CD}(z,t)$, and approximates the number of *overlapping* pulses at each $k \in \mathbb{N}$. Due to the linearity of CD, the dispersed optical envelope is the sum over all dispersed (and shifted) basic pulses weighted with $a_\mathsf{x}[k]$, i.e., convolution with $h_\mathrm{CD}(z,t)$ and $\sum_k$ can be exchanged.

The (retarded) time of the ν^{th} wavelength channel relative to the probe channel (i.e., relative to the Taylor expansion of $\beta(z,\omega)$ at ω_0) is given by

$$t_\nu = t - \tau_{\text{wo}}(z, \Delta\omega_\nu) = t - \Delta\omega_\nu \mathcal{B}(z)\,, \tag{3.59}$$

which provides another view on the channel walk-off. The ν^{th} wavelength channel has hence a z-dependent time base at which the group delay is already canceled from the propagation equation.

Yet another view on the relative retardation between wavelength channels can be obtained by taking the inverse Fourier transform of the CD transfer function *shifted* to the baseband of the ν^{th} interferer, i.e., by the negative channel separation $-\Delta\omega_\nu$. We obtain

$$H_{\text{CD}}(z, \omega + \Delta\omega_\nu) = \exp\left(-\mathrm{j}(\omega + \Delta\omega_\nu)^2 \frac{\mathcal{B}(z)}{2}\right) \tag{3.60}$$

$$\bullet\!-\!\circ$$

$$h_{\text{CD}}(z,t)\ \mathrm{e}^{-\mathrm{j}\Delta\omega_\nu t} = h_{\text{CD}}(z, t - \Delta\omega_\nu \mathcal{B}(z))\ \mathrm{e}^{-\mathrm{j}\mathcal{B}(z)\Delta\omega_\nu^2}\,, \tag{3.61}$$

where we can recover the additional retardation by $\tau_{\text{wo}}(z, \Delta\omega_\nu) = \Delta\omega_\nu \mathcal{B}(z)$ in the argument of the modified chromatic dispersion impulse response. We follow that an interfering wavelength channel at relative frequency offset $\Delta\omega_\nu$ is subject to the (un-shifted) chromatic dispersion impulse response $h_{\text{CD}}(z,t)$, is additionally retarded by $\tau_{\text{wo}} = \Delta\omega_\nu \mathcal{B}(z)$, and phase rotated by the (time-independent) factor $\mathrm{e}^{-\mathrm{j}\mathcal{B}(z)\Delta\omega_\nu^2}$.

We can now introduce the walk-off length $L_{\text{wo},\nu}$ which measures the distance a wavelength channel at frequency ω_ν must co-propagate with the probe channel at $\omega_\mathrm{p} = \omega_0$ to walk-off by *one* symbol period (i.e., by $T = 1/R_\mathrm{s}$, assuming the same symbol rate in both wavelength channels),

$$L_{\text{wo},\nu} \overset{\text{def}}{=} \frac{1}{|\Delta\omega_\nu \bar{\beta}_2| R_\mathrm{s}}\,. \tag{3.62}$$

The walk-off length $L_{\text{wo},\nu}$ versus channel spacing $\Delta\omega_\nu$ is shown in Figure 3.7 for different symbol rates R_s.

Example 3.6: Inter-channel walk-off at multiples of L_{wo}

The walk-off between the probe signal $u_{\mathsf{x},\mathsf{p}}(z,t)$ and a co-propagating wavelength signal $u_{\mathsf{x},1}(z,t)$ is visualized in Figure 3.10 and 3.11.
Similar as in Example 3.5, the probe signal and the co-propagating wavelength signal are visualized in their respective ECB representation. We assume that both signals are modulated with symbol rate R_{s} and RRC pulse-shape with roll-off factor $\rho = 0.2$. To model the joint propagation, the co-propagating wavelength signal is mixed to its carrier frequency $\Delta\omega_1 = 2\pi \cdot 1.2R_{\mathrm{s}}$ relative to the probe at

$$u_{\mathsf{x},1}(0,t) = s_{\mathsf{x},1}(t)\exp(\mathrm{j}\Delta\omega_1 t) = T \cdot \sum_{k\in\mathbb{Z}} b_{\mathsf{x},1}[k]\, h_{\mathrm{T},1}(t-kT)\exp(\mathrm{j}\Delta\omega_1 t)\,.$$

The dispersed signal $u_{\mathsf{x},1}(z,t)$ with relative frequency offset $\Delta\omega_1$ can be expressed as

$$\begin{aligned}
u_{\mathsf{x},1}(z,t) &= h_{\mathrm{CD}}(z,t) * u_{\mathsf{x},1}(0,t) \\
&= \int_{\mathbb{R}} h_{\mathrm{CD}}(z,\tau)\, u_{\mathsf{x},1}(0,t-\tau)\,\mathrm{d}\tau \\
&= T \cdot \sum_{k\in\mathbb{Z}} b_{\mathsf{x},1}[k] \int_{\mathbb{R}} h_{\mathrm{CD}}(z,\tau) h_{\mathrm{T},1}(t-\tau-kT)\ \exp(\mathrm{j}\Delta\omega_1(t-\tau))\,\mathrm{d}\tau \\
&= \sum_{k\in\mathbb{Z}} b_{\mathsf{x},1}[k]\, \underbrace{T\ \exp(-\mathrm{j}\mathcal{B}(z)\Delta\omega_1^2)\cdot h_{\mathrm{CD}}(z,t-kT-\Delta\omega_1\mathcal{B}(z)) * h_{\mathrm{T},1}(t-kT)}_{\tilde{g}_{\mathrm{T},1}(z,t-kT)} \\
&\qquad \times \exp(\mathrm{j}\Delta\omega_1 t)\,,
\end{aligned}$$

where we use (3.61) to explicitly express the temporal retardation within the argument of the CD impulse response.
Similarly as in the previous example, we obtain the z-dependent, *dispersed* (baseband) transmit pulse of the interferer $\tilde{g}_{\mathrm{T},1}(z,t)$, i.e., compared to the previous example, the interferer is affected by a delayed basic pulse, and the time-independent exponential $\mathrm{e}^{-\mathrm{j}\mathcal{B}(z)\Delta\omega_1^2}$.
Assuming $\mathcal{B}(z) = \bar{\beta}_2 z$, both probe and interferer signals are dispersed at $z = 4L_{\mathrm{wo},1}$, and relatively delayed with respect to each other by

$$\Delta\omega_1\mathcal{B}(4L_{\mathrm{wo},1}) = \Delta\omega_1\bar{\beta}_2\ 4L_{\mathrm{wo},1} = 4\,T\,.$$

E.g., for $R_{\mathrm{s}} = 64\,\mathrm{GBd}$ and $\bar{\beta}_2 = -21\,\mathrm{ps}^2/\mathrm{km}$, the walk-off length calculates to

$$L_{\mathrm{wo},1} = 1.54\,\mathrm{km}\,.$$

Hence, in Figure 3.11, the temporal walk-off between the two signals after $4L_{\mathrm{wo},1}$ is equivalent to 4 symbol periods. Due to the anomalous dispersion ($\bar{\beta}_2 < 0$) and the *positive* frequency offset ($\Delta\omega_1 > 0$), the signal $\boldsymbol{u}_1(z,t)$ has a reduced latency, i.e., travels *faster* over the dispersive fiber. The group-delay free time base of the co-propagating wavelength signal t_1 (see top horizontal axis) is shifted by $4T$ with respect to the time base of the probe signal t, see (3.59).
The intra-channel effect of the CD response on the interferer is exactly the same as for the probe signal, i.e., the fundamental pulses, shown in blue, are subject to the same amount of pulse spread (since we neglect the dispersion slope $\bar{\beta}_3$) but are relatively delayed with respect to each other. This view is also in agreement with (3.60)–(3.61).

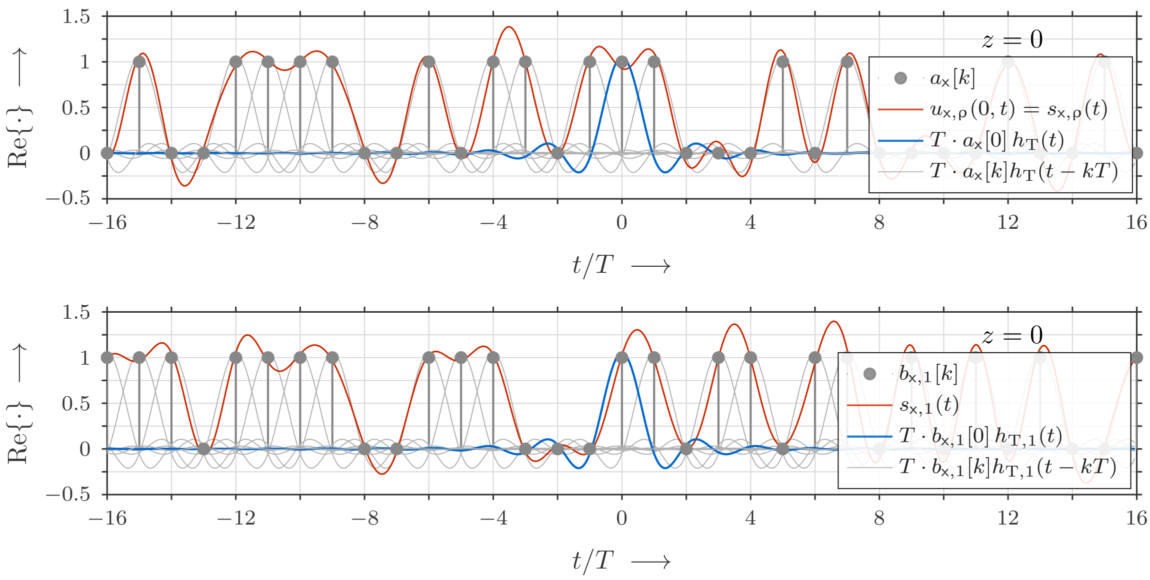


Figure 3.10: Optical field envelope of the probe $u_{\mathrm{x},\rho}(0,t)$ and a co-propagating wavelength signal $u_{\mathrm{x},1}(0,t) = s_{\mathrm{x},1}(t)\exp(\mathrm{j}\Delta\omega_1 t)$ at $z = 0$. Both signals are shown in their respective ECB, i.e., before multiplexing. The probe channel is modulated with $a_\mathrm{x}[k]$ at the reference frequency $\omega_\rho = \omega_0$, the interfering wavelength channel is modulated with $b_{\mathrm{x},1}[k]$ and will be mixed relative to the probe with frequency separation $\Delta\omega_1 = 2\pi \cdot 1.2 R_\mathrm{s}$ in baseband. In both cases, the basic pulse $h_\mathrm{T}(t)$ has an RRC shape with roll-off $\rho = 0.2$.

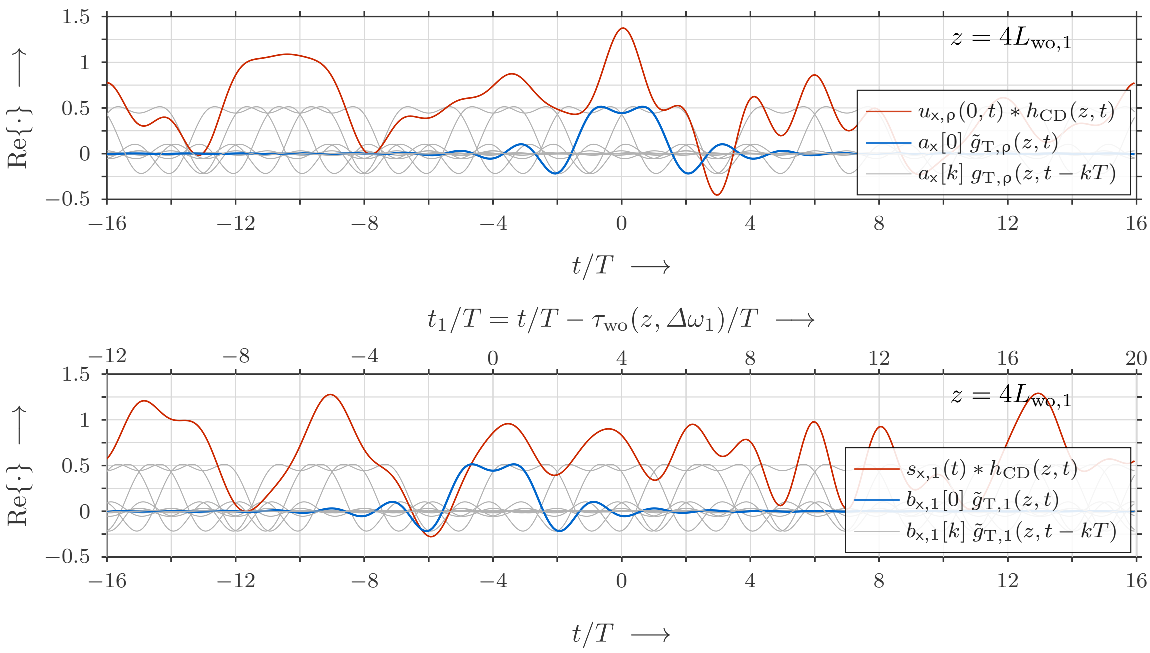


Figure 3.11: *Dispersed* optical field envelope and dispersed basic pulses after $z = 4L_{\mathrm{wo},1}$. The ratio $z/L_{\mathrm{wo},1}$ is a measure of the *temporal* walk-off between the probe channel and the interfering wavelength channel, and quantifies the number of *traversed* pulses. Vice-versa, the (normalized) reference time frame of the interfering wavelength channel t_1/T is retarded by $\Delta\omega_1/T \int_0^z \beta_2(\zeta)\mathrm{d}\zeta = z/L_{\mathrm{wo},1}$ w.r.t. the time frame of the probe t/T.

3.3.3 Linear Channel Transfer Function

We can now define the impulse response and transfer function of the *linear* channel—that is, when the fiber nonlinearity coefficient is zero, i.e., $\gamma = 0$ in (3.21). To that end, we define the optical field envelope $\boldsymbol{u}_{\text{LIN}}(z,t) \circ\!\!-\!\!\bullet \boldsymbol{U}_{\text{LIN}}(z,\omega)$ that propagates solely according to linear effects with the boundary condition $\boldsymbol{u}_{\text{LIN}}(0,t) = \boldsymbol{u}(0,t)$ at the input of the transmission link. The *linear* channel transfer function and impulse response is then given by

$$H_{\text{C}}(z,\omega) \overset{\text{def}}{=} \exp\left(\frac{\mathcal{G}(z) - \mathrm{j}\omega^2\mathcal{B}(z)}{2}\right) \tag{3.63}$$

$$\updownarrow$$

$$h_{\text{C}}(z,t) = \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{\mathrm{j}\mathcal{B}(z)}}\exp\left(\frac{\mathcal{G}(z) + \mathrm{j}t^2/\mathcal{B}(z)}{2}\right), \tag{3.64}$$

which represents the joint effect of chromatic dispersion and the gain/loss variation along the link.

We may also calculate the useful *inverse* of the channel transfer function and impulse response as

$$H_{\text{C}}^{-1}(z,\omega) = \exp(-\mathcal{G}(z))\; H_{\text{C}}^{*}(z,\omega) \tag{3.65}$$

$$\updownarrow$$

$$\mathcal{F}^{-1}\{\, H_{\text{C}}^{-1}(z,\omega)\,\} = \frac{1}{2\pi|\mathcal{B}(z)|} h_{\text{C}}^{-1}(z,t)\,, \tag{3.66}$$

where we use $\sqrt{\mathrm{j}\mathcal{B}(z)}\sqrt{-\mathrm{j}\mathcal{B}(z)} = |\mathcal{B}(z)|$ with $\mathcal{B}(z) \in \mathbb{R}$. In analogy we find

$$h_{\text{C}}^{*}(z,t) = h_{\text{C}}^{*}(z,-t) \circ\!\!-\!\!\bullet\, H_{\text{C}}^{*}(z,\omega) = H_{\text{C}}^{*}(z,-\omega)\,, \tag{3.67}$$

which will be used in the next chapter in the context of the first-order perturbation method.

3.3.4 Optimum Receive Filter for the Linear Fiber Channel

Now, that we defined the *linear* channel impulse response and transfer function, the optical receiver front-end matched to the *linear* fiber-optical channel is derived.

Again, we assume ideal optical-to-electrical (O/E) and analog-to-digital (A/D) conversion. The received continuous-time, optical signal $\boldsymbol{u}(L,t)$ is first matched filtered w.r.t. the *linear* channel response and transmit pulse, and then sampled at the symbol period T to obtain the discrete-time receive symbols $\boldsymbol{y}[k]$. The receiver front-end is shown in Figure 3.12. It compensates for any residual link loss and performs perfect CD compensation. Note, that the analog front-end is usually realized using an oversampled digital representation. E.g., CD compensation is typically performed in the (oversampled) domain of receiver DSP. Here, we prefer to conceptually incorporate it in the continuous-time domain since it significantly simplifies derivation of the channel model.

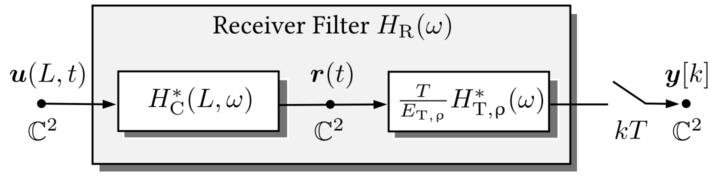


Figure 3.12: Receiver front-end matched to the *linear* optical channel transfer function $H_\mathrm{C}(L,\omega)$ and the transmit basic pulse of the probe $H_{\mathrm{T},\rho}(\omega)$. Matching to the channel transfer function at $z = L$ corresponds to the compensation of residual accumulated chromatic dispersion, i.e., $|H_\mathrm{C}(L,\omega)| = 1$ and $H^*_\mathrm{C}(L,\omega) = H^*_\mathrm{CD}(L,\omega) = H^{-1}_\mathrm{CD}(L,\omega)$. The pre-factor $T/E_{\mathrm{T},\rho}$ re-normalizes the (noiseless part of the) received sequence $\boldsymbol{y}[k]$ to the variance of the signal constellation σ^2_a.

In accordance with the definition of the linear receiver front-end defined in (2.91)–(2.92), the transfer function of the entire cascade of the receiver front-end is given by

$$H_\mathrm{R}(\omega) \overset{\mathrm{def}}{=} \frac{T}{E_{\mathrm{T},\rho}} H^*_\mathrm{C}(L,\omega) H^*_{\mathrm{T},\rho}(\omega) = \frac{1}{E_{\mathrm{T},\rho}} G_\mathrm{R}(\omega) \tag{3.68}$$

$$\bullet\!\!-\!\!\circ$$

$$h_\mathrm{R}(t) = \frac{T}{E_{\mathrm{T},\rho}} h^*_\mathrm{C}(L,-t) * h^*_{\mathrm{T},\rho}(-t) = \frac{1}{E_{\mathrm{T},\rho}} g_\mathrm{R}(t)\,, \tag{3.69}$$

with the (dimensionless) receiver pulse shape $g_\mathrm{R}(t) \circ\!\!-\!\!\bullet\, G_\mathrm{R}(\omega)$. Due to $\mathcal{P}(L) = 1$ and the pre-factor $T/E_{\mathrm{T},\rho}$, the noiseless part of the received signal is scaled to the variance of the constellation σ^2_a. Since we only consider T-spaced sampling any fractional sampling phase-offset or timing synchronization[4] is already incorporated as suited delay in the receive filter $h_\mathrm{R}(t)$, s.t. the transmitted and received sequence of the probe are perfectly aligned in time.

3.3.5 Fiber Nonlinearity

The nonlinear effect considered in this thesis is the *optical Kerr effect* [Agr06, Ch. 6]. The Kerr effect is directly related to the *real* part of the material polarization (cf. [Agr06, Ch. 1.3, and Ch. 10]) and results in a local refractive index change depending on the intensity of the electric field in the fiber. On the other hand, the *imaginary* part of the material polarization is related to stimulated Raman scattering, not considered in this work.

The Kerr effect is reflected by the last term in the Manakov equation (3.21). The expression depends on the product

$$\|\boldsymbol{u}(z,t)\|^2\, \boldsymbol{u}(z,t) = \left(|u_\mathsf{x}(z,t)|^2 + |u_\mathsf{y}(z,t)|^2\right) \boldsymbol{u}(z,t)\,, \tag{3.70}$$

and on the nonlinear coefficient

$$\gamma(z) = \frac{\omega_0 n_2(z)}{\mathrm{c} A_\mathrm{eff}(z)}\,, \tag{3.71}$$

with the Kerr coefficient $n_2(z)$ (typically $3 \times 10^{-20}\,\mathrm{m}^2/\mathrm{W}$ for silica fibers), and the effective fiber area $A_\mathrm{eff}(z)$. A typical value for the (path-average) nonlinear coefficient is $\bar{\gamma} \approx$

[4] In particular, the group delay $L/\nu_\mathrm{g}(\omega_0)$ of the probe signal and any constant common phase rotation $\beta_0(L)$ has already been canceled from the propagation equation in the baseband model.

$1.1\,\mathrm{W}^{-1}\mathrm{km}^{-1}$ for SSMF. Note, that due to the normalization of the signal, the nonlinearity coefficient has strictly speaking units of m^{-1} to be rigorous within the framework of our signal definition, see discussion in Section 2.1.

The optical Kerr effect leads to a number of nonlinear effects, which were first classified in the framework of *dispersion-managed* transmission systems. In dispersion managed transmission, the signal's pulse shape is either maintained or (periodically) recovered during transmission by means of *dispersion-shifted* or *dispersion-compensating* fibers. This leads to highly deterministic nonlinear distortions giving rise to the (traditional, phenomenal) taxonomy of nonlinear effects, namely: self-phase modulation (SPM) [Agr06, Ch. 4], cross-phase modulation (XPM) [Agr06, Ch. 7], cross-polarization modulation (XPolM), and classical four-wave mixing (FWM) [Agr06, Ch. 10].

In the present work, we focus on (dispersion-) *uncompensated* (i.e., *unmanaged*) transmission where the original transmit pulse is constantly changing during propagation due to chromatic dispersion. In this situation, the nonlinear effects, see above, do not manifest themselves as clear phase distortions (i.e., *phase modulation*) after reception. Instead, depending on the considered system scenario, those nonlinear effects can also have features similar to additive noise with Gaussian-like distortions. We hence adopt the taxonomy introduced by Poggiolini *et al.* [Pog12, Sec. VI]. Herein, three different cases of nonlinear (signal-to-signal) distortions are distinguished by its origin rather than its manifestation:

- self-channel interference (SCI), i.e., interference caused by the (probe) wavelength channel itself,
- cross-channel interference (XCI), i.e., interference between the probe channel and a single interfering wavelength channel (aka. *degenerate* cross-channel interference), and
- multi-channel interference (MCI), i.e., interference between the probe channel and two or three other interfering wavelength channels (aka. *non-degenerate* cross-channel interference).

To provide a brief mathematical description of the three categories, we consider only the nonlinear term in the Manakov equation without dispersion and attenuation. This approximation is, e.g., valid for very short fiber segments of length $z \ll L_\mathrm{D}$ and $z \ll L_\mathrm{eff,a}$ where the dispersion- and loss-related terms can be dropped. The Manakov equation reduces to

$$\frac{\partial}{\partial z}\boldsymbol{u}(z,t) = -\mathrm{j}\gamma(z)\frac{8}{9}\left\|\boldsymbol{u}(z,t)\right\|^2\boldsymbol{u}(z,t)\,. \tag{3.72}$$

If we assume that the nonlinear coefficient does not depend on z, e.g., we only consider the path-average $\bar{\gamma}$, then the simple solution is obtained by

$$\boldsymbol{u}(z,t) = \boldsymbol{u}(0,t)\,\exp\left(-\mathrm{j}\bar{\gamma}\frac{8}{9}\left\|\boldsymbol{u}(0,t)\right\|^2 z\right)\,, \tag{3.73}$$

i.e., for the non-dispersive and attenuation-free setting, the signal at $z = L$ is given by the input signal $\boldsymbol{u}(0,t)$ rotated by a fixed value in phase depending on the signal intensity $\left\|\boldsymbol{u}(0,t)\right\|^2$ at the input and the length of the link L. In this context, we define the (ν-dependent) *nonlinear length* [Agr06, Eq. (4.1.2)] as

$$L_{\mathrm{NL},\nu} \stackrel{\text{def}}{=} \frac{1}{\bar{\gamma} P_\nu} \,. \tag{3.74}$$

where $P_\nu = \frac{\sigma_{b,\nu}^2}{T} E_{\mathrm{T},\nu}$, see (2.77), is the optical launch power of the ν^{th} wavelength channel. The nonlinear length $L_{\mathrm{NL},\nu}$ is equivalent to the propagation distance at which the (time-average) phase rotation amounts to $8/9$ rad assuming only a single wavelength channel with launch power P_ν.

We continue to expand the nonlinear *source* term on the right-hand side of (3.72) into expressions that result in self-, cross-, and multi-channel interference. It is defined as

$$\boldsymbol{w}(z,t) \stackrel{\text{def}}{=} \|\boldsymbol{u}(z,t)\|^2 \, \boldsymbol{u}(z,t) = \boldsymbol{u}^{\mathsf{H}}(z,t)\boldsymbol{u}(z,t) \, \boldsymbol{u}(z,t) \,. \tag{3.75}$$

The source term $\boldsymbol{w}(z,t)$ can be expanded at $z=0$ using (3.14)–(3.15) to arrive at

$$\begin{aligned}
\boldsymbol{w}(0,t) &= \sum_{\nu_2=1}^{N_{\mathrm{ch}}} \boldsymbol{u}_{\nu_2}^{\mathsf{H}}(0,t) \sum_{\nu_1=1}^{N_{\mathrm{ch}}} \boldsymbol{u}_{\nu_1}(0,t) \sum_{\nu_3=1}^{N_{\mathrm{ch}}} \boldsymbol{u}_{\nu_3}(0,t) && (3.76)\\
&= \sum_{\nu_2=1}^{N_{\mathrm{ch}}} \boldsymbol{s}_{\nu_2}^{\mathsf{H}}(t) \exp(-\mathrm{j}\Delta\omega_{\nu_2} t) \sum_{\nu_1=1}^{N_{\mathrm{ch}}} \boldsymbol{s}_{\nu_1}(t) \exp(\mathrm{j}\Delta\omega_{\nu_1} t) \sum_{\nu_3=1}^{N_{\mathrm{ch}}} \boldsymbol{s}_{\nu_3}(t) \exp(\mathrm{j}\Delta\omega_{\nu_3} t) \,. && (3.77)
\end{aligned}$$

The expansion at any other z is not exact due to z-dependency of the nonlinear term $\boldsymbol{w}(z,t)$ according to (3.72), but generally the following considerations and the taxonomy derived thereof still hold.

We are now interested in the nonlinear distortion generated in the support of the probe channel. To that end, the following frequency constraint on the relative offsets in (3.77) must be met

$$\Delta\omega_{\rho} \stackrel{!}{=} \Delta\omega_{\nu_1} - \Delta\omega_{\nu_2} + \Delta\omega_{\nu_3} \,, \tag{3.78}$$

where, per assumption, the probe wavelength channel is located at $\Delta\omega_\rho = 0$ in the ECB model. The relevant source term (which has support in the probe's spectral domain) can now be sorted into groups by degeneracy,

$$\begin{aligned}
\boldsymbol{w}(0,t)\Big|_{\mathrm{supp}(\boldsymbol{W})\subset\mathcal{B}_\rho} &= \underbrace{\|\boldsymbol{u}_\rho(0,t)\|^2 \, \boldsymbol{u}_\rho(0,t)}_{\substack{\mathrm{SCI} \\ \nu_1=\nu_2=\nu_3=\rho}} + \underbrace{\sum_{\nu\neq\rho} \|\boldsymbol{u}_\nu(0,t)\|^2 \, \boldsymbol{u}_\rho(0,t)}_{\substack{\mathrm{XCI} \\ \nu_1=\nu_2=\nu,\ \nu_3=\rho}} && (3.79)\\
&\quad + \underbrace{\sum_{\nu\neq\rho} \boldsymbol{u}_\nu(0,t)\boldsymbol{u}_\nu^{\mathsf{H}}(0,t)\boldsymbol{u}_\rho(0,t)}_{\substack{\mathrm{XCI} \\ \nu_1=\rho,\ \nu_2=\nu_3=\nu}} + \underbrace{\sum\sum_{\substack{\nu_2,\nu_3\neq\rho \\ \nu_1=\nu_2+\nu_3-\rho}}\sum \boldsymbol{u}_{\nu_2}^{\mathsf{H}}(0,t)\boldsymbol{u}_{\nu_1}(0,t)\boldsymbol{u}_{\nu_3}(0,t)}_{\mathrm{MCI}} \,,
\end{aligned}$$

where we used $\boldsymbol{u}_\nu^{\mathsf{H}}(z,t)\boldsymbol{u}_\rho(z,t)\boldsymbol{u}_\nu(z,t) = \boldsymbol{u}_\nu(z,t)\boldsymbol{u}_\nu^{\mathsf{H}}(z,t)\boldsymbol{u}_\rho(z,t)$. Here, we also assumed (without loss of generality) that the wavelength channels are arranged on a regular frequency grid, i.e., spaced at integer multiples of a common channel separation $\Delta\omega$. Any other nonlinear contributions which do not manifest themselves within the bandwidth of the probe channel $\mathcal{B}_\rho$ are neglected in the expansion above.

The first term on the right-hand side of (3.79) is the source for SCI (aka. SPM using the traditional taxonomy), and has a doubly-degenerate form ($\nu_1 = \nu_2 = \nu_3 = \rho$) since only the probe channel itself contributes to the nonlinear interaction.

The second and third term on the right-hand side of (3.79) are both contributions due to XCI and can be combined using (A.40) to rearrange

$$\boldsymbol{u}_\nu(z,t)\boldsymbol{u}_\nu^{\mathsf{H}}(z,t) = \frac{1}{2}\left(\|\boldsymbol{u}_\nu(z,t)\|^2\,\mathbf{I} + \vec{\boldsymbol{u}}_\nu(z,t)\cdot\vec{\boldsymbol{\sigma}}\right)\,, \tag{3.80}$$

where we make use of a common notation in optical communication using the dot product of a *Stokes vector* and *Pauli vector* to denote Hermitian matrices in $\mathbb{C}^{2\times 2}$. The unfamiliar reader may consult Appendix A.1 for an introduction on the Jones and Stokes formalism. We can use (3.80) to obtain

$$\begin{aligned}\boldsymbol{w}(0,t)\Big|_{\mathrm{supp}(\boldsymbol{W})\subset\mathcal{B}_\rho} &= \underbrace{\|\boldsymbol{u}_\rho(0,t)\|^2\,\boldsymbol{u}_\rho(0,t)}_{\substack{\text{SCI}\\ \nu_1=\nu_2=\nu_3=\rho}} + \underbrace{\sum_{\nu\neq\rho}\left(\frac{3}{2}\,\|\boldsymbol{u}_\nu(0,t)\|^2\,\mathbf{I} + \frac{1}{2}\vec{\boldsymbol{u}}_\nu(0,t)\cdot\vec{\boldsymbol{\sigma}}\right)\boldsymbol{u}_\rho(0,t)}_{\substack{\text{XCI}\\ \nu_1=\nu_2=\nu,\ \nu_3=\rho}}\\ &\quad + \underbrace{\sum\sum\sum_{\substack{\nu_2,\nu_3\neq\rho\\ \nu_1=\nu_2+\nu_3-\rho}} \boldsymbol{u}_{\nu_2}^{\mathsf{H}}(0,t)\boldsymbol{u}_{\nu_1}(0,t)\boldsymbol{u}_{\nu_3}(0,t)}_{\text{MCI}}\,,\end{aligned} \tag{3.81}$$

where the XCI contribution is single-degenerate ($\nu_1 = \nu_2, \nu_3 = \rho$). The first term in the parenthesis is, in traditional taxonomy, associated with XPM common in both x- and y-polarization. If the co-propagating wavelength channels are not co-polarized but instead randomly polarized w.r.t. the probe channel, the XPM term is in fact the average over all relative polarization states of the probe and interfering channels. Similar to SPM, XPM induces a (both temporally and spatially) local phase modulation of the optical signal $\boldsymbol{u}(z,t)$, see the propagation equation (3.72). These (local) phase distortions will be transformed into a mixture of phase and amplitude distortions by chromatic dispersion.

The second term in the parenthesis is associated with XPolM (i.e., nonlinearly induced birefringence) and relative, polarization-dependent XPM.

The last term on the right-hand side relates to MCI. We will see that typically MCI is negligible for most relevant system scenarios using current technology. Whether MCI can be neglected depends (to first order) on the channel separation, as the so-called *FWM efficiency* decreases rapidly when the contributing wavelength channels are spaced far abart, see next chapter.

3.3.6 The Sampled Manakov Equation

The split-step Fourier method (SSFM) [SHZM03] is a numerical (approximate) solution to the Manakov equation and the de-facto standard method to perform numerical simulations for

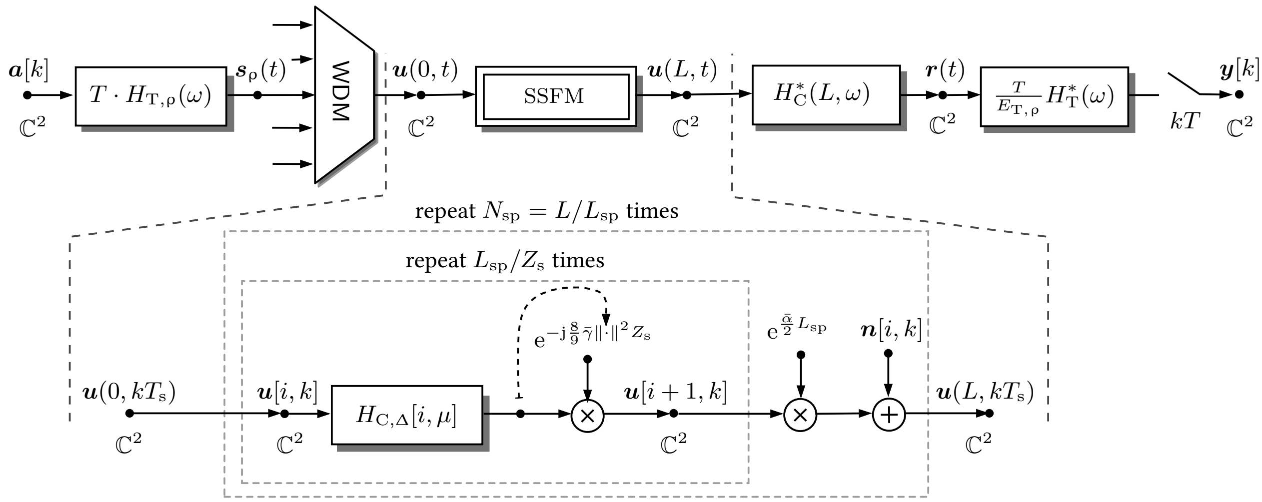


Figure 3.13: Split-step Fourier method for solving the end-to-end optical channel. In the block diagram, a transmission link is considered with lumped end-of-span amplification compensating for the exact span loss via the amplifier (power) gain $\mathrm{e}^{G} = \mathrm{e}^{\bar{\alpha} L_{\mathrm{sp}}}$. The Manakov equation is solved numerically using an oversampled representation of $\boldsymbol{u}[i,k]$ in small steps of Z_{s} where linear propagation via $H_{\mathrm{C},\Delta}[i,\mu]$ and the nonlinear phase rotation are performed sequentially.

optical transmission systems. The algorithmic implementation is depicted as a block diagram in Figure 3.13. Here, the transmission link is divided into fiber spans (due to the amplification scheme selected in this example), and each span itself is divided into small fiber sections of length Z_{s}, known as the *step size* of the split-step algorithm. This is essentially equivalent to *sampling* the optical field envelope in the *spatial* domain at integer multiples of Z_{s} (aka. *constant step-size method* [SHZM03]). Within each section the linear and the nonlinear effects are treated separately (and sequentially). Accordingly, the SSFM is also termed *sampled NLSE*, or alternatively, *sampled Manakov equation* in case of two polarization modes.

The numerical calculations are carried out on an oversampled representation (in both time and space) of the ECB signal

$$\boldsymbol{u}[i,k] \overset{\text{def}}{=} \boldsymbol{u}(iZ_{\mathrm{s}}, kT_{\mathrm{s}}) \,, \tag{3.82}$$

with $i \in \mathbb{N}$ and $k \in \mathbb{Z}$. The (temporal) sampling frequency $\omega_{\mathrm{s}} \overset{\text{def}}{=} 2\pi f_{\mathrm{s}} = 2\pi/T_{\mathrm{s}}$ is typically chosen much higher than the bandwidth B_{WDM}, see (3.20), of the optical field envelope. Since the nonlinear Kerr effect is (to first approximation) a *third-order* nonlinear system, the sampling frequency is commonly chosen *three* times the Nyquist frequency of the optical field envelope [PBP97], i.e., $\omega_{\mathrm{s}} = 3/2 B_{\mathrm{WDM}}$, assuming that the spectral support of the field envelope is centered around zero. Then, (first-order) nonlinear mixing products fall outside the original signal band $\mathcal{B}_{\mathrm{WDM}}$ without cyclic wrapping in frequency domain.

To efficiently carry out the numerical calculations of the *linear* propagation step, see (3.63)–(3.64), the frequency-domain representation of (3.82) is also discretized. This allows to solve the linear step in (discrete-) frequency domain by performing a point-wise multiplication of the DFT transform $\boldsymbol{u}[i,k] \circ\!\!-\!\!\bullet\, \boldsymbol{U}[i,\mu]$ with the sampled (differential) linear transfer function. We define the (differential) linear transfer function between two consecutive discrete, spatial points (i.e., steps) as

$$H_{\mathrm{C},\Delta}[i,\mu] \overset{\mathrm{def}}{=} H_{\mathrm{C}}((i+1)Z_{\mathrm{s}},\mu/T_0)/H_{\mathrm{C}}(iZ_{\mathrm{s}},\mu/T_0)\,, \tag{3.83}$$

where T_0 is the period of the signal (aka. *time window*, or fundamental period of the sequence in time domain) with $M_{\mathrm{s}} = T_0/T_{\mathrm{s}}$ samples per period. In the subsequent step, the (differential) nonlinear phase rotation, see (3.73), is applied by a (point-wise) multiplication in (discrete-) time domain with the phasor

$$\exp\left(-\mathrm{j}\phi_{\mathrm{NL},Z_{\mathrm{s}}}[i,k]\right) \overset{\mathrm{def}}{=} \exp\left(-\mathrm{j}\bar{\gamma}\frac{8}{9}\,\|\boldsymbol{u}[i,k]\|^2\,Z_{\mathrm{s}}\right)\,. \tag{3.84}$$

The discretization in both time and frequency results in a T_0-periodic sequence in time, and a $1/T_{\mathrm{s}}$-periodic sequence in frequency.

The numerical simulations performed in Chapter 5 are implemented via the *symmetric* SSFM [SHZM03] with *adaptive* step size $Z_{\mathrm{s}}(z)$, i.e., instead of uniform spatial sampling, the step size depends on z, particularly on the power profile $\mathcal{P}(z)$ such that the differential nonlinear phase-rotation per step (3.84) does not exceed a certain maximum value. By adjusting the step-size, we require that the nonlinear phase rotation per step is, e.g., no larger than

$$\phi_{\mathrm{NL},Z_{\mathrm{s}}}[i,k] = \bar{\gamma}\frac{8}{9}\,\|\boldsymbol{u}[i,k]\|^2\,Z_{\mathrm{s}} \overset{!}{<} \phi^{\max}_{\mathrm{NL},Z_{\mathrm{s}}}\,. \tag{3.85}$$

The numerical system simulations performed in Chapter 5 are carried out with the setting $\phi^{\max}_{\mathrm{NL},Z_{\mathrm{s}}} = 3.5\times10^{-4}\ \mathrm{rad}$ which will result in accurate solutions for the considered scenarios.

4. The Continuous-Time Perturbation Approach

The principle philosophy of fiber-optical channel models based on the perturbation method is to assume that nonlinear distortions are *weak* compared to its source, i.e., the propagating signal. With *weak* we mean that the power of the nonlinear distortion is considerably smaller than the power of the signal (sometimes called *pseudo-linear* transmission regime). Starting from this premise the well-known RP ansatz for the continuous-time, *optical* end-to-end channel is written as [VB02, Eq. (14)] [Wei06, Eq. (3)] [JK13, Eq. (2)]

$$\boldsymbol{u}(z,t) = \boldsymbol{u}_{\mathrm{LIN}}(z,t) + \Delta\boldsymbol{u}(z,t) \tag{4.1}$$

$$\begin{array}{c}\circ\\ | \\ \bullet\end{array}$$

$$\boldsymbol{U}(z,\omega) = \boldsymbol{U}_{\mathrm{LIN}}(z,\omega) + \Delta\boldsymbol{U}(z,\omega)\,, \tag{4.2}$$

where $\boldsymbol{u}_{\mathrm{LIN}}(z,t) \circ\!\!-\!\!\bullet \boldsymbol{U}_{\mathrm{LIN}}(z,\omega)$ is the signal propagating according to only linear effects, i.e., according to (3.63), (3.64). In this context, the nonlinear distortion

$$\Delta\boldsymbol{u}(z,t) \circ\!\!-\!\!\bullet \Delta\boldsymbol{U}(z,\omega) \in \mathbb{C}^2\,,$$

aka. the *perturbation* is generated locally according to nonlinear signal-signal interaction and is then propagated linearly and independently of the signal $\boldsymbol{u}_{\mathrm{LIN}}(z,t)$ to the end of the optical channel at $z = L$. We assume that the optical perturbation at $z = 0$ is zero, i.e., $\Delta\boldsymbol{u}(0,t) = 0$. The received signal $\boldsymbol{u}(L,t)$ is then given as the sum of the solution for the linearly propagating signal and the perturbation representing the accumulated nonlinear effects, i.e.,

$$\boldsymbol{u}(L,t) = \boldsymbol{u}(0,t) * h_{\mathrm{C}}(L,t) + \Delta\boldsymbol{u}(L,t) \tag{4.3}$$

$$\begin{array}{c}\circ\\ | \\ \bullet\end{array}$$

$$\boldsymbol{U}(L,\omega) = \boldsymbol{U}(0,\omega)\, H_{\mathrm{C}}(L,\omega) + \Delta\boldsymbol{U}(L,\omega)\,. \tag{4.4}$$

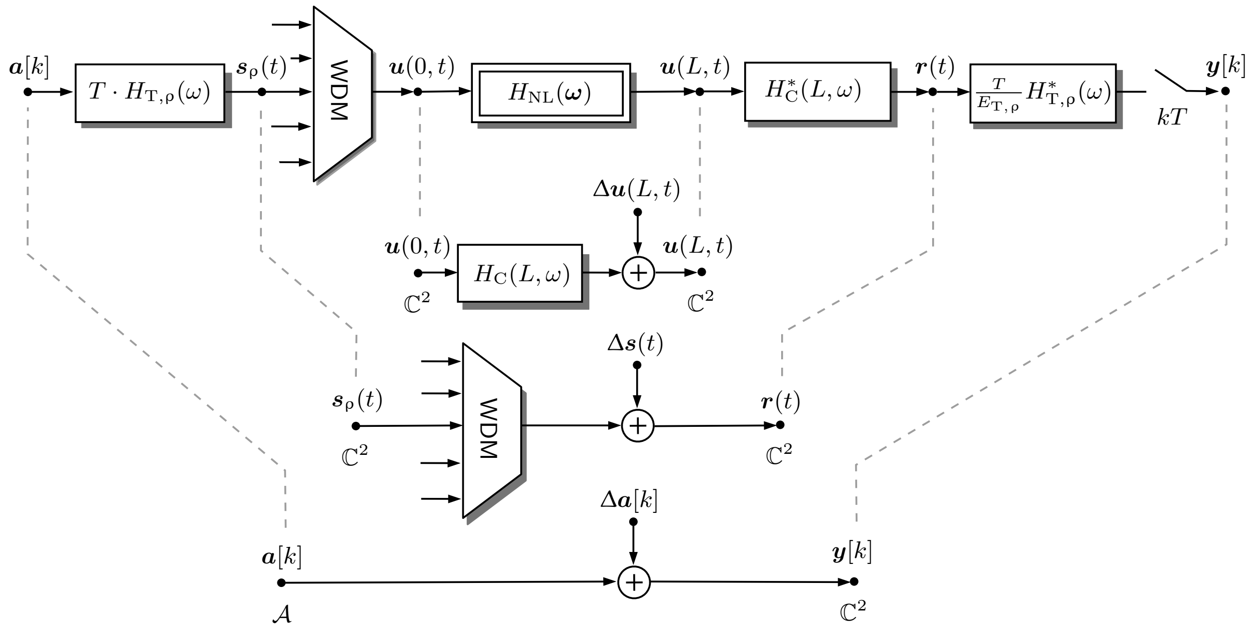


Figure 4.1: Block diagram and terminology used for the corresponding end-to-end perturbation models. The *optical* end-to-end perturbation is denoted by $\Delta \boldsymbol{u}(L,t)$, the (analog) *baseband* end-to-end perturbation is denoted by $\Delta \boldsymbol{s}(t)$, and the *discrete-time* end-to-end perturbation is denoted by $\Delta \boldsymbol{a}[k]$.

Likewise, end-to-end relations can be also formulated for the *baseband* end-to-end channel, and the *discrete-time* end-to-end channel, see, e.g., Figure 1.1. The associated block diagram and terminology used throughout this thesis is highlighted in Figure 4.1.

The ultimate objective of the current and the following chapter is to develop the input/output relation of the equivalent discrete-time end-to-end channel in the form of

$$\boldsymbol{y}[k] = \boldsymbol{a}[k] + \Delta \boldsymbol{a}[k] \tag{4.5}$$

$$\circ\!\!-\!\!\bullet$$

$$\boldsymbol{Y}(\mathrm{e}^{\mathrm{j}\omega T}) = \boldsymbol{A}(\mathrm{e}^{\mathrm{j}\omega T}) + \Delta \boldsymbol{A}(\mathrm{e}^{\mathrm{j}\omega T}), \tag{4.6}$$

where the total perturbation is concentrated in a single *discrete-time* term $\Delta \boldsymbol{a}[k] \circ\!\!-\!\!\bullet \Delta \boldsymbol{A}(\mathrm{e}^{\mathrm{j}\omega T})$.

To that end, we start with the theory of first-order perturbation developed for the continuous-time, optical end-to-end relation (i.e., the inner equivalent block diagram in Figure 4.1) and successively embed the required functional blocks from the analog and discrete-time domain. In this chapter, we focus on the continuous-time end-to-end description, i.e., from the baseband transmit signal of the probe (with subscript ρ) and its interferers $\boldsymbol{s}_\nu(t)$ to the received signal $\boldsymbol{r}(t)$. The *nonlinear impulse response* and the *nonlinear transfer function* $h_{\mathrm{NL}}(\boldsymbol{\tau}) \circ\!\!-\!\!\bullet H_{\mathrm{NL}}(\boldsymbol{\omega})$ are introduced and their relation to the third-order Volterra kernel are discussed. The *map strength* $\mathcal{S}_{T,\rho} \propto L_{\mathrm{eff}}/L_{\mathrm{D}}$ (or equivalently the ν-dependent $\mathcal{S}_{T,\nu} \propto L_{\mathrm{eff}}/L_{\mathrm{wo},\nu}$) will be shown to be a measure of the temporal extent, i.e., the memory of the nonlinear interaction (relative to the probe's symbol rate).

4.1 The First-Order Regular Solution in Frequency Domain

Inserting the ansatz from (4.2) into the Manakov equation (3.21) we obtain

$$\begin{aligned}\frac{\partial}{\partial z}(\boldsymbol{u}_{\mathrm{LIN}}(z,t)+\Delta\boldsymbol{u}(z,t)) = {} & \mathrm{j}\frac{\beta_2(z)}{2}\frac{\partial^2}{\partial t^2}(\boldsymbol{u}_{\mathrm{LIN}}(z,t)+\Delta\boldsymbol{u}(z,t)) \\ & + \frac{g(z)-\alpha(z)}{2}(\boldsymbol{u}_{\mathrm{LIN}}(z,t)+\Delta\boldsymbol{u}(z,t)) \\ & - \mathrm{j}\gamma(z)\frac{8}{9}\left(\|\boldsymbol{u}_{\mathrm{LIN}}(z,t)+\Delta\boldsymbol{u}(z,t)\|^2(\boldsymbol{u}_{\mathrm{LIN}}(z,t)+\Delta\boldsymbol{u}(z,t))\right),\end{aligned} \tag{4.7}$$

where the expansion of the nonlinear source term results in

$$\begin{aligned}\left(\|\boldsymbol{u}_{\mathrm{LIN}}(z,t)+\Delta\boldsymbol{u}(z,t)\|^2(\boldsymbol{u}_{\mathrm{LIN}}(z,t)+\Delta\boldsymbol{u}(z,t))\right) = {} & \|\boldsymbol{u}_{\mathrm{LIN}}(z,t)\|^2\boldsymbol{u}_{\mathrm{LIN}}(z,t) \\ & + 2\mathrm{Re}\{\boldsymbol{u}_{\mathrm{LIN}}^{\mathsf{H}}(z,t)\Delta\boldsymbol{u}(z,t)\}\boldsymbol{u}_{\mathrm{LIN}}(z,t) \\ & + \|\Delta\boldsymbol{u}(z,t)\|^2\boldsymbol{u}_{\mathrm{LIN}}(z,t) \\ & + \|\boldsymbol{u}_{\mathrm{LIN}}(z,t)\|^2\Delta\boldsymbol{u}(z,t) \\ & + 2\mathrm{Re}\{\boldsymbol{u}_{\mathrm{LIN}}^{\mathsf{H}}(z,t)\Delta\boldsymbol{u}(z,t)\}\Delta\boldsymbol{u}(z,t) \\ & + \|\Delta\boldsymbol{u}(z,t)\|^2\Delta\boldsymbol{u}(z,t)\,.\end{aligned} \tag{4.8}$$

The main approximation of the perturbation approach is to neglect all cross-products between the signal $\boldsymbol{u}_{\mathrm{LIN}}(z,t)$ and the perturbation $\Delta\boldsymbol{u}(z,t)$ or the perturbation with itself, such that only the leading nonlinear term $\|\boldsymbol{u}_{\mathrm{LIN}}(z,t)\|^2\boldsymbol{u}_{\mathrm{LIN}}(z,t)$ is considered. This allows us to split (4.7) into the following set of differential equations

$$\frac{\partial}{\partial z}\boldsymbol{u}_{\mathrm{LIN}}(z,t) - \mathrm{j}\frac{\beta_2(z)}{2}\frac{\partial^2}{\partial t^2}\boldsymbol{u}_{\mathrm{LIN}}(z,t) - \frac{g(z)-\alpha(z)}{2}\boldsymbol{u}_{\mathrm{LIN}}(z,t) = 0\,, \tag{4.9}$$

and

$$\frac{\partial}{\partial z}\Delta\boldsymbol{u}(z,t) - \mathrm{j}\frac{\beta_2(z)}{2}\frac{\partial^2}{\partial t^2}\Delta\boldsymbol{u}(z,t) - \frac{g(z)-\alpha(z)}{2}\Delta\boldsymbol{u}(z,t) = -\mathrm{j}\gamma(z)\frac{8}{9}\|\boldsymbol{u}_{\mathrm{LIN}}(z,t)\|^2\boldsymbol{u}_{\mathrm{LIN}}(z,t)\,. \tag{4.10}$$

The first equation is a linear, homogeneous second-order PDE and describes the evolution of the *linear* optical field envelope. The second equation is of the same form, but has an additional *inhomogeneous* term on the right-hand side of the equation, i.e., the linearly propagating field acts as the *source* of the local perturbation in the optical domain.

The solution of the first PDE in (4.9) has been derived in Section 3.3.3 using the linear channel impulse response and transfer function $h_{\mathrm{C}}(z,t) \circ\!\!-\!\!\bullet\, H_{\mathrm{C}}(z,\omega)$ representing the joint effects of chromatic dispersion and the gain/loss variation along the link. We will use the solution $\boldsymbol{u}_{\mathrm{LIN}}(z,t)$ on the right-hand side of (4.10).

The solution to the second PDE in (4.10) is obtained by first performing a signal transformation on the optical perturbation such that the gain/loss profile is instead represented by a re-scaled nonlinear source term. To that end, we define the (path-) *normalized* optical perturbation $\Delta\tilde{\boldsymbol{u}}(z,t)$ to have constant average power over z with

$$\Delta\tilde{\boldsymbol{u}}(z,t) \stackrel{\mathrm{def}}{=} \Delta\boldsymbol{u}(z,t)/\sqrt{\mathcal{P}(z)}\,, \tag{4.11}$$

where $\mathcal{P}(z) = \exp(\mathcal{G}(z))$ is the normalized power profile defined in (3.24)–(3.25). The same normalization strategy is, e.g., also pursued to obtain the *power-normalized* nonlinear Schrödinger equation (or equivalently the normalized Manakov equation), see [Agr06].

Starting from (4.10), the *chain rule of calculus* applied to $\Delta\boldsymbol{u}(z,t) = \sqrt{\mathcal{P}(z)}\Delta\tilde{\boldsymbol{u}}(z,t)$, using (4.11), yields

$$\frac{\partial}{\partial z}\left(\sqrt{\mathcal{P}(z)}\,\Delta\tilde{\boldsymbol{u}}(z,t)\right) = \frac{g(z)-\alpha(z)}{2}\sqrt{\mathcal{P}(z)}\,\Delta\tilde{\boldsymbol{u}}(z,t) + \sqrt{\mathcal{P}(z)}\,\frac{\partial}{\partial z}\Delta\tilde{\boldsymbol{u}}(z,t)\,, \tag{4.12}$$

where we use the derivative of the power profile from (3.24). The expression can be substituted in (4.10) and both sides are divided by $\sqrt{\mathcal{P}(z)} = \mathrm{e}^{\frac{\mathcal{G}(z)}{2}}$ to arrive at[1]

$$\frac{\partial}{\partial z}\Delta\tilde{\boldsymbol{u}}(z,t) - \mathrm{j}\frac{\beta_2(z)}{2}\frac{\partial^2}{\partial t^2}\Delta\tilde{\boldsymbol{u}}(z,t) = -\mathrm{j}\gamma(z)\frac{8}{9}\,\frac{1}{\sqrt{\mathcal{P}(z)}}\,\left\|\boldsymbol{u}_{\mathrm{LIN}}(z,t)\right\|^2\boldsymbol{u}_{\mathrm{LIN}}(z,t)\,, \tag{4.13}$$

where the gain/loss profile $\mathcal{P}(z)$ now effectively scales the strength of the nonlinear source term on the right-hand side. For ease of notation, we assume that the *local* z-variation in $\gamma(z)$ can be equivalently expressed in a variation of either a local gain $g(z)$ or the local fiber attenuation coefficient $\alpha(z)$. Hence, we only consider the *path-average* nonlinear coefficient $\bar{\gamma}$ from (3.22) and capture the z-dependency of $\gamma(z)$ via $\mathcal{P}(z)$ in the following.

Fourier transforming equation (4.13), we obtain a first-order, linear, inhomogeneous ordinary differential equation

$$\frac{\partial}{\partial z}\Delta\tilde{\boldsymbol{U}}(z,\omega) + \mathrm{j}\omega^2\frac{\beta_2(z)}{2}\Delta\tilde{\boldsymbol{U}}(z,\omega) = -\mathrm{j}\bar{\gamma}\frac{8}{9}\,\mathrm{e}^{-\frac{\mathcal{G}(z)}{2}}\,\mathcal{F}\{\,\left\|\boldsymbol{u}_{\mathrm{LIN}}(z,t)\right\|^2\boldsymbol{u}_{\mathrm{LIN}}(z,t)\,\}\,, \tag{4.14}$$

where we used the symbolic correspondence $\frac{\partial^2}{\partial t^2} \circ\!\!-\!\!\bullet\ -\omega^2$.

The method of the *integrating factor* [RE10, P. 449] can be applied to solve the equation (4.14) where the integrating factor is equal to

$$\exp\left(\mathrm{j}\omega^2\frac{1}{2}\int_0^z\beta_2(\zeta)\mathrm{d}\zeta\right) = \exp\left(\mathrm{j}\omega^2\frac{\mathcal{B}(z)}{2}\right)\,. \tag{4.15}$$

We obtain the solution [ME12, Eq. (31)], [Joh12, Eq. (17)]

$$\Delta\tilde{\boldsymbol{U}}(z,\omega) = -\mathrm{j}\bar{\gamma}\frac{8}{9}\,\mathrm{e}^{-\mathrm{j}\omega^2\frac{\mathcal{B}(z)}{2}}\int_0^z\mathrm{e}^{+\mathrm{j}\omega^2\frac{\mathcal{B}(\zeta)}{2}}\,\mathrm{e}^{-\frac{\mathcal{G}(\zeta)}{2}}\,\mathcal{F}\{\,\left\|\boldsymbol{u}_{\mathrm{LIN}}(\zeta,t)\right\|^2\boldsymbol{u}_{\mathrm{LIN}}(\zeta,t)\,\}\,\mathrm{d}\zeta \tag{4.16}$$

$$= -\mathrm{j}\bar{\gamma}\frac{8}{9}\,H_{\mathrm{CD}}(z,\omega)\int_0^z H_{\mathrm{C}}^{-1}(\zeta,\omega)\,\mathcal{F}\{\,\left\|\boldsymbol{u}_{\mathrm{LIN}}(\zeta,t)\right\|^2\boldsymbol{u}_{\mathrm{LIN}}(\zeta,t)\,\}\,\mathrm{d}\zeta\,, \tag{4.17}$$

where we used the chromatic dispersion transfer function $H_{\mathrm{CD}}(z,\omega)$ from (3.48), and the inverse of the linear channel transfer function $H_{\mathrm{C}}(z,\omega)$ from (3.63).

A pictorial explanation of the equation is given in Figure 4.2 (cf. also the *parallel fiber model* in [VSB02, Sec. IV]). The optical signal at the input of the first fiber span $\boldsymbol{u}_{\mathrm{LIN}}(0,t)$ propagates linearly to the (arbitrary, fixed) location $\mathrm{d}\zeta$ where the *local* perturbation is generated according

[1] The procedure via the normalized optical perturbation is equivalent to the method of *integrating factor* [RE10, P. 449], which exploits the product and chain rule similarly.

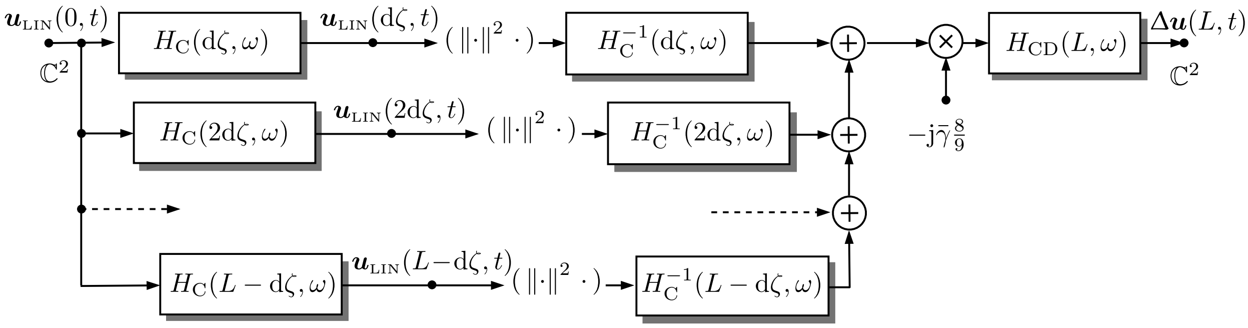


Figure 4.2: The *parallel* fiber-optical channel model [VSB02] derived from first-order regular perturbation. The local perturbations from the parallel branches are summed over ζ each associated with a unique local position along the transmission link. In the limit $\text{d}\zeta \to 0$, the summation becomes an integral as in (4.17).

to the double product of the local, linearly propagating signal, i.e., $\|\boldsymbol{u}_{\text{LIN}}(\text{d}\zeta,t)\|^2\ \boldsymbol{u}_{\text{LIN}}(\text{d}\zeta,t)$. The generated perturbation is then propagated back to the fiber input at $z=0$ via the inverse linear transfer function $H_{\text{C}}^{-1}(\text{d}\zeta,\omega)$ and then summed with all remaining perturbations from locations other than $\text{d}\zeta$. The sum over all local perturbations is then jointly propagated to the end of the transmission link at $z=L$ by the dispersion transfer function $H_{\text{CD}}(L,\omega)$. In the limit $\text{d}\zeta \to 0$, the branching and summation over the z-dimension in Figure 4.2 becomes continuous. By definition, we have $\mathcal{P}(L)=1$, and it follows that $\Delta\boldsymbol{u}(L,t)=\Delta\tilde{\boldsymbol{u}}(L,t)$.

The authors of [VSB02] call this channel model derived from first-order RP a *parallel fiber model*, since, in contrast to the *sequential* SSFM, once a local perturbation is generated, it does no longer interact nonlinearly during transmission. When second- (or higher-) order terms are included into the RP approach, then local perturbations may act as nonlinear source, see right-hand side of (4.8), for higher-order perturbations. In the limit, i.e., when the order of the perturbation approach goes to infinity, the RP solution approaches the SSFM solution (for which the step-size $\Delta z \to 0$), see also discussion in [VSB02, Sec. IV]. The term *regular* in RP implies that the asymptotic solution approaches the true solution of the nonlinear system.

Note, that the accumulated chromatic dispersion is generally compensated by the receiver (front-end) using bulk electronic dispersion compensation such that the system *net* dispersion after the receiver front-end is zero, see Section 3.3.4. Then, the effective perturbation after reception, i.e., after the channel matched filter, is the perturbation propagated back to the fiber input at $z=0$. We can hence cancel the dispersion operator $H_{\text{CD}}(L,\omega)$ at the channel output in Figure 4.2 and only consider the system relevant end-to-end model (i.e., including receiver-side chromatic dispersion compensation).

We now continue to modify the right-hand side of (4.17) by carrying out the Fourier transform $\mathcal{F}\{\,\cdot\,\}$ to obtain

$$\begin{gathered}\|\boldsymbol{u}_{\text{LIN}}(z,t)\|^2\,\boldsymbol{u}_{\text{LIN}}(z,t) = \left(\,|u_{\text{LIN,x}}(z,t)|^2 + |u_{\text{LIN,y}}(z,t)|^2\,\right)\boldsymbol{u}_{\text{LIN}}(z,t)\\ \circ\!\!-\!\!\bullet\\ \Big(U_{\text{LIN,x}}(z,\omega) * U^*_{\text{LIN,x}}(z,-\omega) + U_{\text{LIN,y}}(z,\omega) * U^*_{\text{LIN,y}}(z,-\omega)\Big) * \boldsymbol{U}_{\text{LIN}}(z,\omega)\,,\end{gathered} \tag{4.18}$$

using $\boldsymbol{u}_{\text{LIN}}(z,t) = [u_{\text{LIN,x}}(z,t), u_{\text{LIN,y}}(z,t)]^{\mathsf{T}}$. We may write the relation using the short-hand notation $\boldsymbol{u}_{\mathrm{LIN}}^{\mathsf{H}}(z,t)\boldsymbol{u}_{\mathrm{LIN}}(z,t)\boldsymbol{u}_{\mathrm{LIN}}(z,t) \circ\!\!-\!\!\bullet\ \boldsymbol{U}_{\mathrm{LIN}}^{\mathsf{H}}(z,-\omega) * \boldsymbol{U}_{\mathrm{LIN}}(z,\omega) * \boldsymbol{U}_{\mathrm{LIN}}(z,\omega)$, see below.

We find that the (double) product of the time-domain optical signal turns into a (double) convolution in frequency domain. The definition of the convolution, see (2.38), is used to exemplarily show the result for the first (of the four) convolution cross products

$$\begin{aligned} U_{\mathrm{LIN,x}}(\omega) * U_{\mathrm{LIN,x}}^{*}(-\omega) * U_{\mathrm{LIN,x}}(\omega) &= \left(\int_{\mathbb{R}} U_{\mathrm{LIN,x}}(\omega_1)U_{\mathrm{LIN,x}}^{*}(\omega_1-\omega)\mathrm{d}\omega_1\right) * U_{\mathrm{LIN,x}}(\omega) \\ &= \int_{\mathbb{R}^2} U_{\mathrm{LIN,x}}(\omega_1)U_{\mathrm{LIN,x}}^{*}(\omega_1-\omega_2')U_{\mathrm{LIN,x}}(\omega-\omega_2')\,\mathrm{d}\omega_1\mathrm{d}\omega_2' \\ &= \int_{\mathbb{R}^2} U_{\mathrm{LIN,x}}(\omega_1)U_{\mathrm{LIN,x}}^{*}(\omega_2)U_{\mathrm{LIN,x}}(\omega-\omega_0+\omega_2)\,\mathrm{d}^2\boldsymbol{\omega} \\ &= \int_{\mathbb{R}^2} U_{\mathrm{LIN,x}}(\omega+\upsilon_1)U_{\mathrm{LIN,x}}^{*}(\omega+\upsilon_1+\upsilon_2)U_{\mathrm{LIN,x}}(\omega+\upsilon_2)\,\mathrm{d}^2\boldsymbol{\upsilon}\,, \end{aligned} \tag{4.19}$$

where we omit the z-dependency of the optical signal for short notation. Similar expressions are obtained for the three remaining x/y-cross products. In the literature, one may find also find variations of the form in (4.19) depending on the exact definition of the convolution integral.

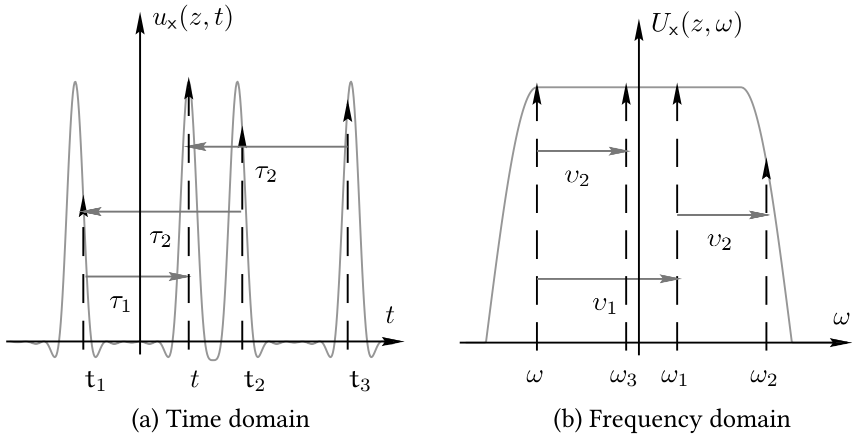


Figure 4.3: Definitions of the auxiliary (i.e., helper) time and frequency variables. Note, that both τ_1, τ_2 and υ_1, υ_2 can take any value, positive and negative, in $\mathbb{R}$. The gray curve drawn left and right is a fictitious signal to indicate the spectral support in time- and frequency domain.

We typically use the common (*absolute*) frequency variables $[\omega_1,\omega_2,\omega_3]^{\mathsf{T}}$ and the common (*relative* to ω) auxiliary frequency variables $[\upsilon_1,\upsilon_2]^{\mathsf{T}}$ related by definition as

$$\omega_1 \stackrel{\text{def}}{=} \omega + \upsilon_1 \tag{4.20}$$

$$\omega_2 \stackrel{\text{def}}{=} \omega + \upsilon_1 + \upsilon_2 \tag{4.21}$$

$$\omega_3 \stackrel{\text{def}}{=} \omega - \omega_1 + \omega_2 = \omega + \upsilon_2\,, \tag{4.22}$$

to express the optical signal $\boldsymbol{U}(z,\cdot)$. Figure 4.3 summarizes the definitions of the time- and frequency variables that are used throughout this text.

We can now write the optical perturbation $\boldsymbol{U}(z,\omega)$ at $z=L$ with $\sqrt{\mathcal{P}(L)}=1$ using again the concise vectorial notation to obtain [VSB02, Eq. (9)]

$$\Delta\boldsymbol{U}(L,\omega) = -\,\mathrm{j}\bar{\gamma}\frac{8}{9}\frac{1}{(2\pi)^2}H_{\mathrm{CD}}(L,\omega)\int_0^L H_{\mathrm{C}}^{-1}(\zeta,\omega) \times \int_{\mathbb{R}^2}\boldsymbol{U}_{\mathrm{LIN}}(\zeta,\underbrace{\omega_3}_{\omega_3=\omega-\omega_1+\omega_2})\boldsymbol{U}_{\mathrm{LIN}}^{\mathsf{H}}(\zeta,\omega_2)\boldsymbol{U}_{\mathrm{LIN}}(\zeta,\omega_1)\,\mathrm{d}^2\boldsymbol{\omega}\,\mathrm{d}\zeta\,. \tag{4.23}$$

The integral over $\mathbb{R}^2$ in (4.23) can be equivalently performed w.r.t. $[\omega_1,\omega_2]^{\mathsf{T}}$ or $[\upsilon_1,\upsilon_2]^{\mathsf{T}}$. We recall from (3.63) that the *linearly propagating* optical signal $\boldsymbol{U}_{\mathrm{LIN}}(\zeta,\omega)$ can be expressed as the product of the *linear transfer function* $H_{\mathrm{C}}(\zeta,\omega)$ and the optical signal at the *input* of the transmission $\boldsymbol{U}_{\mathrm{LIN}}(0,\omega)\equiv\boldsymbol{U}(0,\omega)$.

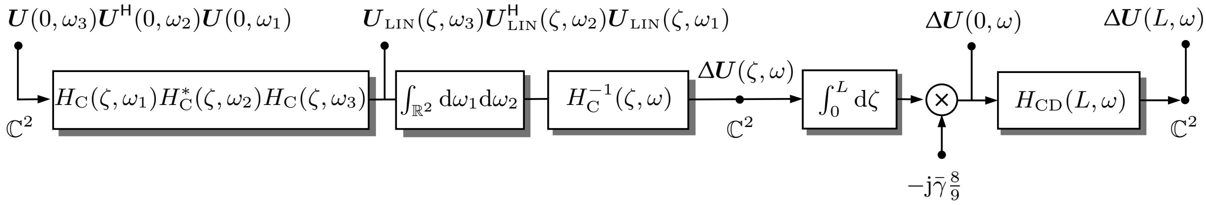


Figure 4.4: Frequency-domain representation of the *parallel* fiber-optical channel model. The two-fold product in time domain turns into a two-fold convolution in frequency domain over the auxiliary frequency variables $[\omega_1,\omega_2]^{\mathsf{T}}$. In contrast to Fig. 4.2, the integration over both frequency and space is not explicitly illustrated by the branches.

In Figure 4.4, we show the block diagram related to (4.23). The two-fold product of optical signals at the input of the fiber $\boldsymbol{U}(0,\omega_3)\boldsymbol{U}^{\mathsf{H}}(0,\omega_2)\boldsymbol{U}(0,\omega_1)$ is weighted with the respective two-fold product of linear channel transfer functions associated with position ζ and frequency $[\omega_1,\omega_2,\omega_3]^{\mathsf{T}}$. The third frequency ω_3 is constrained with respect to the FWM selection rule according to (4.22) such that the mixing between the optical signals generates the *local* perturbation at frequency $\omega=\omega_3+\omega_1-\omega_2$. The local perturbation at ω is hence obtained by integrating over all possible pairs of $[\omega_1,\omega_2]^{\mathsf{T}}$. The result, i.e., the *local* perturbation due to all FWM combinations, is then propagated back to the fiber input at $z=0$ via the inverse linear transfer function $H_{\mathrm{C}}^{-1}(\zeta,\omega)$. The total (path-) *accumulated* perturbation is attained via integration of $\Delta\boldsymbol{U}(\zeta,\omega)$ over all positions ζ, i.e., via summing the local perturbations generated along the transmission link from $z=0$ to $z=L$. The final dispersion operator $H_{\mathrm{CD}}(L,\omega)$ transforms the accumulated perturbation to the receiver at $z=L$.

The linear transfer function from (3.63) is substituted into (4.23) and the order of integration is exchanged. We yield [VSB02, Eq. (12)] [Mec11, Eq. (6.9)]

$$\Delta\boldsymbol{U}(L,\omega) = -\,\mathrm{j}\bar{\gamma}\frac{8}{9}\frac{1}{(2\pi)^2}H_{\mathrm{CD}}(L,\omega)\int_{\mathbb{R}^2}\int_0^L H_{\mathrm{C}}^{-1}(\zeta,\omega)H_{\mathrm{C}}(\zeta,\omega_3)H_{\mathrm{C}}^*(\zeta,\omega_2)H_{\mathrm{C}}(\zeta,\omega_1) \times \boldsymbol{U}(0,\underbrace{\omega_3}_{\omega_3=\omega-\omega_1+\omega_2})\boldsymbol{U}^{\mathsf{H}}(0,\omega_2)\boldsymbol{U}(0,\omega_1)\,\mathrm{d}\zeta\,\mathrm{d}^2\boldsymbol{\omega} \tag{4.24}$$

$$= -\,\mathrm{j}\bar{\gamma}\frac{8}{9}\frac{1}{(2\pi)^2}H_{\mathrm{CD}}(L,\omega)\int_{\mathbb{R}^2}\boldsymbol{U}(0,\omega_3)\boldsymbol{U}^{\mathsf{H}}(0,\omega_2)\boldsymbol{U}(0,\omega_1) \times \int_0^L \underbrace{H_{\mathrm{C}}^{-1}(\zeta,\omega)H_{\mathrm{C}}(\zeta,\omega_3)H_{\mathrm{C}}^*(\zeta,\omega_2)H_{\mathrm{C}}(\zeta,\omega_1)}_{\exp\left(\mathcal{G}(\zeta)-\mathrm{j}\frac{\mathcal{B}(\zeta)}{2}(-\omega^2+\omega_3^2-\omega_2^2+\omega_1^2)\right)}\,\mathrm{d}\zeta\,\mathrm{d}^2\boldsymbol{\omega}\,, \tag{4.25}$$

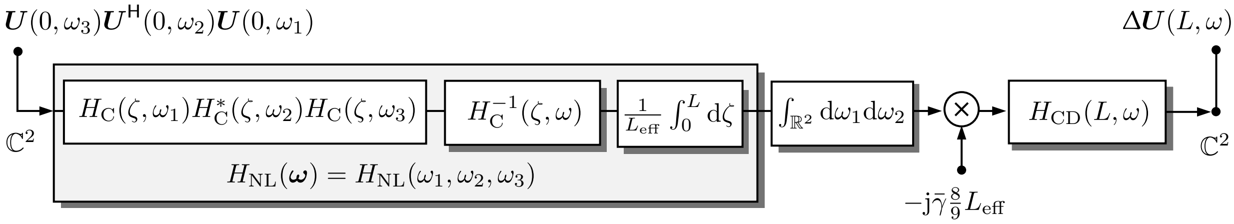


Figure 4.5: Frequency-domain representation of the *parallel* fiber-optical channel model where the inner integral over the spatial variable ζ reflects the (path-) accumulated phasor (due to the linear propagation) for any $[\omega_1,\omega_2,\omega_3]^{\mathsf{T}}$ tuple which is subsumed in the so-called *nonlinear transfer function* $H_{\mathrm{NL}}(\omega_2-\omega_3,\omega_2-\omega_1)$. The outer (double) integral sums the accumulated phasors for all $[\omega_1,\omega_2]^{\mathsf{T}}$ pairs.

where we can apply the useful equivalence

$$\upsilon_1\upsilon_2 = (\underbrace{\omega_1-\omega}_{(\omega_2-\omega_3)})(\omega_2-\omega_1) = \frac{1}{2}(\omega^2-\omega_1^2+\omega_2^2-(\underbrace{\omega-\omega_1+\omega_2}_{\omega_3})^2)\,, \tag{4.26}$$

to arrive at the solution to the first-order RP method in frequency domain [VSB02, Eq. (12)], [LHP+05, Eq. (2)], [Wei06, Eq. (4)], [LLH+12, Eq. (24)–(27)]

$$\begin{aligned}\Delta\boldsymbol{U}(L,\omega) =& -\mathrm{j}\bar{\gamma}\frac{8}{9}L_{\mathrm{eff}}\frac{1}{(2\pi)^2}H_{\mathrm{CD}}(L,\omega)\int_{\mathbb{R}^2}\boldsymbol{U}(0,\omega+\upsilon_2)\boldsymbol{U}^{\mathsf{H}}(0,\omega+\upsilon_1+\upsilon_2)\boldsymbol{U}(0,\omega+\upsilon_1)\\ &\times\underbrace{\frac{1}{L_{\mathrm{eff}}}\int_0^L\exp\left(\mathcal{G}(\zeta)+\mathrm{j}\upsilon_1\upsilon_2\mathcal{B}(\zeta)\right)\mathrm{d}\zeta}_{H_{\mathrm{NL}}(\boldsymbol{\upsilon})=H_{\mathrm{NL}}(\upsilon_1,\upsilon_2)}\;\mathrm{d}^2\boldsymbol{\upsilon} \qquad (4.27)\\ =& -\mathrm{j}\bar{\gamma}\frac{8}{9}L_{\mathrm{eff}}\frac{1}{(2\pi)^2}H_{\mathrm{CD}}(L,\omega)\int_{\mathbb{R}^2}\boldsymbol{U}(0,\underbrace{\omega-\omega_1+\omega_2}_{\omega_3})\boldsymbol{U}^{\mathsf{H}}(0,\omega_2)\boldsymbol{U}(0,\omega_1)\\ &\times\underbrace{\frac{1}{L_{\mathrm{eff}}}\int_0^L\exp\left(\mathcal{G}(\zeta)+\mathrm{j}(\omega_2-\omega_3)(\omega_2-\omega_1)\mathcal{B}(\zeta)\right)\mathrm{d}\zeta}_{H_{\mathrm{NL}}(\boldsymbol{\omega})=H_{\mathrm{NL}}(\omega_1,\omega_2,\omega_3)}\;\mathrm{d}^2\boldsymbol{\omega}\,. \qquad (4.28)\end{aligned}$$

Equations (4.27)–(4.28) constitute the central result of the continuous-time first-order RP method in frequency domain. It is remarkable, that, in this form, the integrand can be factored into a term that only depends on the optical signal at the *input* of the transmission system $\boldsymbol{U}(0,\omega)$, and a term that only depends on the characteristics of the transmission link, i.e., the gain/loss and dispersion profile of the link. The latter is commonly referred to as *nonlinear transfer function* and either defined as a function of difference frequencies $H_{\mathrm{NL}}(\boldsymbol{\upsilon}) = H_{\mathrm{NL}}(\upsilon_1,\upsilon_2)$, see (4.27), or as a function of absolute frequencies $H_{\mathrm{NL}}(\boldsymbol{\omega}) = H_{\mathrm{NL}}(\omega_1,\omega_2,\omega_3)$, see (4.28). Section 4.1 provides an in-depth discussion on the characteristics of the nonlinear transfer function.

Relation to the Third-Order Volterra Operator It is instructive to relate the solution of the perturbation approach in (4.27)–(4.28) to the theory of Volterra series as introduced in Section 2.1.3.2. The first application of VSTF for solving the nonlinear Schrödinger equation were published in [PBP97, XBP01]. In [VSB02], Vannucci *et al.* shows that the first-order perturbation approach is equivalent to the third-order Volterra operator. This can be intuitively understood considering the following: Figure 4.6 shows the basic (third-order) nonlinear system

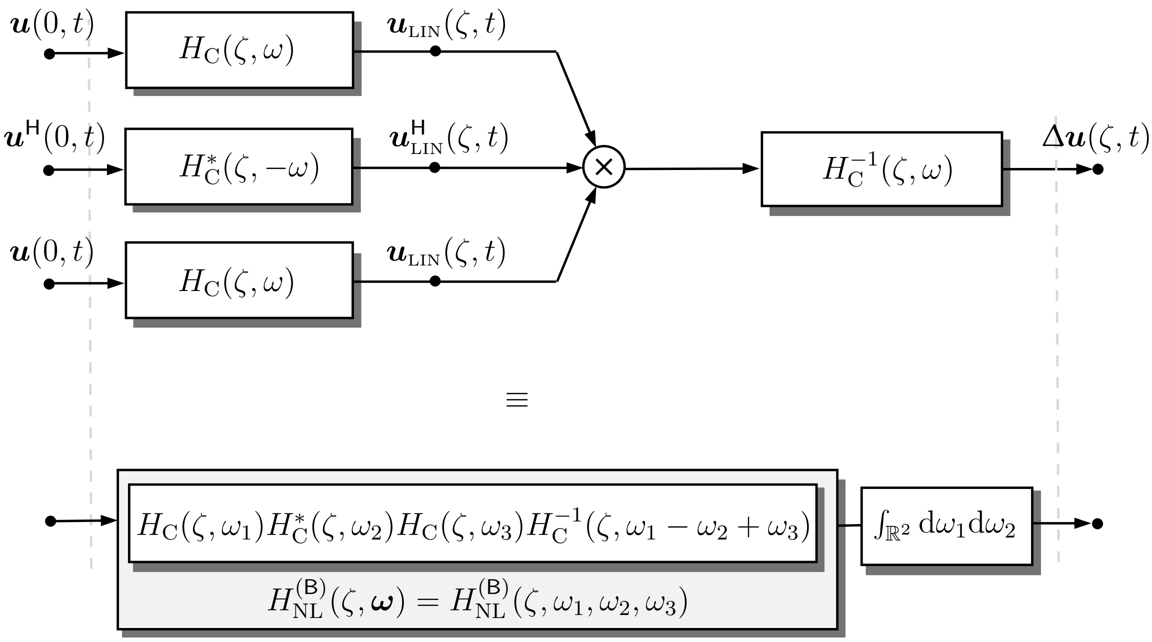


Figure 4.6: Block diagram of the basic third-order Volterra system (compare with Figure 2.1) constituting a single (spatial) branch of the *parallel fiber model* from Figure 4.2.

corresponding to a single (spatial) branch related to position ζ of the *parallel fiber model* from Figure 4.2. Here, the optical signal $\boldsymbol{u}(\zeta,t)$ and its Hermitian conjugate $\boldsymbol{u}^{\mathsf{H}}(\zeta,t)$ are explicitly drawn in a direct correspondence to the *basic third-order Volterra system* in Figure 2.1. Using the relation from (2.66), one can immediately derive the frequency-domain, third-order Volterra kernel of this particular (ζ-dependent) *basic* system as

$$H^{(\mathsf{B})}_{\mathrm{NL}}(\zeta,\boldsymbol{\omega}) = H^{(\mathsf{B})}_{\mathrm{NL}}(\zeta,\omega_1,\omega_2,\omega_3) = \underbrace{H_{\mathrm{C}}(\zeta,\omega_1)}_{H^{(1)}_{\mathrm{LTI}}(\omega_1)}\,\underbrace{H^{*}_{\mathrm{C}}(\zeta,\omega_2)}_{H^{(2)}_{\mathrm{LTI}}(\omega_2)}\,\underbrace{H_{\mathrm{C}}(\zeta,\omega_3)}_{H^{(3)}_{\mathrm{LTI}}(\omega_3)}\,\underbrace{H_{\mathrm{C}}^{-1}(\zeta,\underbrace{\omega_1-\omega_2+\omega_3}_{\omega})}_{H^{(4)}_{\mathrm{LTI}}(\omega)}\,, \tag{4.29}$$

where the Hermitian conjugation induces the sign swap of ω_2 compared to (2.66). The kernel is apparently partially symmetric, i.e.,

$$H^{(\mathsf{B})}_{\mathrm{NL}}(\zeta,\omega_1,\omega_2,\omega_3) = H^{(\mathsf{B})}_{\mathrm{NL}}(\zeta,\omega_3,\omega_2,\omega_1)\,, \tag{4.30}$$

due to the equivalence $H^{(1)}_{\mathrm{LTI}}(\omega) = H^{(3)}_{\mathrm{LTI}}(\omega) = H_{\mathrm{C}}(\zeta,\omega)$.

We can now use the theory of *parallel* concatenated Volterra systems, see (2.69), to obtain the (normalized) nonlinear transfer function as the integral over all spatial positions $0 \le \zeta \le L$ in the limit $\mathrm{d}\zeta \to 0$, and obtain

$$\begin{aligned} H_{\mathrm{NL}}(\boldsymbol{\omega}) &= \frac{1}{L_{\mathrm{eff}}}\int_0^L H^{(\mathsf{B})}_{\mathrm{NL}}(\zeta,\boldsymbol{\omega})\,\mathrm{d}\zeta \\ &= \frac{1}{L_{\mathrm{eff}}}\int_0^L H_{\mathrm{C}}(\zeta,\omega_1)H^{*}_{\mathrm{C}}(\zeta,\omega_2)H_{\mathrm{C}}(\zeta,\omega_3)H_{\mathrm{C}}^{-1}(\zeta,\underbrace{\omega_1-\omega_2+\omega_3}_{\omega})\,\mathrm{d}\zeta\,, \end{aligned} \tag{4.31}$$

where we used the effective length of the link L_{eff} to normalize the nonlinear transfer function to $H_{\text{NL}}(\mathbf{0}) = 1$. The same result can be found in the derivation of the first-order perturbation method, e.g., (4.24)–(4.25).

The Nonlinear Transfer-Function We now formally define the (normalized) *nonlinear transfer function* in terms of *absolute frequencies* as[2]

$$\begin{aligned} H_{\text{NL}}(\boldsymbol{\omega}) &= H_{\text{NL}}(\omega_1, \omega_2, \omega_3) \\ &\stackrel{\text{def}}{=} \frac{1}{L_{\text{eff}}} \int_0^L H_{\text{C}}(\zeta, \omega_1) H_{\text{C}}^*(\zeta, \omega_2) H_{\text{C}}(\zeta, \omega_3) H_{\text{C}}^{-1}(\zeta, \omega_1 - \omega_2 + \omega_3)\, \mathrm{d}\zeta \qquad (4.32) \\ &= \frac{1}{L_{\text{eff}}} \int_0^L \exp\big(\, \mathcal{G}(\zeta) + \mathrm{j}(\omega_2 - \omega_3)(\omega_2 - \omega_1)\mathcal{B}(\zeta)\,\big)\, \mathrm{d}\zeta\,, \qquad (4.33) \end{aligned}$$

such that the transfer function is *normalized* and *dimensionless*, i.e., with (3.27) with have

$$H_{\text{NL}}(\mathbf{0}) = \frac{1}{L_{\text{eff}}} \int_0^L \exp\left(\mathcal{G}(\zeta)\right)\, \mathrm{d}\zeta = 1\,. \qquad (4.34)$$

The form of (4.32) is in direct correspondence with the definition of the Volterra operator in its *standard* form, see (2.58), which explicitly recovers the dependence on the linear transfer function $H_{\text{C}}(\zeta, \omega)$ evaluated at the four involved frequencies and then *path-averaged* over the transmission link, cf. Figure 4.6.

It is noteworthy, that the nonlinear transfer function contains all the relevant information about the transmission link characterized by the dispersion profile $\mathcal{B}(z)$ (including CD pre-compensation $\mathcal{B}_0$, cf. (3.47)) and the gain/loss profile $\mathcal{P}(z)$.

Equivalently, the nonlinear transfer function can also be defined in terms of *relative frequencies* (aka. *difference frequencies*) as

$$H_{\text{NL}}(\boldsymbol{\upsilon}) = H_{\text{NL}}(\upsilon_1, \upsilon_2) \stackrel{\text{def}}{=} \frac{1}{L_{\text{eff}}} \int_0^L \exp\big(\, \mathcal{G}(\zeta) + \mathrm{j}\upsilon_1\upsilon_2\mathcal{B}(\zeta)\,\big)\, \mathrm{d}\zeta\,, \qquad (4.35)$$

which reduces the dimension of its domain $\mathrm{dom}(H_{\text{NL}})$ by one due to the simple relation $\upsilon_1 = \omega_2 - \omega_3$ and $\upsilon_2 = \omega_2 - \omega_1$ such that

$$H_{\text{NL}}(\upsilon_1, \upsilon_2) = H_{\text{NL}}(\omega_2 - \omega_3, \omega_2 - \omega_1)\,, \qquad (4.36)$$

and the change of variables creates symmetry w.r.t. $[\upsilon_1, \upsilon_2]^{\mathsf{T}}$, i.e., $H_{\text{NL}}(\upsilon_1, \upsilon_2) = H_{\text{NL}}(\upsilon_2, \upsilon_1)$. The form in (4.35) already appears in, e.g., (4.27) and will be mainly used in the following, since, as will be shown below, it relates via a 2D Fourier transform to the corresponding kernel $h_{\text{NL}}(\boldsymbol{\tau}) = h_{\text{NL}}(\tau_1, \tau_2)$ of the third-order time-domain Volterra operator.

Similar to above, $H_{\text{NL}}(\upsilon_1, \upsilon_2)$ depends in fact on the product of (relative) frequencies

$$\xi \stackrel{\text{def}}{=} \upsilon_1\upsilon_2 = (\omega_2 - \omega_3)(\omega_2 - \omega_1) \qquad (4.37)$$

[2]We will define the function $H_{\text{NL}}(\cdot)$, where, depending on the (number of) arguments, one of the following definitions in (4.33), (4.35), or (4.38) is to be used.

which again reduces the dimension of $\mathrm{dom}(H_{\mathrm{NL}})$ to finally one, and we define [Wei06, Eq. (5)]

$$H_{\mathrm{NL}}(\xi) \stackrel{\text{def}}{=} \frac{1}{L_{\mathrm{eff}}} \int_0^L \exp\left(\mathcal{G}(\zeta) + \mathrm{j}\,\xi\,\mathcal{B}(\zeta) \right) \mathrm{d}\zeta\,. \tag{4.38}$$

The form in (4.38) is related via a one-dimensional (1D) Fourier transform w.r.t. ξ to the so-called *power-weighted dispersion distribution* [Wei06], which was used to, e.g., optimize dispersion-managed systems. Equation (4.38) is also well-suited to visualize the nonlinear transfer function w.r.t. the scalar variable ξ, or to investigate the *poles* and *zeros* of the nonlinear transfer function, as the *image* of $H_{\mathrm{NL}}(\cdot)$, i.e., set of all output values, is apparently the same for all three variants (4.32), (4.35), and (4.38).

The nonlinear transfer function can be interpreted as a measure of the so-termed *phase matching condition.* The phase mismatch due to dispersion, i.e., the difference in the propagation constant, see (3.3), between the four involved frequencies of the nonlinear mixing, is given by [Agr10, Eq. (6.3.19)]

$$\begin{aligned}\beta(z,\omega) - \beta(z,\omega_1) + \beta(z,\omega_2) - \beta(z,\omega_3) &= (\omega^2 - \omega_1^2 + \omega_2^2 - (\underbrace{\omega - \omega_1 + \omega_2}_{\omega_3})^2)\,\frac{\beta_2(z)}{2} \\ &= (\omega_1 - \omega)(\omega_2 - \omega_1)\beta_2(z) = \upsilon_1\upsilon_2\,\beta_2(z)\,, \end{aligned} \tag{4.39}$$

where on the right-hand side of (4.39) only the second-order Taylor coefficient $\beta_2(z)$ is considered. This phase matching condition is present in the argument of the exponential in (4.35) in terms of accumulated dispersion $\upsilon_1\upsilon_2\,\mathcal{B}(z)$. Phase matching is obtained when the nonlinear transfer function is maximized; compare Figure 4.7. Simply speaking, the mixing is *good*, i.e., $|H_{\mathrm{NL}}(\boldsymbol{\upsilon})| \to 1$, if the involved frequency components propagate *pair-wise* at approximately the same (relative) *velocity* over a section of the link with relatively high signal power, i.e., $\mathcal{G}(z) \to 0$. This is achieved for low chromatic dispersion $\beta_2(z)$, i.e., close to the *zero-dispersion wavelength* where $\beta_2(z) \to 0$, or either for $\upsilon_1 \to 0$ or $\upsilon_2 \to 0$, which relates to the pair-wise propagation of the two frequency components, cf. Figure 4.3. The squared magnitude $|H_{\mathrm{NL}}(\boldsymbol{\upsilon})|^2$ is also called the *FWM efficiency* [Agr10, (6.3.18)].

The spectral width of the nonlinear transfer function $|H_{\mathrm{NL}}(\xi)|$ in relation to the probe's spectral extend, see (2.82), is a measure of *intra-channel* (i.e., SCI) nonlinear effects [FBP09, Fis09]. Recall, e.g., from Figure 3.5 (right) and (3.57), that the length of the CD impulse response (normalized to the symbol period) scales with $(1+\rho)\,L/L_{\mathrm{D}}$. Hence, the ratio $(1+\rho)\,L_{\mathrm{eff}}/L_{\mathrm{D}}$ is proportional to the number of overlapping pulses due to dispersion (within the probe channel) which interact nonlinearly over the link (i.e., where the signal power is *effectively* high). It will turn out that the width of the nonlinear transfer function (over $\xi/(2\pi R_{\mathrm{s}})^2$) scales *inversely* with the number of overlapping pulses within L_{eff}. This was first reported by Louchet *et al.* [LHP$^+$05] where the perception of a *nonlinear diffusion bandwidth* (i.e., a bandwidth proportional to $\bar{\alpha}/\bar{\beta}_2$ of an equivalent single-span model)[3] was developed.

[3]The notion of *diffusion* stems form an analogy with a diffusion process in which the *diffusion length* (here, equivalent to the *nonlinear diffusion bandwidth*) is a measure of how far a concentration propagates, i.e. here, a measure of the nonlinear interaction range of a given spectral component [Lou06, P. 33].

Example 4.1: The single-span nonlinear transfer function

We consider the nonlinear transfer function $H_{\mathrm{NL}}(\boldsymbol{\upsilon})$ of a single span with typical fiber parameters (cf. ITU recommendation G.652) and end-of-span, lumped amplification. Figure 4.7 shows the squared-magnitude $|H_{\mathrm{NL}}(\boldsymbol{\upsilon})|^2$ in logarithmic scale.

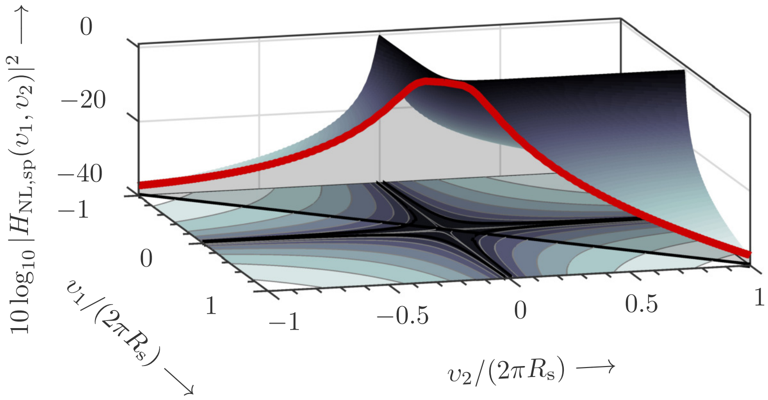


Figure 4.7: Squared-magnitude in logarithmic scale of the *single-span* nonlinear transfer function for $\bar{\beta}_2 = -21\,\mathrm{ps}^2/\mathrm{km}$, $\mathcal{B}_0 = 0\,\mathrm{ps}^2$, $10\log_{10} \mathrm{e}^{\bar{\alpha}} = 0.2\,\mathrm{dB/km}$ and $L = L_{\mathrm{sp}} = 100\,\mathrm{km}$ over the *difference* frequencies υ_1 and υ_2 normalized to $R_\mathrm{s} = 64\,\mathrm{GBd}$ [FFF20]. The red line denotes $H_{\mathrm{NL,,sp}}(\xi)$ which only depends on the scalar $\xi = \upsilon_1\upsilon_2$. (Part for $\upsilon_1 > \upsilon_2$ not shown).

For this example, we set the dispersion profile to

$$\mathcal{B}(z) = \bar{\beta}_2 z\,,$$

i.e., unmanaged/no inline dispersion compensation with $\bar{\beta}_2 = -21\ \mathrm{ps}^2/\mathrm{km}$, and the (logarithmic) power profile to

$$\mathcal{G}(z) = -\bar{\alpha} z + \bar{\alpha} L_{\mathrm{sp}}\,\delta(z - L_{\mathrm{sp}})\,,$$

with $10\log_{10} \mathrm{e}^{\bar{\alpha}} = 0.2\ \mathrm{dB/km}$ and $L = L_{\mathrm{sp}} = 100\ \mathrm{km}$. For visualization, the difference frequencies $[\upsilon_1, \upsilon_2]^\mathsf{T}$ are normalized by the probe's symbol rate $2\pi R_\mathrm{s} = 2\pi \cdot 64\ \mathrm{GBd}$ to relate the spectral width of $H_{\mathrm{NL}}(\boldsymbol{\upsilon})$ to the spectral width of the channel under consideration (or, more precisely, to the spectral width of the probe's Nyquist interval).
It can be observed that $H_{\mathrm{NL}}(\upsilon_1, \upsilon_2)$ has features of a hyperbolic function in two dimensions (cf. the projected contour) as it depends on the product $\xi = \upsilon_1\upsilon_2$, see also discussion in [Pog12, VIII. A]. The bold red line drawn into the diagonal cross section in Figure 4.7 corresponds to $H_{\mathrm{NL}}(\xi/(2\pi R_\mathrm{s})^2)$ over the (normalized) scalar variable $\xi = \upsilon_1\upsilon_2$.

This insight has led to the definition of the *map strength* [AHB01, SBO07, BSO08]. The map strength $\mathcal{S}$ is a measure of the temporal extend (i.e., the induced memory) of the system's response due to *intra-channel* nonlinear effects (assuming no in-line dispersion compensation). It is defined as

$$\mathcal{S} \overset{\text{def}}{=} \bar{\beta}_2 L_{\mathrm{eff}}/(2\pi)\,, \tag{4.40}$$

and has units s^2 (*squared seconds*). Accordingly, the map strength *normalized* to the probe's symbol rate $R_\mathrm{s} = 1/T$ is defined as

$$\mathcal{S}_{T,\rho} \stackrel{\text{def}}{=} (2\pi R_\mathrm{s})^2\, \mathcal{S} = 2\pi R_\mathrm{s}^2 \bar{\beta}_2 L_\mathrm{eff} = \operatorname{sign}(\bar{\beta}_2)\, L_\mathrm{eff}/L_\mathrm{D}\,, \tag{4.41}$$

which immediately recovers the ratio $L_\mathrm{eff}/L_\mathrm{D}$, see (3.27), (3.56), and, hence, inversely relates to the spectral width of the (normalized) nonlinear transfer function $H_\mathrm{NL}(\xi/(2\pi R_\mathrm{s})^2)$, see discussion below.

Example 4.2: The spectral width of the nonlinear transfer function

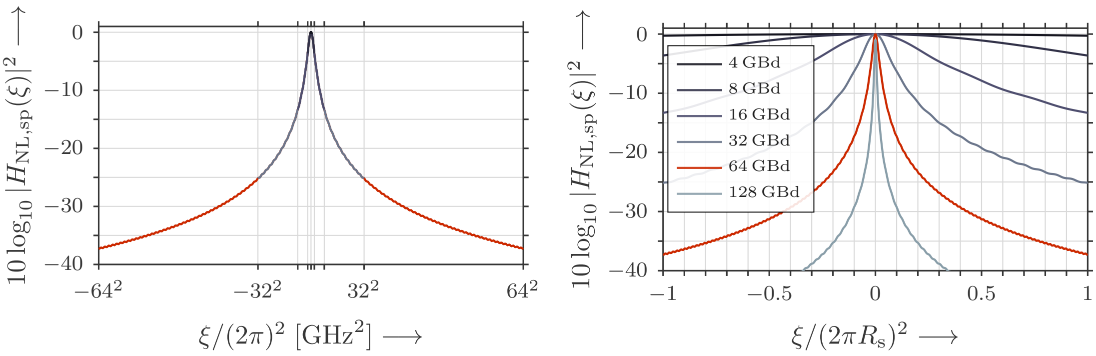


Figure 4.8: Squared-magnitude in logarithmic scale of the *single-span* nonlinear transfer function for $\bar{\beta}_2 = -21\ \mathrm{ps^2/km}$, $\mathcal{B}_0 = 0\ \mathrm{ps}^2$, $10\log_{10}\mathrm{e}^{\bar{\alpha}} = 0.2\ \mathrm{dB/km}$ and $L = L_\mathrm{sp} = 100\ \mathrm{km}$ over $\xi = \upsilon_1\upsilon_2$. The normalization by $(2\pi R_\mathrm{s})^2$ relates $H_\mathrm{NL}(\xi)$ to the probe's spectral width [FFF20].

In Figure 4.8, the nonlinear transfer function $|H_\mathrm{NL}(\xi)|^2$ is shown in logarithmic scale over the variable $\xi/(2\pi)^2$ in units of Hz^2 (*squared Hertz*), and over the normalized variable $\xi/(2\pi R_\mathrm{s})^2$. The corresponding segment of $H_\mathrm{NL}(\xi)$ is color-coded for the following set of symbol rates

$$R_\mathrm{s} = \{\, 4, 8, 16, 32, 64, 128 \,\}\ \mathrm{GBd}\,.$$

Apart from the symbol rate R_s used in the normalization of the axis, the same link parameters are used as in Example 4.1. We observe that the nonlinear transfer function has the form an equivalent low-pass filter w.r.t. ξ, with a weak ripple superimposed. The width of $|H_\mathrm{NL}(\xi/(2\pi R_\mathrm{s})^2)|^2$ is proportional to $1/\mathcal{S}_{T,\rho} \propto L_\mathrm{D}/L_\mathrm{eff} \propto R_\mathrm{s}^{-2}$, i.e., doubling R_s *reduces* the spectral width by a factor of 4.
The trace with $R_\mathrm{s} = 64\,\mathrm{GBd}$ is identical to Figure 4.7. For this case, span length, the dispersion length, and the effective length take the following values

$$L = L_\mathrm{sp} = 100\ \mathrm{km}\,, \qquad L_\mathrm{D} = 1.85\ \mathrm{km}\,, \qquad L_\mathrm{eff} = 21.50\ \mathrm{km}\,,$$

e.g., also see the red marker in Figure 3.7 (left). The normalized map strength takes the value

$$\mathcal{S}_{T,\rho} = -11.62\,.$$

The number of overlapping basic pulses subject to chromatic dispersion over the effective length $(1+\rho)L_\mathrm{eff}/L_\mathrm{D} = (1+\rho)|\mathcal{S}_{T,\rho}|$ is approximately 14 given $\rho = 0.2$. This is in the same order of the system's intra-channel nonlinear memory given the selected parameters.

Closed-Form Single-Span Solutions Closed-form analytical solutions to the nonlinear transfer function in (4.35) can be obtained for single-span or homogeneous multi-span systems.

Considering only a single span and *average* link parameters (i.e., no z-dependency of the link parameters $\alpha(z)$, $g(z)$, $\beta_2(z)$), the power and dispersion profile reduce to the simple expressions already given in Example 4.1. Then, the *single-span* nonlinear transfer function $H_{\mathrm{NL,sp}}(\upsilon_1, \upsilon_2)$ can be written in closed-form by

$$H_{\mathrm{NL,sp}}(\upsilon_1, \upsilon_2) = \frac{1}{L_{\mathrm{eff}}} \int_0^{L_{\mathrm{sp}}} \exp\left(-\bar{\alpha}\zeta + \mathrm{j}\upsilon_1\upsilon_2\bar{\beta}_2\zeta\right) \mathrm{d}\zeta \tag{4.42}$$

$$= \frac{1}{L_{\mathrm{eff}}} \frac{1 - \exp\left(-\bar{\alpha}L_{\mathrm{sp}} + \mathrm{j}\upsilon_1\upsilon_2\bar{\beta}_2 L_{\mathrm{sp}}\right)}{\bar{\alpha} - \mathrm{j}\upsilon_1\upsilon_2\bar{\beta}_2}, \tag{4.43}$$

where L_{eff} is the effective length of a *single-span*, i.e., $L_{\mathrm{eff}} = L_{\mathrm{eff,sp}}$, as defined in (3.29).

Asymptotic Limits — Lossless Transmission It is instructive to study the asymptotic limits of $H_{\mathrm{NL,sp}}(\upsilon_1, \upsilon_2)$. We first study the case of lossless transmission with $\mathcal{G}(z) = 0$ (equivalently realized by a lossless fiber with $\alpha(z) \to 0$, or by ideal distributed amplification with $g(z) = \alpha(z)$). We define the nonlinear transfer function of a *lossless* single-span link as [AH02a, (6)]

$$H_{\mathrm{NL,sp}}^{\alpha\to 0}(\upsilon_1, \upsilon_2) \overset{\mathrm{def}}{=} \lim_{\bar{\alpha}\to 0} H_{\mathrm{NL,sp}}(\upsilon_1, \upsilon_2) = \frac{1}{L_{\mathrm{sp}}} \frac{1 - \exp\left(\mathrm{j}\upsilon_1\upsilon_2\bar{\beta}_2 L_{\mathrm{sp}}\right)}{-\mathrm{j}\upsilon_1\upsilon_2\bar{\beta}_2} \tag{4.44}$$

$$= \exp\left(\mathrm{j}\upsilon_1\upsilon_2\bar{\beta}_2\frac{L_{\mathrm{sp}}}{2}\right) \mathrm{si}\left(\upsilon_1\upsilon_2\bar{\beta}_2\frac{L_{\mathrm{sp}}}{2}\right), \tag{4.45}$$

where we used that $L_{\mathrm{sp}} = \lim_{\alpha\to 0} L_{\mathrm{eff,sp}}$ from (3.31). Similarly, the map strength $\mathcal{S}$, as given in (4.40), in the limit of a lossless transmission is defined as the *zero-attenuation* map strength

$$\mathcal{S}_0 \overset{\mathrm{def}}{=} \lim_{\alpha\to 0} \mathcal{S} = \bar{\beta}_2 L_{\mathrm{sp}}/(2\pi) \tag{4.46}$$

$$\mathcal{S}_{T,0} \overset{\mathrm{def}}{=} (2\pi R_{\mathrm{s}})^2 \lim_{\alpha\to 0} \mathcal{S} = 2\pi R_{\mathrm{s}}^2 \bar{\beta}_2 L_{\mathrm{sp}} = \mathrm{sign}(\bar{\beta}_2)\, L_{\mathrm{sp}}/L_{\mathrm{D}}\,. \tag{4.47}$$

Since the (normalized) map strength is a measure of the number of nonlinear interacting pulses over the effective length, in the limit of lossless transmission it simply becomes the ratio of the span length and the dispersion length. The map strength $\mathcal{S}_0$ fully characterizes the nonlinear transfer function for lossless transmission, and we obtain the concise expression [AHB01, Eq. (3)]

$$H_{\mathrm{NL,sp}}^{\alpha\to 0}(\upsilon_1, \upsilon_2) = \exp\left(\mathrm{j}\pi\upsilon_1\upsilon_2\,\mathcal{S}_0\right)\ \mathrm{sinc}\left(\upsilon_1\upsilon_2\,\mathcal{S}_0\right) \tag{4.48}$$

$$= \frac{\sin(2\pi\mathcal{S}_0\,\upsilon_1\upsilon_2)}{2\pi\mathcal{S}_0\,\upsilon_1\upsilon_2} + \mathrm{j}\frac{1}{2\pi\mathcal{S}_0\,\upsilon_1\upsilon_2} - \mathrm{j}\frac{\cos(2\pi\mathcal{S}_0\,\upsilon_1\upsilon_2)}{2\pi\mathcal{S}_0\,\upsilon_1\upsilon_2} \tag{4.49}$$

$$= \frac{\sin(2\pi\mathcal{S}_0\,\upsilon_1\upsilon_2)}{2\pi\mathcal{S}_0\,\upsilon_1\upsilon_2} + \mathrm{j}\frac{\sin^2(\pi\mathcal{S}_0\,\upsilon_1\upsilon_2)}{\pi\mathcal{S}_0\,\upsilon_1\upsilon_2}, \tag{4.50}$$

where we use the definition $\mathrm{sinc}(x) = \mathrm{si}(\pi x) = \sin(\pi x)/(\pi x)$.

The nonlinear transfer function can become a purely real-valued function if the exponential pre-factor in (4.45), (4.48) vanishes. This can be achieved if dispersion pre-compensation is taken into account with $\mathcal{B}_0 = -\bar{\beta}_2 \frac{L_{\mathrm{sp}}}{2} = -\pi \mathcal{S}_0$, i.e., the dispersion profile $\mathcal{B}(z)$ is *symmetrized* when the signal is pre-dispersed with half the accumulated net dispersion, cf. sqrt-sqrt equalization of linear dispersive channel [Fis02]. For (homogeneous) multi-span transmission the value is multiplied by the number of spans N_{sp} which is equivalent to the *straight-line rule*—an approximate rule to symmetrize the dispersion profile [FABH02, BSO08, BSB08].

In the context of soliton transmission, the nonlinear kernel $H_{\mathrm{NL,sp}}^{\alpha\to 0}(\upsilon_1, \upsilon_2)$ with a symmetrized dispersion profile has already been derived through the *nonlocal dispersion-managed NLSE* in early publications by Ablowitz and coworkers in [AB98, AHB01, AH02a]. An equivalent expression has been obtained by a similar approach—the *path-averaged propagation model*—by Turitsyn and coworkers [GT96, TTMF00, TFS+00].

Asymptotic Limits — Infinite Span-length Given a transmission link with a single, asymptotically long fiber (i.e., $L_{\mathrm{sp}} \to \infty$) or equivalently for $L_{\mathrm{sp}} \gg L_{\mathrm{eff,a}}$, the single-span nonlinear transfer function can be approximated to be independent of the length L_{sp}, and only dependent on the asymptotic length $L_{\mathrm{eff,a}}$, as defined in (3.30).

The nonlinear transfer function is then obtained as [LHP+05, (9)] [SBO07, (31)] [BSO08, (8)] [FBP09, (7)]

$$H_{\mathrm{NL,sp}}^{L_{\mathrm{sp}}\to\infty}(\upsilon_1, \upsilon_2) \stackrel{\mathrm{def}}{=} \lim_{L_{\mathrm{sp}}\to\infty} H_{\mathrm{NL,sp}}(\upsilon_1, \upsilon_2) = \frac{1}{L_{\mathrm{eff,a}}} \int_0^\infty \exp(-\bar{\alpha}\zeta + \mathrm{j}\upsilon_1\upsilon_2\bar{\beta}_2\zeta)\mathrm{d}\zeta \tag{4.51}$$

$$= \frac{1}{1 - \mathrm{j}\upsilon_1\upsilon_2\bar{\beta}_2 L_{\mathrm{eff,a}}}, \tag{4.52}$$

where we use $L_{\mathrm{eff,a}} \stackrel{\mathrm{def}}{=} \lim_{L_{\mathrm{sp}}\to\infty} L_{\mathrm{eff}} = 1/\bar{\alpha}$. We can again take the limit $L_{\mathrm{sp}} \to \infty$ for the map strength $\mathcal{S}$ and define the *asymptotic* map strength as

$$\mathcal{S}_{\mathrm{a}} \stackrel{\mathrm{def}}{=} \lim_{L_{\mathrm{sp}}\to\infty} \mathcal{S} = \bar{\beta}_2 L_{\mathrm{eff,a}}/(2\pi) \tag{4.53}$$

$$\mathcal{S}_{T,\mathrm{a}} \stackrel{\mathrm{def}}{=} (2\pi R_{\mathrm{s}})^2 \lim_{L_{\mathrm{sp}}\to\infty} \mathcal{S} = \mathrm{sign}(\bar{\beta}_2)\, L_{\mathrm{eff,a}}/L_{\mathrm{D}}, \tag{4.54}$$

to rewrite equation (4.52) as a function of the (asymptotic) map strength, i.e.,

$$H_{\mathrm{NL,sp}}^{L_{\mathrm{sp}}\to\infty}(\upsilon_1, \upsilon_2) = \frac{1}{1 - \mathrm{j}2\pi\upsilon_1\upsilon_2\mathcal{S}_{\mathrm{a}}}, \tag{4.55}$$

The expression in (4.55) is equivalent to a first-order low-pass w.r.t. ξ. It has a (two-sided) 3-dB bandwidth, i.e., a full-width half maximum, of $|2/(2\pi\mathcal{S}_{\mathrm{a}})|$ cf. [LHP+05, FBP09]. The width over the *normalized* frequency axis $\xi/(2\pi R_{\mathrm{s}})^2$ is $|2/(2\pi\mathcal{S}_{T,\mathrm{a}})|$, cf. Example 4.3.

Example 4.3: Asymptotic limits of the nonlinear transfer function

In Figure 4.9, the magnitude (in logarithmic scale) and the real- and imaginary part (in linear scale) of the single-span nonlinear transfer function $H_{\mathrm{NL,sp}}(\xi)$ are shown over $\xi/(2\pi R_{\mathrm{s}})^2$ exemplarily for $R_{\mathrm{s}} = 32\,\mathrm{GBd}$.

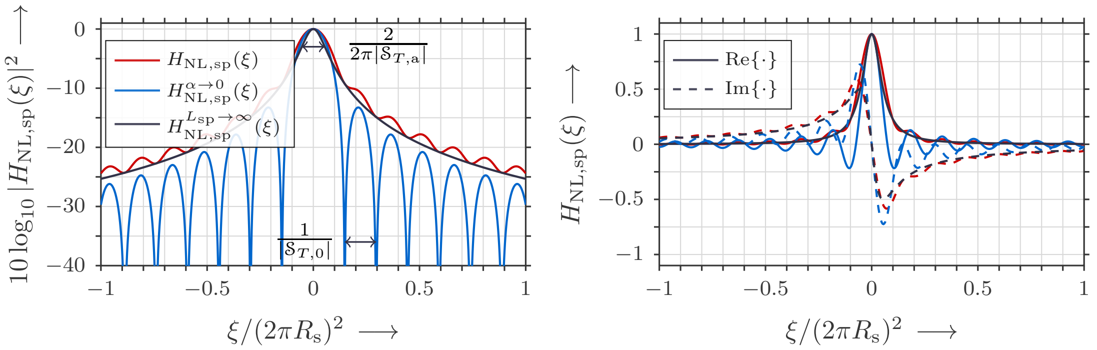


Figure 4.9: Squared-magnitude in logarithmic scale (left) and real- and imaginary-part in linear scale (right) of the *single-span* nonlinear transfer function for $\beta_2 = -21\,\mathrm{ps}^2/\mathrm{km}$, $\mathcal{B}_0 = 0\,\mathrm{ps}^2$, $10\log_{10}\mathrm{e}^{\bar{\alpha}} = 0.2\ \mathrm{dB/km}$, $R_{\mathrm{s}} = 32\,\mathrm{GBd}$, and $L_{\mathrm{sp}} = 50\ \mathrm{km}$ over $\xi/(2\pi R_{\mathrm{s}})^2 = \upsilon_1\upsilon_2/(2\pi R_{\mathrm{s}})^2$. The asymptotic limits of $H_{\mathrm{NL,sp}}(\xi)$ for $\alpha \to 0$ and $L_{\mathrm{sp}} \to \infty$ are also shown. Here, $L_{\mathrm{D}} = 7.4\,\mathrm{km}$ and $L_{\mathrm{eff,a}} = 21.71\,\mathrm{km}$.

We find the following length scales

$$L = L_{\mathrm{sp}} = 50\ \mathrm{km}\,, \qquad L_{\mathrm{D}} = 7.40\ \mathrm{km}\,, \qquad L_{\mathrm{eff}} = 19.54\ \mathrm{km} \qquad L_{\mathrm{eff,a}} = 21.71\ \mathrm{km}\,,$$

given $10\log_{10}\mathrm{e}^{\bar{\alpha}} = 0.2\,\mathrm{dB/km}$ and $\bar{\beta}_2 = -21\ \mathrm{ps}^2/\mathrm{km}$.
Additionally, the asymptotic limits of the nonlinear transfer function $H^{\alpha\to 0}_{\mathrm{NL,sp}}(\xi)$ and $H^{L_{\mathrm{sp}}\to\infty}_{\mathrm{NL,sp}}(\xi)$ are also shown. The related map strengths compute to

$$\mathcal{S}_{T,\rho} = -2.64\,, \qquad \mathcal{S}_{T,0} = -6.76 = -\frac{1}{0.148}\,, \qquad \mathcal{S}_{T,\mathrm{a}} = -2.93 \approx -\frac{1}{\pi\cdot 0.10}\,.$$

For this set of system parameters an interesting feature of the nonlinear transfer function can be observed. It can be seen that $H_{\mathrm{NL,sp}}(\xi)$ is not monotonically decreasing with ξ, i.e., there are values of $\xi = \upsilon_1\upsilon_2$ where the nonlinear coupling accumulates *worse* even though the phase matching proportional to ξ in (4.39) is *better*. In particular, those dips occur at integer multiples of

$$\frac{1}{|\mathcal{S}_{T,0}|} = \frac{L_{\mathrm{D}}}{L_{\mathrm{sp}}} = 0.148\,.$$

Those dips are identical with the *zeros* of the corresponding sinc-shaped transfer function $H^{\alpha\to 0}_{\mathrm{NL,sp}}(\xi)$. As a consequence, in spite of lossless transmission (i.e., strong nonlinear interaction), there are combinations of υ_1 and υ_2 for which the nonlinear distortion accumulates to *zero* over the span length L_{sp}, e.g.,

$$\frac{(2\pi\cdot 8\ \mathrm{GHz})(2\pi\cdot 18.94\,\mathrm{GHz})}{(2\pi R_{\mathrm{s}})^2} = 0.148, \quad \text{or} \quad \frac{(2\pi\cdot 12.31\ \mathrm{GHz})(2\pi\cdot 12.31\,\mathrm{GHz})}{(2\pi R_{\mathrm{s}})^2} = 0.148\,.$$

On the other hand, the single-span nonlinear transfer function $H^{L_{\mathrm{sp}}\to\infty}_{\mathrm{NL,sp}}(\xi)$ has the characteristics of a first-order low-pass and the full-width half maximum, given in normalized units

$\xi/(2\pi R_\mathrm{s})^2$, computes to

$$\frac{1}{\pi|\mathcal{S}_{T,\mathrm{a}}|} = \frac{2L_\mathrm{D}}{2\pi L_\mathrm{eff,a}} = 0.10\,.$$

In Figure 4.10, the corresponding Nyquist plot of the nonlinear transfer function $H_\mathrm{NL,sp}(\xi)$ is shown. Here, the gray curve shows the scatter plot of $H_\mathrm{NL,sp}(\xi)$ in Cartesian coordinates (i.e., $z = L$ is fixed, while $\xi = \upsilon_1\upsilon_2$ is varied), while the color-coded traces show the accumulation of the nonlinear perturbation over ζ at a particular point $\xi = \upsilon_1\upsilon_2$ (i.e., ξ is fixed, and z is varied form 0 to L). The colored bullet markers indicate the point $z = L_\mathrm{sp}$ and all lie on the gray curve. E.g., the zero-crossings of $H^{\alpha\to 0}_\mathrm{NL,sp}(\xi)$ are clearly visible—the first one occurring at around $\xi_0 = (2\pi \cdot 12.31\,\mathrm{GHz})^2$, see above. The scatter plot of $H^{L_\mathrm{sp}\to\infty}_\mathrm{NL,sp}(\xi)$ forms a perfect circle in the Cartesian space.

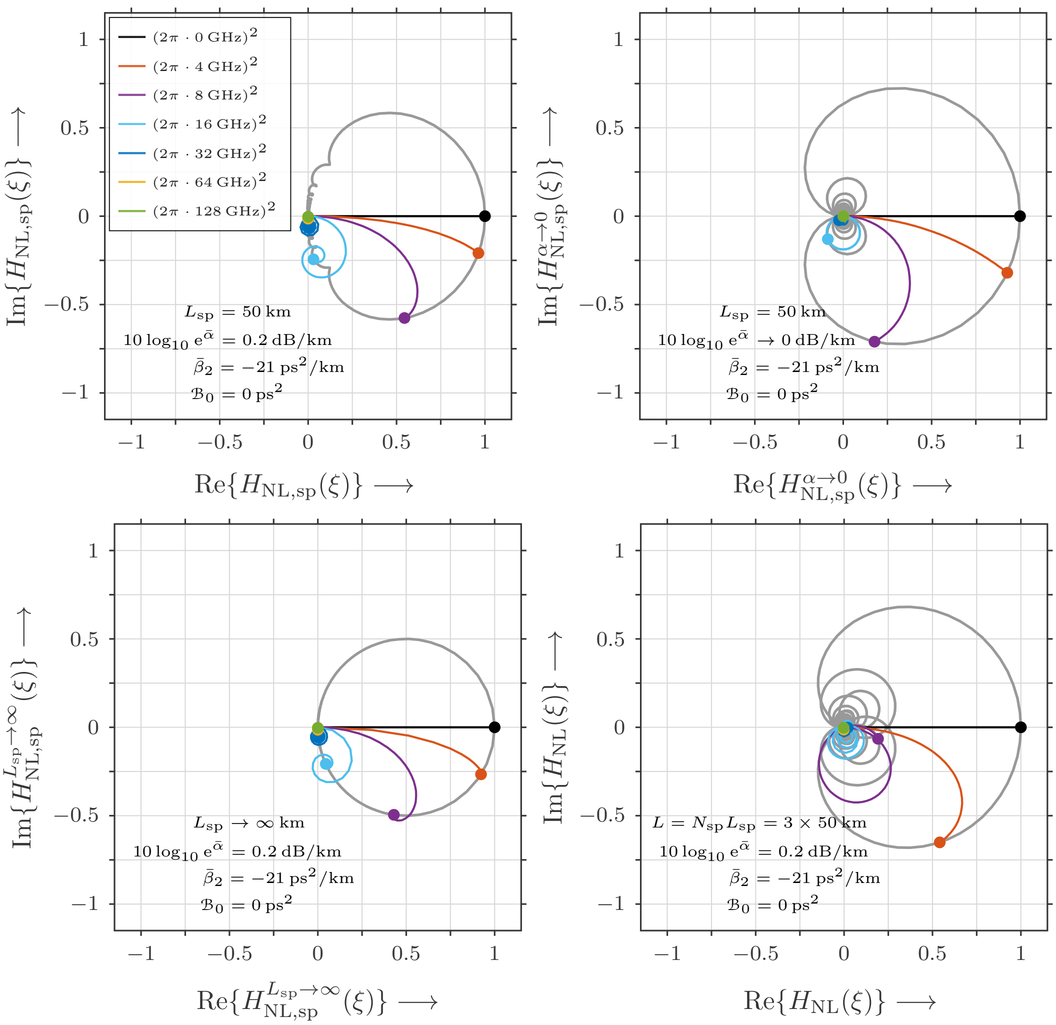


Figure 4.10: *Nyquist plot* or *locus curve* of $H_\mathrm{NL,sp}(\xi)$ and its asymptotic limits ($\alpha \to 0$ and $L_\mathrm{sp} \to \infty$) with the system parameters as in Example 4.3 and Figure 4.9. Additionally, the case for a homogeneous multi-span system with $N_\mathrm{sp} = 3$ is shown, also see Figure 4.11. The grey curve shows the nonlinear transfer function in Cartesian coordinates, while the color-coded traces show how the value of $H_\mathrm{NL}(\xi)$ at a particular point $\xi = \upsilon_1\upsilon_2$ accumulates via the integral over ζ in (4.42). The bullet markers indicate the location $z = N_\mathrm{sp}L_\mathrm{sp} = L$.

Homogeneous Multi-Span Solution In accordance with the parallel channel model, see Figure 4.2, the perturbation generated at one local position ζ_0 is independent (and hence uncorrelated) of the perturbation generated at any other position $\zeta \neq \zeta_0$. In the framework of the first-order RP method, this assumption clearly extends to perturbations generated in different spans.

Under the assumption that all spans are identical (i.e., have the same dispersion profile $\mathcal{B}(z)$, power profile $\mathcal{G}(z)$, and have the same length L_{sp}), the perturbation generated in different spans differs only in the amount of accumulated dispersion, i.e., first the required accumulated dispersion for the linear source term $\boldsymbol{u}_{\text{LIN}}(z,t)$ to propagate to the respective span, and then for the local perturbation to propagate back to the input of the transmission link. This is expressed by an additional phase factor (accounting for the accumulated dispersion) added to the nonlinear transfer function of a single span.

This observation enables to factor the integral in (4.35) into a term representing the effects of a single span $H_{\text{NL,sp}}(\upsilon_1, \upsilon_2)$, and a second term representing the sum of all phase contributions. Again, assuming only path-average span parameters as in (4.42), we express the nonlinear transfer function of the whole link as the product of the single-span transfer function and the sum of all phasors

$$H_{\text{NL}}(\upsilon_1, \upsilon_2) = H_{\text{NL,sp}}(\upsilon_1, \upsilon_2) \left(1 + \mathrm{e}^{\mathrm{j}\upsilon_1\upsilon_2\bar{\beta}_2 L_{\text{sp}}} + \mathrm{e}^{\mathrm{j}2\upsilon_1\upsilon_2\bar{\beta}_2 L_{\text{sp}}} + \ldots + \mathrm{e}^{\mathrm{j}(N_{\text{sp}}-1)\upsilon_1\upsilon_2\bar{\beta}_2 L_{\text{sp}}}\right), \tag{4.56}$$

where the sum of phasors is a truncated geometric series and can be given as closed-form analytic expression. This leads to the definition of the *phased-array factor* [LHP+05, Eqn. (6)] [PBC+12, Eqn. (92)] as

$$H_{\text{PAF}}(\upsilon_1, \upsilon_2) \stackrel{\text{def}}{=} \sum_{n=0}^{N_{\text{sp}}-1} \exp(\mathrm{j}\, n\, \upsilon_1\upsilon_2\bar{\beta}_2 L_{\text{sp}}) \tag{4.57}$$

$$= \frac{1 - \exp(\mathrm{j}\upsilon_1\upsilon_2\bar{\beta}_2 N_{\text{sp}} L_{\text{sp}})}{1 - \exp(\mathrm{j}\upsilon_1\upsilon_2\bar{\beta}_2 L_{\text{sp}})} \tag{4.58}$$

$$= \exp(\mathrm{j}\pi\upsilon_1\upsilon_2\mathcal{S}_0(N_{\text{sp}}-1)) \, \frac{\sin(\pi\upsilon_1\upsilon_2\,\mathcal{S}_0 N_{\text{sp}})}{\sin(\pi\upsilon_1\upsilon_2\,\mathcal{S}_0)}\,. \tag{4.59}$$

In Figure 4.11, we follow up on the Example 4.3, where $H_{\text{NL}}(\xi)$ and $H_{\text{PAF}}(\xi)$ are shown for $N_{\text{sp}} = 3$ and $N_{\text{sp}} = 4$ given the same single-span system parameters. The phase-array factor is a periodic function w.r.t. $\xi = \upsilon_1\upsilon_2$ with a period of $1/|\mathcal{S}_0|$ and has $N_{\text{sp}} - 1$ *zeros* (again w.r.t. ξ) over a single period. The number of sidelobes (within the considered frequency axis in Figure 4.11) increases linearly with the span number. At the same time, the spectral width of both the main- and the sidelobes is decreasing. The resulting steep slopes within the transfer function will impose severe constraints on the required frequency resolution if the multi-span transfer function should be *sampled* and numerical evaluated.

In Figure 4.10, the corresponding scatter plot for $N_{\text{sp}} = 3$ of the multi-span example is also shown. The high dynamic range of the multi-span transfer function is also visible in the scatter diagram.

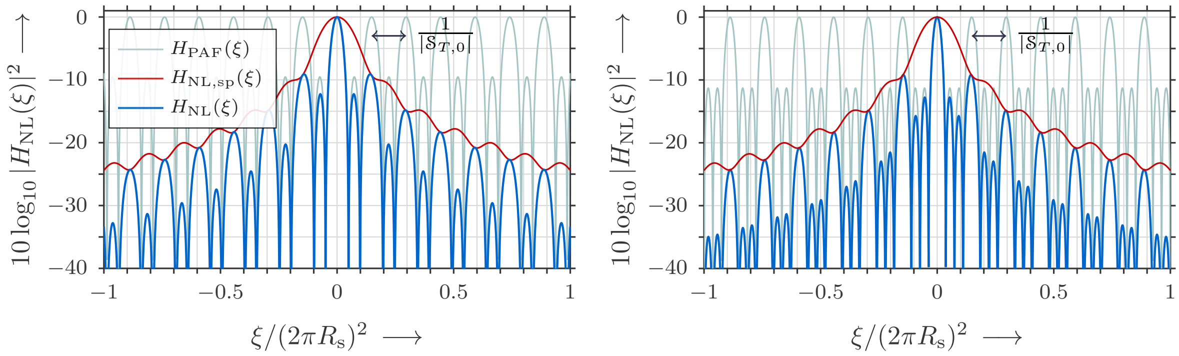


Figure 4.11: Magnitude in logarithmic scale of the *multi-span* nonlinear transfer function for $\beta_2 = -21\ \mathrm{ps}^2/\mathrm{km}$, $\mathcal{B}_0 = 0\ \mathrm{ps}^2$, $10\log_{10}\mathrm{e}^{\alpha} = 0.2\ \mathrm{dB/km}$, $R_\mathrm{s} = 32\,\mathrm{GBd}$, $L_\mathrm{sp} = 50\ \mathrm{km}$, $N_\mathrm{sp} = 3$, and $L = 150\ \mathrm{km}$ (left) and $N_\mathrm{sp} = 4$, and $L = 200\ \mathrm{km}$ (right) over $\xi/(2\pi R_\mathrm{s})^2$. Additionally, the single-span transfer function $H_{\mathrm{NL,sp}}(\xi)$ and the *phase array factor* $H_{\mathrm{PAF}}(\xi)$ are shown.

4.2 The First-Order Regular Solution in Time Domain

To derive the (analog) *baseband* end-to-end channel model and its time-domain equivalent, see Figure 4.1, the channel matched filter $H_\mathrm{C}^*(L,\omega)$ is subsequently applied to both $\boldsymbol{U}_{\mathrm{LIN}}(L,\omega)$ and $\Delta\boldsymbol{U}(L,\omega)$ from (4.27). Note, that the following equivalence holds

$$H_\mathrm{C}^*(L,\omega) \equiv H_\mathrm{CD}^*(L,\omega)\,, \tag{4.60}$$

due to the receive-side normalization with $\mathcal{G}(L) = 0$, see also (3.63) and (3.48).

The perturbation $\Delta\boldsymbol{S}(\omega)$, i.e., the perturbation in the *analog* domain following our terminology, is hence obtained by

$$\Delta\boldsymbol{S}(\omega) = H_\mathrm{C}^*(L,\omega)\,\Delta\boldsymbol{U}(L,\omega)\,, \tag{4.61}$$

which cancels out the leading term $H_\mathrm{CD}(L,\omega)$ in (4.27) since $|H_\mathrm{CD}(L,\omega)|^2 = |H_\mathrm{C}(L,\omega)|^2 = 1$. We find the expression of the electrical perturbation in both frequency and time domain as

$$\Delta\boldsymbol{S}(\omega) = -\mathrm{j}\bar{\gamma}\frac{8}{9}L_\mathrm{eff}\frac{1}{(2\pi)^2}\int_{\mathbb{R}^2}\boldsymbol{U}(0,\underbrace{\omega+\upsilon_2}_{\omega_3=\omega-\omega_1+\omega_2})\boldsymbol{U}^\mathsf{H}(0,\underbrace{\omega+\upsilon_1+\upsilon_2}_{\omega_2})\boldsymbol{U}(0,\underbrace{\omega+\upsilon_1}_{\omega_1})H_\mathrm{NL}(\upsilon_1,\upsilon_2)\ \mathrm{d}^2\boldsymbol{\upsilon} \tag{4.62}$$

$$\bullet\!\!-\!\!\circ$$

$$\Delta\boldsymbol{s}(t) = -\mathrm{j}\bar{\gamma}\frac{8}{9}L_\mathrm{eff}\int_{\mathbb{R}^2}\boldsymbol{u}(0,\underbrace{t-\tau_1}_{\mathrm{t}_1})\boldsymbol{u}^\mathsf{H}(0,\underbrace{t-\tau_1-\tau_2}_{\mathrm{t}_2})\boldsymbol{u}(0,\underbrace{t-\tau_2}_{\mathrm{t}_3=t-\mathrm{t}_1+\mathrm{t}_2})\,h_\mathrm{NL}(\tau_1,\tau_2)\ \mathrm{d}^2\boldsymbol{\tau}\,, \tag{4.63}$$

which recovers the time-domain representation of the third-order Volterra operator, see, e.g., (2.63)–(2.64). The Fourier transform $\Delta\boldsymbol{s}(t) \circ\!\!-\!\!\bullet \Delta\boldsymbol{S}(\omega)$ is carried out explicitly in Appendix A.2.

The frequency matching with $\omega_3 \overset{\mathrm{def}}{=} \omega - \omega_1 + \omega_2$ is translated to a *temporal matching*[4] $\mathrm{t}_3 \overset{\mathrm{def}}{=} t - \mathrm{t}_1 + \mathrm{t}_2$, i.e., the selection rules of FWM apply both in time and frequency, cf. [AH00].

[4]Not to be confused with the phase matching condition in (4.35), (4.39).

In symmetry with (4.20)–(4.22), we define the *absolute* and *relative* time variables as

$$\mathrm{t}_1 \overset{\text{def}}{=} t - \tau_1 \tag{4.64}$$

$$\mathrm{t}_2 \overset{\text{def}}{=} t - \tau_1 - \tau_2 \tag{4.65}$$

$$\mathrm{t}_3 \overset{\text{def}}{=} t - \mathrm{t}_1 + \mathrm{t}_2 = t - \tau_2, \tag{4.66}$$

see also Figure 4.3 (a).

The time-domain perturbation $\Delta\boldsymbol{s}(t)$ has the same form as its frequency-domain counterpart, i.e., the integrand is constituted by the respective time-domain representations of the optical signal at the input $\boldsymbol{u}(0,t)$ weighted with the time-domain kernel $h_{\mathrm{NL}}(\tau_1,\tau_2)$, which we will call the *nonlinear impulse response* of the continuous-time end-to-end channel, see Section 4.2.

In contrast to the *standard* form of the third-order Volterra operator in (2.48), the integral in (4.63) is only two-fold and the time-domain kernel has only two degrees of freedom $\boldsymbol{\tau} = [\tau_1,\tau_2]^\mathsf{T}$ instead of three (cf. [AH02a, Wei06]). This is induced by the temporal matching, see above, alternatively expressed as $\tau_3 = \tau_1 + \tau_2$ using the notation of the Volterra theory. In the Appendix A.3, we present an alternative derivation of the time-domain perturbation $\Delta\boldsymbol{s}(t)$ based on the time-domain Volterra theory. Therein, we show how the temporal matching constraint causes the three-fold integral from the Volterra ansatz to collapse into a two-fold integral, and at the same time reduces the dimension of the kernel's domain $\mathrm{dom}(h_{\mathrm{NL}})$ from three to two. Importantly, the time- and frequency-domain kernel expressed in *relative* time and frequency are now related via a 2D Fourier transform [BSO08, Eq. (6)]

$$H_{\mathrm{NL}}(\boldsymbol{\upsilon}) = \mathcal{F}_{\boldsymbol{\upsilon}\leftrightarrow\boldsymbol{\tau}}\{\, h_{\mathrm{NL}}(\boldsymbol{\tau})\,\}\,, \tag{4.67}$$

see also Appendix A.2 for the full proof.

The Nonlinear Impulse Response The time-domain kernel $h_{\mathrm{NL}}(\tau_1,\tau_2)$ can be given by, e.g., carrying out the inverse 2D Fourier transform (cf. [AH02a, Appx.] and [BSO08, Eq. (6)]) or by following the time-domain derivation in Appendix A.3. It is explicitly given as

$$h_{\mathrm{NL}}(\boldsymbol{\tau}) = \mathcal{F}^{-1}_{\boldsymbol{\upsilon}\leftrightarrow\boldsymbol{\tau}}\{\, H_{\mathrm{NL}}(\boldsymbol{\upsilon})\,\} = \frac{1}{L_{\mathrm{eff}}}\int_0^L \frac{1}{2\pi|\mathcal{B}(\zeta)|}\exp\left(\mathcal{G}(\zeta) - \mathrm{j}\frac{\tau_1\tau_2}{\mathcal{B}(\zeta)}\right)\mathrm{d}\zeta\,, \tag{4.68}$$

with the tuples $\boldsymbol{\tau} = [\tau_1,\tau_2]^\mathsf{T}$ and $\boldsymbol{\upsilon} = [\upsilon_1,\upsilon_2]^\mathsf{T}$. The time-domain kernel maintains its hyperbolic form and symmetry as it is a function of the product $\tau_1\tau_2$. Figure 4.12 shows the corresponding single-span time-domain kernel $h_{\mathrm{NL,sp}}(\tau_1,\tau_2)$ from Example 4.1. Here, the kernel itself has a singularity at $\tau_1\tau_2 = 0$ due to the singularity of the integrand at $\zeta = 0$.

Similar as in (4.32), we can also express $h_{\mathrm{NL}}(\boldsymbol{\tau})$ in terms of *absolute time variables*, i.e.,

$$h_{\mathrm{NL}}(\mathbf{t}) = h_{\mathrm{NL}}(\mathrm{t}_1,\mathrm{t}_2,\mathrm{t}_3) = \frac{1}{L_{\mathrm{eff}}}\int_0^L \frac{1}{2\pi|\mathcal{B}(\zeta)|}\exp\left(\mathcal{G}(\zeta) - \mathrm{j}\frac{(\mathrm{t}_3-\mathrm{t}_2)(\mathrm{t}_1-\mathrm{t}_2)}{\mathcal{B}(\zeta)}\right)\mathrm{d}\zeta \tag{4.69}$$

$$= \frac{1}{L_{\mathrm{eff}}}\int_0^L h_{\mathrm{C}}(\zeta,\mathrm{t}_1)h^*_{\mathrm{C}}(\zeta,\mathrm{t}_2)h_{\mathrm{C}}(\zeta,\mathrm{t}_3)h^{-1}_{\mathrm{C}}(\zeta,\underbrace{\mathrm{t}_1-\mathrm{t}_2+\mathrm{t}_3}_{t})\,\mathrm{d}\zeta\,. \tag{4.70}$$

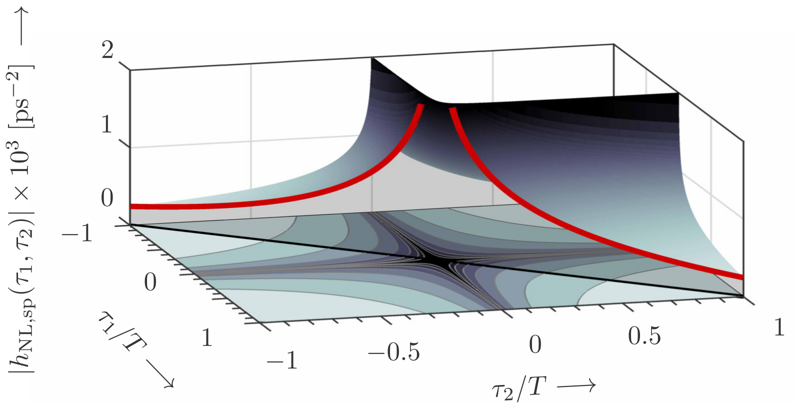


Figure 4.12: Magnitude in linear scale of the *single-span* nonlinear impulse response for $\bar{\beta}_2 = -21\ \mathrm{ps}^2/\mathrm{km}$, $\mathcal{B}_0 = 0\ \mathrm{ps}^2$, $10\log_{10} \mathrm{e}^{\bar{\alpha}} = 0.2\ \mathrm{dB/km}$ and $L = L_{\mathrm{sp}} = 100\ \mathrm{km}$ over the time *differences* τ_1 and τ_2 normalized to $R_\mathrm{s} = 64$ GBd. The red line denotes $h_{\mathrm{NL,sp}}(\psi)$ which only depends on the scalar $\psi = \tau_1\tau_2$. (Part for $\tau_1 > \tau_2$ not shown). The time-domain kernel $h_{\mathrm{NL,sp}}(\tau_1,\tau_2)$ is also the inverse 2D Fourier transform of the kernel $H_{\mathrm{NL,sp}}(\upsilon_1,\upsilon_2)$ in frequency domain from Figure 4.7.

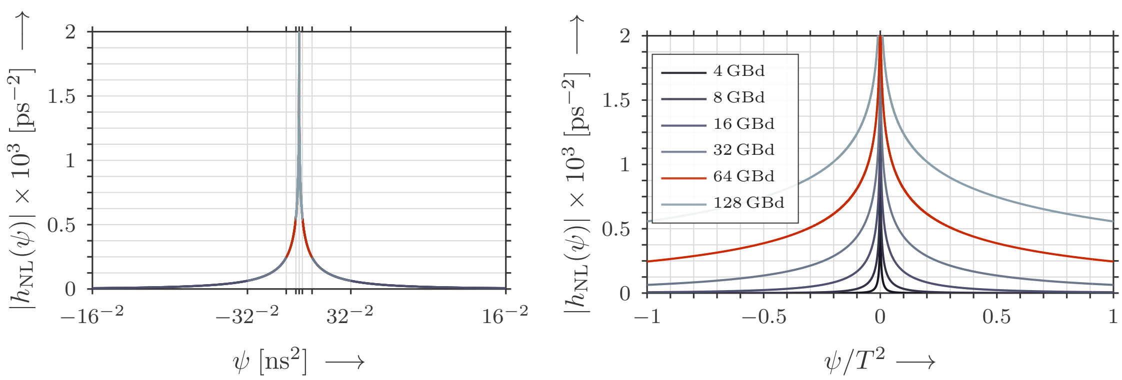


Figure 4.13: Squared-magnitude in linear scale of the *single-span* nonlinear impulse response $h_{\mathrm{NL}}(\psi)$ for $\bar{\beta}_2 = -21\ \mathrm{ps}^2/\mathrm{km}$, $\mathcal{B}_0 = 0\ \mathrm{ps}^2$, $10\log_{10} \mathrm{e}^{\bar{\alpha}} = 0.2\ \mathrm{dB/km}$ and $L_{\mathrm{sp}} = 100\ \mathrm{km}$ over $\psi = \tau_1\tau_2$. The normalization with $R_\mathrm{s}^2 = 1/T^2$ relates $h_{\mathrm{NL}}(\psi)$ to the probe's temporal width, i.e., symbol duration.

where we recover[5] the explicit dependency on the linear channel impulse response $h_\mathrm{C}(z,t)$. Here we use again the equivalence from (4.26) applied to the time variables

$$\tau_1\tau_2 = (t-\mathrm{t}_1)(\mathrm{t}_1-\mathrm{t}_2) = (\mathrm{t}_3-\mathrm{t}_2)(t-\mathrm{t}_3) = \frac{1}{2}\big(t^2-\mathrm{t}_1^2+\mathrm{t}_2^2-(\underbrace{t-\mathrm{t}_1+\mathrm{t}_2}_{\mathrm{t}_3})^2\big)\,. \tag{4.71}$$

Note the duality to (4.32), where in both representations the nonlinear transfer function can be understood as the *path-average* (cf. [GT96]) over an expression related to the linear channel response $h_\mathrm{C}(z,t) \circ\!\!-\!\!\bullet\, H_\mathrm{C}(z,\omega)$. Other than the general basic third-order system in (2.65)–(2.66), the inner temporal convolution is not explicitly carried out due to the temporal matching constraint $\mathrm{t}_3 = t - \mathrm{t}_1 + \mathrm{t}_2$, see Appendix A.3 for details.

In Figure 4.13, we show the time-domain kernel using the same parameters as in Example 4.2 over the scalar variable

$$\psi \overset{\mathrm{def}}{=} \tau_1\tau_2\,, \tag{4.72}$$

[5]We used the simple relations $\sqrt{\mathrm{j}}\,\sqrt{\mathrm{j}}^* = 1$ and $\sqrt{\mathcal{B}(\zeta)}\,\sqrt{\mathcal{B}(\zeta)}^* = |\mathcal{B}(\zeta)|$.

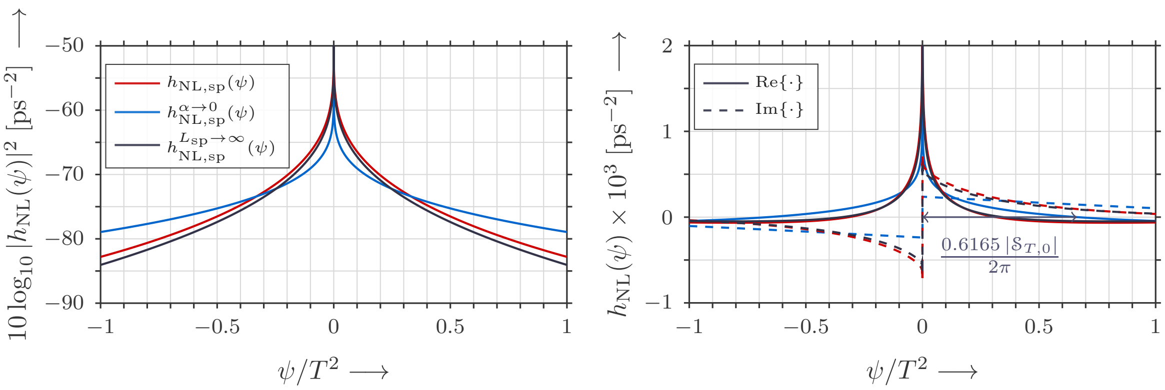


Figure 4.14: Squared-magnitude in logarithmic scale (left) and real- and imaginary-part in linear scale (right) of the *single-span* nonlinear impulse response $h_{\mathrm{NL,sp}}(\psi)$ for $\bar{\beta}_2 = -21\ \mathrm{ps}^2/\mathrm{km}$, $\mathcal{B}_0 = 0\ \mathrm{ps}^2$, $10\log_{10}\mathrm{e}^{\bar{\alpha}} = 0.2\ \mathrm{dB/km}$, $R_\mathrm{s} = 32$ GBd, and $L_{\mathrm{sp}} = 50$ km over $\psi/T^2 = \tau_1\tau_2/T^2$. The asymptotic limits of $h_{\mathrm{NL,sp}}(\psi)$ for $\alpha \to 0$ and $L_{\mathrm{sp}} \to \infty$ are also shown. See also the corresponding nonlinear transfer function $H_{\mathrm{NL,sp}}(\xi)$ in Figure 4.9.

with and without normalization to the symbol rate $R_\mathrm{s} = 1/T$. The red traces (i.e., $R_\mathrm{s} = 64$ GBd) in Figure 4.12 and Figure 4.13 are identical. The nonlinear memory (i.e., the temporal width of the nonlinear impulse response relative to the symbol rate of the probe) is proportional to the map strength $\mathcal{S}_{T,\rho} \propto R_\mathrm{s}^2$, i.e., doubling R_s *increases* the temporal width of $|h_{\mathrm{NL}}(\psi/T^2)|$ by a factor of 4, see Figure 4.13 (right).

Asymptotic Limits — Lossless Transmission We now consider the asymptotic limit of the single-span time-domain kernel $h_{\mathrm{NL,sp}}(\boldsymbol{\tau})$ for lossless transmission, i.e., $\alpha(z) \to 0$. We assume again only path-average span parameters and follow from (4.68) that [AH02a]

$$h^{\alpha\to 0}_{\mathrm{NL,sp}}(\tau_1,\tau_2) = \frac{1}{2\pi|\bar{\beta}_2|L_{\mathrm{sp}}}\int_0^{L_{\mathrm{sp}}} \frac{1}{\zeta}\exp\left(-\mathrm{j}\frac{\tau_1\tau_2}{\bar{\beta}_2\zeta}\right)\,\mathrm{d}\zeta\,, \tag{4.73}$$

which can be given by the analytic expression [AH02a]

$$\begin{aligned} h^{\alpha\to 0}_{\mathrm{NL,sp}}(\tau_1,\tau_2) &= \frac{1}{2\pi|\bar{\beta}_2|L_{\mathrm{sp}}}\mathrm{E}_1\left(-\mathrm{j}\frac{\tau_1\tau_2}{\bar{\beta}_2 L_{\mathrm{sp}}}\right) & (4.74)\\ &= \frac{1}{(2\pi)^2|\mathcal{S}_0|}\mathrm{E}_1\left(-\mathrm{j}\frac{\tau_1\tau_2}{2\pi\,\mathcal{S}_0}\right) & (4.75)\\ &= \frac{1}{2\pi|\bar{\beta}_2|L_{\mathrm{sp}}}\left(\mathrm{Ci}\left(-\frac{\tau_1\tau_2}{\bar{\beta}_2 L_{\mathrm{sp}}}\right) + \mathrm{j}\mathrm{Si}\left(-\frac{\tau_1\tau_2}{\bar{\beta}_2 L_{\mathrm{sp}}}\right) - \mathrm{j}\frac{\pi}{2}\right), & (4.76)\end{aligned}$$

where $\mathrm{E}_1(\cdot)$ is the *exponential integral function*[6], and $\mathrm{Ci}(\cdot)$ and $\mathrm{Si}(\cdot)$ are the *cosine* and *sine integral functions* (cf. also *Nielsen spiral*).

Figure 4.14 follows up on Example 4.3 and shows the single-span nonlinear impulse response including its asymptotic limits. Additionally, in Figure 4.15, the related Nyquist plots are shown. E.g., the first *zero* of $\mathrm{Ci}(\cdot)$ is at 0.6165, and the first *zero* of $\mathrm{Si}(\cdot) - \pi/2$ is at 1.9264. Those zero-crossing can also be seen in Figure 4.14 and 4.15, where, e.g., for the kernel

[6] Here, we assume $\mathrm{E}_1(\mathrm{j}x)$ with $x > 0$, i.e., in the anomalous dispersion regime with $\bar{\beta}_2 < 0$ and $\tau_1\tau_2 > 0$.

$h_{\mathrm{NL,sp}}^{\alpha\to 0}(\psi)$, the first zero-crossing of the real part occurs at $0.6165\,|\mathcal{S}_{T,0}|/(2\pi)$—again showing how the temporal width of the nonlinear impulse response scales with the map strength $\mathcal{S}_T$. The authors of [AB98] and [AHB01] show that, by symmetrizing the dispersion profile with $\mathcal{B}_0 = -\bar{\beta}_2 \frac{L_{\mathrm{sp}}}{2} = -\pi\mathcal{S}_0$, both frequency- and time-domain kernels become real-valued functions (up to a constant term) with

$$H_{\mathrm{NL,sym}}^{\alpha\to 0}(\upsilon_1, \upsilon_2) = \mathrm{sinc}\left(\upsilon_1\upsilon_2\, 2\,\mathcal{S}_0\right) \tag{4.77}$$

$$h_{\mathrm{NL,sym}}^{\alpha\to 0}(\tau_1, \tau_2) = \frac{1}{(2\pi)^2|\mathcal{S}_0|}\left(\mathrm{Ci}\left(-\frac{\tau_1\tau_2}{2\pi\mathcal{S}_0}\right) - \mathrm{j}\frac{\pi}{2}\right)\,, \tag{4.78}$$

where the 2D Fourier transform is performed w.r.t. $\boldsymbol{\tau} \leftrightarrow \boldsymbol{\upsilon}$.

Asymptotic Limits — Infinite Span-length The single-span nonlinear transfer function in the limit $L_{\mathrm{sp}} \to \infty$ can only be given by the integral expression

$$h_{\mathrm{NL,sp}}^{L_{\mathrm{sp}}\to\infty}(\tau_1, \tau_2) = \frac{1}{2\pi|\bar{\beta}_2|L_{\mathrm{eff,a}}}\int_0^\infty \frac{1}{\zeta}\exp\left(-\bar{\alpha}\zeta - \mathrm{j}\frac{\tau_1\tau_2}{\bar{\beta}_2\zeta}\right)\mathrm{d}\zeta\,, \tag{4.79}$$

which is solved numerically for the examples in Figure 4.14 and Figure 4.15.

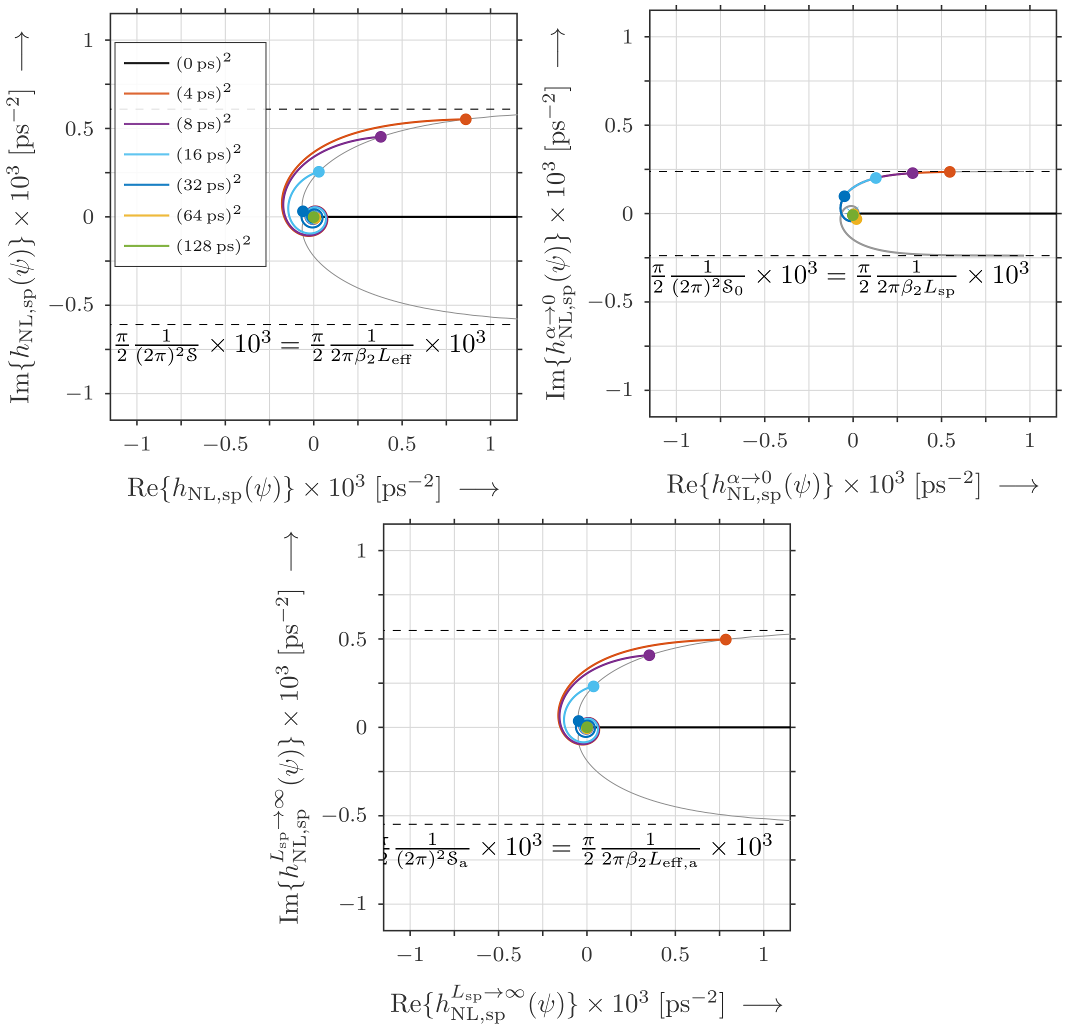


Figure 4.15: *Nyquist plot* or *locus curve* of $h_{\mathrm{NL,sp}}(\psi)$ and its asymptotic limits ($\alpha \to 0$ and $L_{\mathrm{sp}} \to \infty$) with the system parameters as in Example 4.3 and Figure 4.10. The grey curve shows the nonlinear impulse response in Cartesian coordinates, while the color-coded traces show how the value of $h_{\mathrm{NL,sp}}(\psi)$ at a particular point $\psi = \tau_1\tau_2$ accumulates via the integral over ζ in (4.68). The bullet markers indicate the location $\zeta = L_{\mathrm{sp}}$.

4.3 The (Analog) Baseband End-to-End Channel

We now continue to express the perturbation $\Delta\boldsymbol{s}(t) \circ\!\!-\!\!\bullet \Delta\boldsymbol{S}(\omega)$ as a function of the electrical baseband transmit signals $\boldsymbol{s}_\nu(t) \circ\!\!-\!\!\bullet \boldsymbol{S}_\nu(\omega)$. The general idea is to expand the optical field envelope of the WDM signal $\boldsymbol{u}(0,t) \circ\!\!-\!\!\bullet \boldsymbol{U}(0,\omega)$, similarly as in (3.77), and in turn, shift the nonlinear transfer function by the respective channel separation $\Delta\omega_\nu$ to align with the baseband signals, see below.

We are again interested in the perturbation that is imposed on the probe channel, i.e., the frequency-domain expression in (4.62) is evaluated for all $\omega \in \mathcal{B}_\rho$ in (4.62). Similar as in (3.79), we will dissect the total perturbation $\Delta\boldsymbol{s}(t) \circ\!\!-\!\!\bullet \Delta\boldsymbol{S}(\omega)$ into contributions originating from SCI and XCI, whereas MCI will be neglected.

The reason for the latter is as follows. We notice from Figure 4.8 and Figure 4.11 that, given R_s (and consequently the channel spacing between different wavelength channels) is sufficiently large, the *FWM efficiency* $|H_\mathrm{NL}(\xi/(2\pi R_\mathrm{s})^2)|^2$ may become negligibly small for $\xi \gg (2\pi R_\mathrm{s})^2$. In other words, the *phase matching condition* in (4.39), proportional to $\xi = \upsilon_1\upsilon_2$, is not properly met if both υ_1 and υ_2 are larger than $2\pi R_\mathrm{s}$. Conversely, if the nonlinear transfer function (normalized to the probe's symbol rate) is narrow-band, then we can factor the integrand in (4.62), (4.63) into an SCI and XCI term, i.e., mixing terms that originate either from within the probe channel (both $\upsilon_1 < 2\pi R_\mathrm{s}$ and $\upsilon_2 < 2\pi R_\mathrm{s}$) or from within the probe channel and a single interfering wavelength channel (either $\upsilon_1 < 2\pi R_\mathrm{s}$ or $\upsilon_2 < 2\pi R_\mathrm{s}$). Mixing terms originating from MCI are only relevant if R_s (and hence the channel separation $\Delta\omega_\nu$) or the dispersion coefficient $\beta_2(z)$ is small[7]. We hence neglect any FWM terms involving more than two wavelength channels.

Considering only mixing terms falling into the spectral support of the probe channel $\omega \in \mathcal{B}_\rho$, we can expand the triple product in (4.62) as

$$\begin{aligned}\boldsymbol{U}(0,\underbrace{\omega_3}_{\omega-\omega_1+\omega_2})\boldsymbol{U}^\mathsf{H}(0,\omega_2)\boldsymbol{U}(0,\omega_1)\Big|_{\omega\in\mathcal{B}_\rho} &= \underbrace{\boldsymbol{U}_\rho(0,\omega_3)\boldsymbol{U}_\rho^\mathsf{H}(0,\omega_2)\boldsymbol{U}_\rho(0,\omega_1)}_{\text{SCI}} \\ &+ \sum_{\nu\neq\rho}\underbrace{\Big(\boldsymbol{U}_\nu(0,\omega_3)\boldsymbol{U}_\nu^\mathsf{H}(0,\omega_2)\boldsymbol{U}_\rho(0,\omega_1) + \boldsymbol{U}_\rho(0,\omega_3)\boldsymbol{U}_\nu^\mathsf{H}(0,\omega_2)\boldsymbol{U}_\nu(0,\omega_1)\Big)}_{\text{XCI}}, \quad (4.80)\end{aligned}$$

where the XCI term has two contributions—the first results from an interaction where ω_3 and ω_2 are from the ν^th interfering wavelength channel and ω and ω_1 are within the probe's support (see lower blue region in Figure 4.16 (left)). The second involves an interaction where ω_2 and ω_1 are from the interfering wavelength channel, and ω, ω_3 are from the probe channel (see blue region on the right in Figure 4.16 (left)).

We can exploit the symmetry[8] of the nonlinear transfer function $H_\mathrm{NL}(\upsilon_1,\upsilon_2) = H_\mathrm{NL}(\upsilon_2,\upsilon_1)$ to simplify the XCI expression in (4.80). We obtain the integrand of the Volterra operator in

[7]E.g., for a channel spacing smaller than 25 GHz and typical $\bar{\beta}_2 = -21\ \mathrm{ps}^2/\mathrm{km}$ [FBP09, Fis09].

[8]Note, that the symmetry of the nonlinear transfer function is broken if dispersion slope $\beta_3(z) \neq 0$ is considered [AFAK+18].

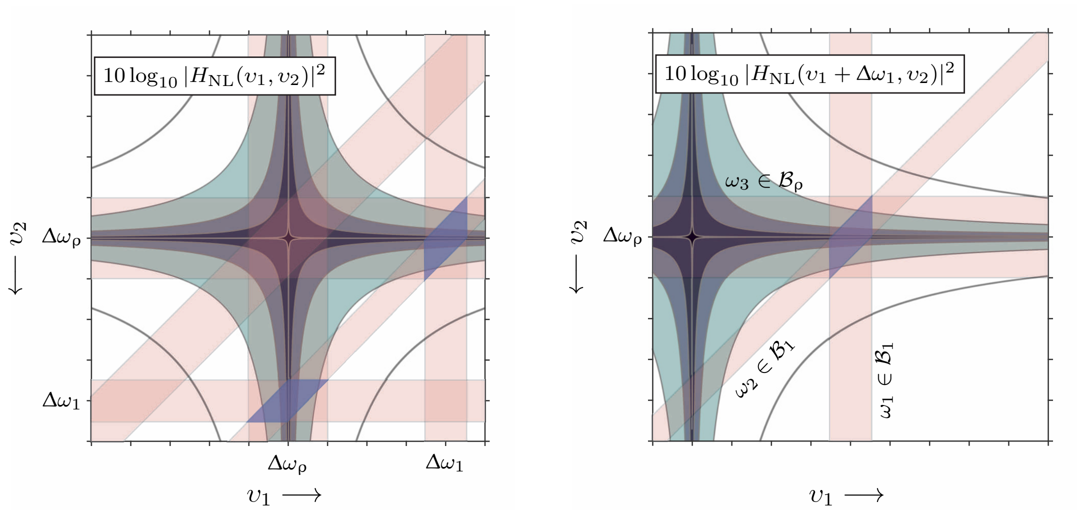


Figure 4.16: Squared-magnitude of the un-shifted (left) and shifted (right) nonlinear transfer function $H_{\mathrm{NL}}(\upsilon_1,\upsilon_2)$ in logarithmic scale. On the left side, the shaded vertical, horizontal, and diagonal regions in red show the spectral support of $\boldsymbol{U}(0,\omega_1)$, $\boldsymbol{U}(0,\omega_3)$, and $\boldsymbol{U}(0,\omega_2)$ for $\omega=0$ (i.e., $\omega_2=\upsilon_1+\upsilon_2$) given two wavelength channels at $\Delta\omega_\rho$ and $\Delta\omega_1$ with bandwidth B_ρ and $B_1=0.5B_\rho$. The intersections of those regions (blue areas) are the integration domains that contribute to the XCI terms in (4.80). On the right side, the nonlinear transfer function is aligned with the baseband signals. The shaded regions in red show the spectral support of $\boldsymbol{S}_\nu(\omega_1)$, $\boldsymbol{S}_\rho(\omega_3)$, and $\boldsymbol{S}_\nu(\omega_2)$ for $\omega=0$. Only a single blue shaded area remains and must be considered as XCI contribution in (4.81).

(4.62) using the definition of the baseband signal of each wavelength channel[9]

$$\begin{aligned}
&\boldsymbol{U}(0,\,\omega_3)\boldsymbol{U}^{\mathsf{H}}(0,\omega_2)\boldsymbol{U}(0,\omega_1)H_{\mathrm{NL}}(\omega_2-\omega_3,\omega_2-\omega_1)\Big|_{\omega\in\mathcal{B}_\rho} \\
&= \boldsymbol{S}_\rho(\omega_1)\boldsymbol{S}_\rho^{\mathsf{H}}(\omega_2)\boldsymbol{S}_\rho(\omega_3)H_{\mathrm{NL}}(\omega_2-\omega_1,\omega_2-\omega_3) \\
&\quad+\sum_{\nu\neq\rho}\Big(\boldsymbol{S}_\nu(\omega_1)\boldsymbol{S}_\nu^{\mathsf{H}}(\omega_2)+\boldsymbol{S}_\nu^{\mathsf{H}}(\omega_2)\boldsymbol{S}_\nu(\omega_1)\mathbf{I}\Big)\,\boldsymbol{S}_\rho(\omega_3)H_{\mathrm{NL}}(\underbrace{\omega_2-\omega_3}_{\upsilon_1}+\Delta\omega_\nu,\underbrace{\omega_2-\omega_1}_{\upsilon_2})\,,
\end{aligned} \tag{4.81}$$

which now corresponds to the case that ω_3 always lays in the support of the probe[10], i.e., $\omega_3\in\mathcal{B}_\rho$. The signals of the interfering wavelength channels are now represented in their respective ECB, see (3.15)–(3.16). In turn, the nonlinear transfer function $H_{\mathrm{NL}}(\upsilon_1+\Delta\omega_\nu,\upsilon_2)$ is now shifted according to the relative frequency offset $\Delta\omega_\nu$ to align properly with the baseband signal of the interferer, see Figure 4.16 (right).

The shifted argument of the nonlinear transfer function (given in terms of relative frequencies $[\upsilon_1,\upsilon_2]^{\mathsf{T}}$) can be rewritten according to (4.26) as

$$(\upsilon_1+\Delta\omega_\nu)\upsilon_2=\frac{1}{2}\big(\underbrace{\omega^2}_{\in\mathcal{B}_\rho}-(\underbrace{\omega_1+\Delta\omega_\nu}_{\in\mathcal{B}_\nu})^2+(\underbrace{\omega_2+\Delta\omega_\nu}_{\in\mathcal{B}_\nu})^2-(\underbrace{\omega-(\omega_1-\Delta\omega_\nu)+(\omega_2+\Delta\omega_\nu)}_{\omega_3\in\mathcal{B}_\rho})^2\big)\,, \tag{4.82}$$

[9]Since $\boldsymbol{U}_\nu^{\mathsf{H}}\boldsymbol{U}_\nu$ is a scalar, we have $\boldsymbol{U}_\rho\boldsymbol{U}_\nu^{\mathsf{H}}\boldsymbol{U}_\nu=\boldsymbol{U}_\nu^{\mathsf{H}}\boldsymbol{U}_\nu\boldsymbol{U}_\rho$. The 2×2 identity matrix $\mathbf{I}$ is required to factor the XCI expression in a ν- and ρ-dependent term.

[10]An alternative formulation with ω_1 in the support of the probe is obtained by exchanging the subscripts of ω_1 and ω_3 in frequency domain and t_1 and t_3 in time domain.

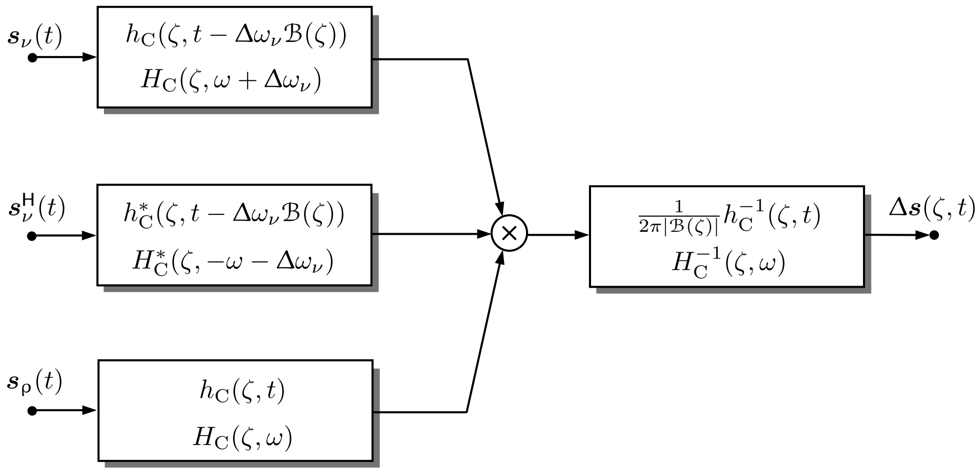


Figure 4.17: Block diagram of the basic third-order Volterra system (compare with Figure 2.1 and Figure 4.6) constituting a single spatial branch of the parallel fiber model shifted to the baseband of the ν^th wavelength channel. The input signals are given by the analog baseband signals of the probe $s_\rho(t)$ and interferer $s_\nu(t)$. The linear channel transfer function acting on the interferer must be shifted in frequency by the relative channel separation $\Delta\omega_\nu$. This corresponds to a retardation of the impulse response $h_\mathrm{C}(\zeta, t)$ in time by the walk-off $\tau_\mathrm{wo} = \Delta\omega_\nu \mathcal{B}(\zeta)$ depending on the *local* amount of accumulated dispersion, see (3.60)–(3.61) and Figure 3.11, where dispersion slope $\beta_3(z)$ is neglected.

i.e., the shifted (and now ν-dependent) nonlinear transfer function is either expressed in *relative* frequencies and defined as

$$H_{\mathrm{NL},\nu}(\boldsymbol{\upsilon}) \overset{\mathrm{def}}{=} H_\mathrm{NL}(\upsilon_1 + \Delta\omega_\nu, \upsilon_2)\,, \tag{4.83}$$

or in terms of *absolute* frequencies, as in (4.32)–(4.33), by

$$\begin{aligned} H_{\mathrm{NL},\nu}(\boldsymbol{\omega}) \overset{\mathrm{def}}{=} H_\mathrm{NL}(\omega_1 + \Delta\omega_\nu, \omega_2 + \Delta\omega_\nu, \omega_3) = &\frac{1}{L_\mathrm{eff}} \int_0^L H_\mathrm{C}(\zeta, \omega_1 + \Delta\omega_\nu) H_\mathrm{C}^*(\zeta, \omega_2 + \Delta\omega_\nu) \\ &\times H_\mathrm{C}(\zeta, \omega_3) H_\mathrm{C}^{-1}(\zeta, \omega_1 - \omega_2 + \omega_3)\, \mathrm{d}\zeta\,. \end{aligned} \tag{4.84}$$

The corresponding block diagram of the basic third-order Volterra system of a single spatial branch is shown in Figure 4.17. Here, the input signals to the nonlinear system are given by the analog baseband signals $s_\rho(t)$ and $s_\nu(t)$. The relative frequency offset between the two wavelength channels results in a frequency shift by $\Delta\omega_\nu$ of the *linear* channel transfer function $H_\mathrm{C}(\zeta, \omega)$ acting on the interfering wavelength channel. A frequency shift of the channel transfer function corresponds to a retardation in time of the channel impulse response[11] by the temporal channel walk-off $\tau_\mathrm{wo}(z, \Delta\omega_\nu)$ defined in (3.58). The temporal walk-off scales linearly with the channel separation $\Delta\omega_\nu$ and the accumulated chromatic dispersion $\mathcal{B}(z)$, and so does the memory of the nonlinear system described by the kernel $H_{\mathrm{NL},\nu}(\boldsymbol{\omega})$.

We define the ν-dependent (normalized) map strength to measure *inter-channel* (i.e., XCI) nonlinear memory using the definition of the walk-off length $L_{\mathrm{wo},\nu}$ from (3.62) as

$$\mathcal{S}_{T,\nu} \overset{\mathrm{def}}{=} \Delta\omega_\nu (2\pi R_\mathrm{s})\, \mathcal{S} = \Delta\omega_\nu R_\mathrm{s} \bar{\beta}_2 L_\mathrm{eff} = \mathrm{sign}(\bar{\beta}_2 \Delta\omega_\nu)\, L_\mathrm{eff}/L_{\mathrm{wo},\nu}, \quad \nu \neq \rho\,. \tag{4.85}$$

[11]Here, dispersion slope $\beta_3(z)$ is neglected. This approximation is valid for $2\pi B_\mathrm{WDM}\bar{\beta}_3 < \bar{\beta}_2$, e.g., $B_\mathrm{WDM} <$ 20 THz for typical fiber parameter as in G.652 [AFAK+18]. If dispersion slope is present, a ν-dependent GVD parameter can be used to approximate the dispersion slope to first-order.

where the temporal walk-off between the probe and the ν^{th} interfering wavelength channels is the relevant length scale. Here, the ratio $L_{\text{eff}}/L_{\text{wo},\nu}$ is equivalent to the number of traversed pulses over the effective length (i.e., where nonlinear interaction is most relevant) between the probe and the ν^{th} wavelength channel, see, e.g., Figure 3.11 with $z = L_{\text{eff}}$.

The expansion of the time-domain integrand from (4.63) can be done similarly as for the frequency-domain integrand. We have

$$\boldsymbol{u}(0,\, t-\tau_1)\boldsymbol{u}^{\mathsf{H}}(0, t-\tau_1-\tau_2)\boldsymbol{u}(0, t-\tau_2)h_{\text{NL}}(\tau_1,\tau_2)\Big|_{\text{supp}(\Delta\boldsymbol{S})\subset\mathcal{B}_{\rho}} \tag{4.86}$$

$$= \boldsymbol{s}_{\rho}(t-\tau_1)\boldsymbol{s}_{\rho}^{\mathsf{H}}(t-\tau_1-\tau_2)\boldsymbol{s}_{\rho}(t-\tau_2)h_{\text{NL},\rho}(\tau_1,\tau_2) \tag{4.87}$$
$$+\sum_{\nu\neq\rho}\Big(\boldsymbol{s}_{\nu}(t-\tau_1)\boldsymbol{s}_{\nu}^{\mathsf{H}}(t-\tau_1-\tau_2)+\boldsymbol{s}_{\nu}^{\mathsf{H}}(t-\tau_1-\tau_2)\boldsymbol{s}_{\nu}(t-\tau_1)\mathbf{I}\Big)\,\boldsymbol{s}_{\rho}(t-\tau_2)h_{\text{NL},\nu}(\tau_1,\tau_2)\,,$$

where we use the now ν-dependent nonlinear impulse response given by

$$h_{\text{NL},\nu}(\boldsymbol{\tau}) = h_{\text{NL}}(\tau_1,\tau_2)\exp(-\mathrm{j}\Delta\omega_{\nu}\tau_2)\,. \tag{4.88}$$

which is related by a 2D inverse Fourier transform to the *shifted* nonlinear transfer function from (4.83), i.e.,

$$H_{\text{NL},\nu}(\boldsymbol{\upsilon}) = \mathcal{F}_{[\upsilon_1,\upsilon_2]^{\mathsf{T}}\leftrightarrow[\tau_2,\tau_1]^{\mathsf{T}}}\{\, h_{\text{NL},\nu}(\boldsymbol{\tau})\,\}\,. \tag{4.89}$$

In symmetry with (4.84) and (4.70), the ν-dependent nonlinear impulse response can also be given as a function of $\mathbf{t} = [\mathrm{t}_1, \mathrm{t}_2, \mathrm{t}_3]^{\mathsf{T}}$ by

$$\begin{aligned} h_{\text{NL},\nu}(\mathbf{t}) =& \frac{1}{L_{\text{eff}}}\int_0^L \frac{1}{2\pi|\mathcal{B}(\zeta)|}h_{\text{C}}(\zeta,\mathrm{t}_1-\Delta\omega_{\nu}\mathcal{B}(\zeta))\,h_{\text{C}}^{*}(\zeta,\mathrm{t}_2-\Delta\omega_{\nu}\mathcal{B}(\zeta)) \\ & \times h_{\text{C}}(\zeta,\mathrm{t}_3)\,h_{\text{C}}^{-1}(\zeta,\underbrace{\mathrm{t}_1-\mathrm{t}_2+\mathrm{t}_3}_{t})\;\mathrm{d}\zeta\,. \end{aligned} \tag{4.90}$$

where the temporal matching with $\mathrm{t}_3 = t - \mathrm{t}_1 + \mathrm{t}_2$ is still present.

In summary, the analog baseband end-to-end relation of the continuous-time third-order nonlinear system can be given in both frequency and time domain by

$$\Delta\boldsymbol{S}(\omega)\Big|_{\omega\in\mathcal{B}_{\rho}} = \Delta\boldsymbol{S}^{\text{SCI}}(\omega)+\Delta\boldsymbol{S}^{\text{XCI}}(\omega) \tag{4.91}$$

$$= -\mathrm{j}\bar{\gamma}\frac{8}{9}L_{\text{eff}}\frac{1}{(2\pi)^2}\int_{\mathbb{R}^2}\boldsymbol{S}_{\rho}(\omega_1)\boldsymbol{S}_{\rho}^{\mathsf{H}}(\omega_2)\boldsymbol{S}_{\rho}(\omega_3)H_{\text{NL},\rho}(\boldsymbol{\omega})\,\mathrm{d}^2\boldsymbol{\omega} \tag{4.92}$$
$$-\,\mathrm{j}\bar{\gamma}\frac{8}{9}L_{\text{eff}}\frac{1}{(2\pi)^2}\sum_{\nu\neq\rho}\int_{\mathbb{R}^2}\Big(\boldsymbol{S}_{\nu}(\omega_1)\boldsymbol{S}_{\nu}^{\mathsf{H}}(\omega_2)+\boldsymbol{S}_{\nu}^{\mathsf{H}}(\omega_2)\boldsymbol{S}_{\nu}(\omega_1)\mathbf{I}\Big)$$
$$\times\,\boldsymbol{S}_{\rho}(\omega_3)H_{\text{NL},\nu}(\boldsymbol{\omega})\,\mathrm{d}^2\boldsymbol{\omega}$$

$$\Delta\boldsymbol{s}(t)\Big|_{\text{supp}(\Delta\boldsymbol{S})\subset\mathcal{B}_{\rho}} = \Delta\boldsymbol{s}^{\text{SCI}}(t)+\Delta\boldsymbol{s}^{\text{XCI}}(t) \tag{4.93}$$

$$= -\mathrm{j}\bar{\gamma}\frac{8}{9}L_{\text{eff}}\int_{\mathbb{R}^2}\boldsymbol{s}_{\rho}(\mathrm{t}_1)\boldsymbol{s}_{\rho}^{\mathsf{H}}(\mathrm{t}_2)\boldsymbol{s}_{\rho}(\mathrm{t}_3)h_{\text{NL},\rho}(\mathbf{t})\,\mathrm{d}^2\mathbf{t} \tag{4.94}$$
$$-\,\mathrm{j}\bar{\gamma}\frac{8}{9}L_{\text{eff}}\sum_{\nu\neq\rho}\int_{\mathbb{R}^2}\Big(\boldsymbol{s}_{\nu}(\mathrm{t}_1)\boldsymbol{s}_{\nu}^{\mathsf{H}}(\mathrm{t}_2)+\boldsymbol{s}_{\nu}^{\mathsf{H}}(\mathrm{t}_2)\boldsymbol{s}_{\nu}(\mathrm{t}_1)\mathbf{I}\Big)\,\boldsymbol{s}_{\rho}(\mathrm{t}_3)h_{\text{NL},\nu}(\mathbf{t})\,\mathrm{d}^2\mathbf{t}\,.$$

At this point, considering (4.62) and (4.81), we formulated the relation between the perturbation at the probe $\Delta\boldsymbol{S}(\omega)$ after chromatic dispersion compensation and the transmit spectra of the probe $\boldsymbol{S}_{\mathsf{p}}(\omega)$ and the interferers $\boldsymbol{S}_{\nu}(\omega)$ in their respective baseband. The remaining operation to arrive at a discrete-time end-to-end relation will be discussed in the following chapter.

5. The Discrete-Time Perturbation Approach

Within the preceding chapter we have looked at the continuous-time end-to-end relation between the transmit signals and the perturbation imposed on the probe signal. In the present chapter, we take the solution from (4.92)–(4.94) to derive the discrete-time end-to-end relation anticipated in (4.5)–(4.6). Our analysis in the previous chapter revealed that the perturbation can be split into its self-channel interference (SCI) and cross-channel interference (XCI) contribution. Applied to the discrete-time ansatz we have

$$\Delta\boldsymbol{a}[k] = \Delta\boldsymbol{a}^{\text{SCI}}[k] + \Delta\boldsymbol{a}^{\text{XCI}}[k] \tag{5.1}$$

$$\circ\!\!-\!\!\bullet$$

$$\Delta\boldsymbol{A}(\mathrm{e}^{\mathrm{j}\omega T}) = \Delta\boldsymbol{A}^{\text{SCI}}(\mathrm{e}^{\mathrm{j}\omega T}) + \Delta\boldsymbol{A}^{\text{XCI}}(\mathrm{e}^{\mathrm{j}\omega T})\,, \tag{5.2}$$

which will be used in the following to ease the derivation. Due to the underlying *regular perturbation* method (truncated after the first-order term), the perturbation signal in (5.1)–(5.2) is purely *additive* w.r.t. the transmit sequence $\boldsymbol{a}[k]$.

The time-domain formulation of the T-spaced discrete-time end-to-end channel was already developed in [ME12] with a focus on *cross-channel* nonlinear interference, see discussion in Section 1.1. The corresponding frequency-domain view connecting the $1/T$-periodic spectrum of the transmit and receive symbol sequences via an adequate frequency-domain Volterra kernel is one of the key results of the present work, see Section 1.2. The fundamental requirement for such a discrete-time Volterra system in frequency domain is that frequency components that appear—due to aliasing—at the sampled output of the *continuous-time* system, must also be produced by the discrete-time system which only operates within the Nyquist interval.

In the nonlinear transmission regime, a subset of the distortions can be attributed to pure phase and polarization rotations. Large *multiplicative* distortions can, however, not be modeled well using only the first-order terms in the series expansion of the perturbation approach. This in turn motivates the extension of the original *regular* perturbation approach to a combined so-called *regular-logarithmic* approach taking the multiplicative nature of certain distortions properly into account. In the context of the discrete-time RP method, the above considerations resulted in *enhanced* numerical implementations of the RP method including the so-called *pulse collision picture*, see Section 1.1. The same considerations can also be applied to discrete-time Volterra systems in $1/T$-periodic frequency-domain. The result of that, published in [FFF20], is no longer equivalent to, e.g., the pulse collision picture (since multiplication does not commute in time and frequency domain), but performs equally well in terms of accuracy.

The last section of this chapter concerns with the numerical validation of the derived models. An algorithmic implementation using the sampled and periodic spectrum of the transmit sequence and the frequency-domain kernel is presented.

5.1 The First-Order Regular Solution

We start with the solution of the electrical end-to-end model from (4.92) and (4.94). The matched filter corresponding to the transmit filter of the probe signal $h^*_{\mathrm{T},\rho}(-t) \circ\!\!-\!\!\bullet\, H^*_{\mathrm{T},\rho}(\omega)$ is applied to the sum of the linearly propagating transmit signal $\boldsymbol{s}(t) \circ\!\!-\!\!\bullet\, \boldsymbol{S}(\omega)$ and the perturbation $\Delta\boldsymbol{s}(t) \circ\!\!-\!\!\bullet\, \Delta\boldsymbol{S}(\omega)$, see Figure 4.1. We assume that the matched filter is ideally aligned with the transmit filter (or pulse) of the probe signal. The end-to-end filter cascade from the perspective of the probe signal

$$T \cdot H_{\mathrm{T},\rho}(\omega) H_{\mathrm{C}}(L,\omega)\, H^*_{\mathrm{C}}(L,\omega) H^*_{\mathrm{T},\rho}(\omega)\,, \tag{5.3}$$

forms an overall Nyquist response cancelling out inter-symbol interference on the perturbation-free part of the receive signal[1]. The matched filter is typically realized using an oversampled representation of $H^*_{\mathrm{C}}(L,\omega) H^*_{\mathrm{T},\rho}(\omega)$ in the receiver-side DSP. In a second step, T-spaced *sampling*, see (2.30), translates the continuous-time receive signal to the discrete-time domain recovering the original transmit sequence $\boldsymbol{a}[k]$ superimposed with the sampled perturbation $\Delta\boldsymbol{a}[k]$. Any additional noise source is neglected at this point.

In the following, we will only consider the additive perturbation $\Delta\boldsymbol{a}^{\mathrm{SCI}}[k]$ originating from SCI (justified due to the linearity of the prior operations), and then generalize to XCI. The sampling operation, see (2.30), in time domain (corresponding to the aliasing operation, see (2.29), in frequency domain) acting on the continuous-time perturbation (after the matched filter) results in

$$\Delta\boldsymbol{A}^{\mathrm{SCI}}(\mathrm{e}^{\mathrm{j}\omega T}) = \frac{T}{E_{\mathrm{T}}} \textsc{Alias}_{\,\omega_{\mathrm{Nyq}}}\{\, \Delta\boldsymbol{S}^{\mathrm{SCI}}(\omega) \cdot H^*_{\mathrm{T},\rho}(\omega)\,\} \tag{5.4}$$

$$\bullet\!\!-\!\!\circ$$

$$\Delta\boldsymbol{a}^{\mathrm{SCI}}[k] = \frac{T}{E_{\mathrm{T}}} \textsc{Sample}_{\,T}\{\, \Delta\boldsymbol{s}^{\mathrm{SCI}}(t) * h^*_{\mathrm{T},\rho}(-t)\,\}\,, \tag{5.5}$$

i.e., in frequency domain any signal components outside the Nyquist interval will be folded into the Nyquist interval $\mathbb{T}$, (2.28), which results in the frequency-continuous, and now $1/T$-periodic spectrum $\Delta\boldsymbol{A}^{\mathrm{SCI}}(\mathrm{e}^{\mathrm{j}\omega T})$.

[1] In practice, this condition is only approximately fulfilled by the receiver-side DSP which tries to optimize between ISI compensation and noise amplification.

The remaining objective is to relate the received discrete-time perturbation to its discrete-time transmit sequence and spectrum of the probe $\boldsymbol{a}[k] \circ\!\!-\!\!\bullet \boldsymbol{A}(e^{j\omega T})$, and to that of the ν^{th} interferer $\boldsymbol{b}_\nu[k] \circ\!\!-\!\!\bullet \boldsymbol{B}_\nu(e^{j\omega T})$. The general idea is outlined in Figure 5.1, exemplarily for the SCI contribution. The block diagram from Figure 4.2 is modified such that the transmit pulse $T \cdot H_{\mathrm{T},\nu}(\omega)$, here, for $\nu = \rho$, and the receive-side matched filter $H_{\mathrm{R}}(\omega) = \frac{T}{E_{\mathrm{T},\rho}} H_{\mathrm{C}}^*(L,\omega) H_{\mathrm{T},\rho}(\omega)$ are incorporated into the parallel, spatial branches. The block $H_{\mathrm{CD}}(L,\omega)$ at the output of Figure 4.2 cancels with $H_{\mathrm{C}}^*(L,\omega)$, as part of the matched filter, since

$$H_{\mathrm{C}}^*(L,\omega) H_{\mathrm{CD}}(L,\omega) = \exp\left(\mathcal{G}(L)/2\right) = 1, \quad \forall \omega\,, \tag{5.6}$$

with $\mathcal{G}(L) = 0$ by assumption in (3.25) and (3.63).

To that end, the dimensionless, *dispersed* and *attenuated* transmit pulse $G_{\mathrm{T},\nu}(z,\omega)$ is defined as the concatenation of the transmit pulse shape and the linear channel transfer function. We define[2]

$$\begin{aligned} G_{\mathrm{T},\nu}(z,\omega) &\overset{\text{def}}{=} T \cdot H_{\mathrm{T},\nu}(\omega)\, H_{\mathrm{C}}(z, \omega + \Delta\omega_\nu) && (5.7)\\ &\;\bullet\!\!-\!\!\circ \\ g_{\mathrm{T},\nu}(z,t) &= T \cdot \exp(-j\mathcal{B}(z)\Delta\omega_\nu^2) \int_{-\infty}^{\infty} h_{\mathrm{T},\nu}(\tau)\, h_{\mathrm{C}}(z, t - \tau - \Delta\omega_\nu \mathcal{B}(z))\, \mathrm{d}\tau\,, && (5.8) \end{aligned}$$

taking into account the transmit pulse shape (including the launch power via $P_\nu = \frac{\sigma_{b,\nu}^2}{T} E_{\mathrm{T},\nu}$) and the relative channel offset $\Delta\omega_\nu$ of the ν^{th} wavelength channel. The definition in (5.7)–(5.8) is in line with Example 3.5 and Example 3.6, where we use the path-normalized (i.e., un-attenuated) pulse $\tilde{g}_{\mathrm{T},\nu}(z,t)$. In Figure 5.1, the block diagram is drawn for the case of SCI, where, with $\nu = \rho$, the relative channel offset $\Delta\omega_\nu$ vanishes in the linear channel transfer function $H_{\mathrm{C}}(z,\omega)$.

Vice-versa, we have the adversary, z-dependent receive shape $G_{\mathrm{R}}(z,\omega)$ which is defined as the concatenation of the channel inverse and the matched filter w.r.t. the transmit pulse shape of the probe. It follows

$$\begin{aligned} G_{\mathrm{R}}(z,\omega) &\overset{\text{def}}{=} T \cdot H_{\mathrm{T},\rho}^*(\omega) H_{\mathrm{C}}^{-1}(z,\omega) = \exp(-\mathcal{G}(z))\, G_{\mathrm{T},\rho}^*(z,\omega) && (5.9)\\ &\;\bullet\!\!-\!\!\circ \\ g_{\mathrm{R}}(z,t) &= T \int_{-\infty}^{\infty} \frac{1}{2\pi|\mathcal{B}(z)|} h_{\mathrm{T},\rho}^*(\tau) h_{\mathrm{C}}^{-1}(z, t+\tau) \mathrm{d}\tau = \exp(-\mathcal{G}(z))\, g_{\mathrm{T},\rho}^*(z,-t)\,, && (5.10) \end{aligned}$$

where the receive filter $G_{\mathrm{R}}(z,\omega)$ is always matched to the *dispersed* transmit pulse of the probe $G_{\mathrm{T},\rho}(z,\omega)$, hence, omitting the additional ρ-subscript.

[2]Note, that in the original paper [DFMS16] and its predecessor [ME12] only the time-domain derivation of the theory is provided. Therein, the z-dependent transmit/receive pulses are defined as the concatenation of $h_{\mathrm{T}}(t)$ with the impulse response of the chromatic dispersion $h_{\mathrm{CD}}(z,t)$, whereas we use the linear channel transfer function $h_{\mathrm{C}}(z,t)$ which is in the context of the Volterra theory presented in Section 2.1.3.2 and Appendix A.4 the more natural notation. Beyond that, the authors of [DFMS16] use a different normalization of their signals and systems, and present their derivations only in a reduced degree of detail.

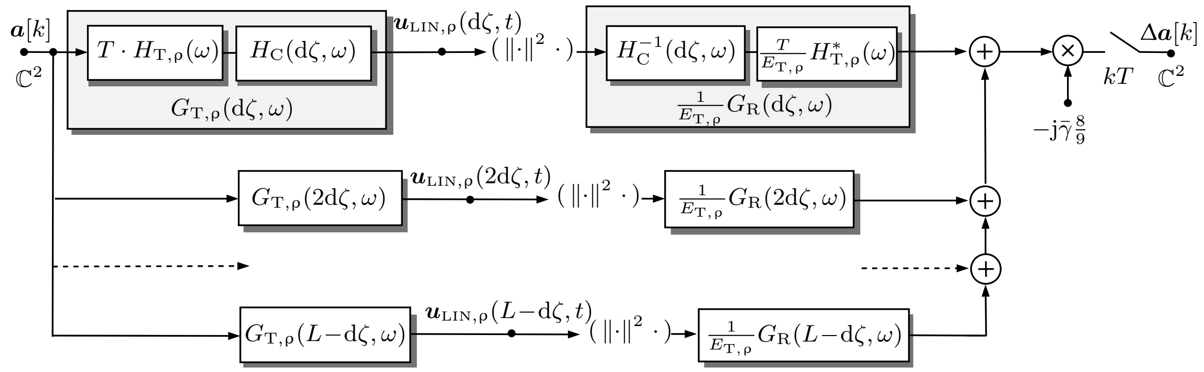


Figure 5.1: The *modified* parallel fiber-optical channel model, see also Figure 4.2, exemplarily for the SCI contribution. The transmit pulse $H_{\mathrm{T},\rho}(\omega)$ and the receive-side matched filter $H^*_{\mathrm{T},\rho}(\omega)$ are part of the parallel branches and combined with the linear channel transfer function $H_{\mathrm{C}}(\zeta,\omega)$ to the z-dependent transmit and receive pulses $G_{\mathrm{T},\rho}(\zeta,\omega)$ and $G_{\mathrm{R}}(\zeta,\omega)$.

Using the definitions of $G_{\mathrm{T},\nu}(z,\omega)$ and $G_{\mathrm{R}}(z,\omega)$, the third-order kernel of the end-to-end nonlinear system follows directly from the analysis in the last chapter, cf., e.g., (4.29)–(4.31) and (4.84). We define the *normalized* nonlinear end-to-end transfer function $H_\nu(\boldsymbol{\omega}) = H_\nu(\omega_1,\omega_2,\omega_3)$ as [DFMS15, Eq. (12)] [DW17, Eq. (14)]

$$\begin{aligned} H_\nu(\boldsymbol{\omega}) &\overset{\text{def}}{=} \frac{1}{L_{\text{eff}}}\frac{T}{E_{\mathrm{T},\nu}E_{\mathrm{T},\rho}}\int_0^L G_{\mathrm{T},\nu}(\zeta,\omega_1)G^*_{\mathrm{T},\nu}(\zeta,\omega_2)G_{\mathrm{T},\rho}(\zeta,\omega_3)G_{\mathrm{R}}(\zeta,\omega_1-\omega_2+\omega_3)\,\mathrm{d}\zeta \qquad (5.11)\\ &= \frac{T}{E_{\mathrm{T},\nu}}\;T\cdot H_{\mathrm{T},\nu}(\omega_1)\;T\cdot H^*_{\mathrm{T},\nu}(\omega_2)\;T\cdot H_{\mathrm{T},\rho}(\omega_3)\;\frac{T}{E_{\mathrm{T},\rho}}H^*_{\mathrm{T},\rho}(\omega_1-\omega_2+\omega_3)H_{\mathrm{NL},\nu}(\boldsymbol{\omega})\,, \end{aligned}$$

where we use the pre-factor $T/E_{\mathrm{T},\nu} = 1/P_\nu$ to normalize the end-to-end transfer function, see below. The nonlinear end-to-end transfer function characterizes the nonlinear cross-talk from the ν^{th} wavelength channel to the probe channel taking the transmit and receive filter into account. In particular, $H_\rho(\boldsymbol{\omega})$, i.e., with $\nu = \rho$, describes SCI and $H_\nu(\boldsymbol{\omega})$ with $\nu \neq \rho$ describes XCI.

An alternative block diagram of the intra-channel end-to-end relation with emphasis on the frequency-domain view is shown at the top of Figure 5.2. The block diagram is in direct analogy to Figure 4.5. The nonlinear transfer function $H_\rho(\boldsymbol{\omega})$ relates the periodic spectrum of the transmit *symbol* sequence to the received signal *before* sampling. The nonlinear end-to-end transfer function in (5.11) depends on the characteristics of the transmission link, comprised by the (*shifted*) nonlinear transfer function $H_{\mathrm{NL},\nu}(\boldsymbol{\omega})$ from (4.84), taking the frequency offset $\Delta\omega_\nu$ between probe and interferer into account, and the characteristics of transmit pulses $H_{\mathrm{T},\rho}(\omega)$ and $H_{\mathrm{T},\nu}(\omega)$ (assuming the matched filter receiver front-end).

We now consolidate the T-spaced sampling operation with the prior analysis by considering the *periodic continuation*, i.e., the *aliased* discrete-time equivalent of $H_\nu(\boldsymbol{\omega})$, see (2.29). It is given by

$$H_\nu(\mathrm{e}^{\mathrm{j}\boldsymbol{\omega}T}) = \frac{1}{T^3}\sum_{\boldsymbol{m}\in\mathbb{Z}^3} H_\nu(\boldsymbol{\omega} - \frac{2\pi\boldsymbol{m}}{T})\,, \qquad (5.12)$$

where the three-fold aliasing is done along each frequency dimension with $\boldsymbol{\omega} = [\omega_1,\omega_2,\omega_3]^\mathsf{T}$ and $\boldsymbol{m} = [m_1,m_2,m_3]^\mathsf{T} \in \mathbb{Z}^3$.

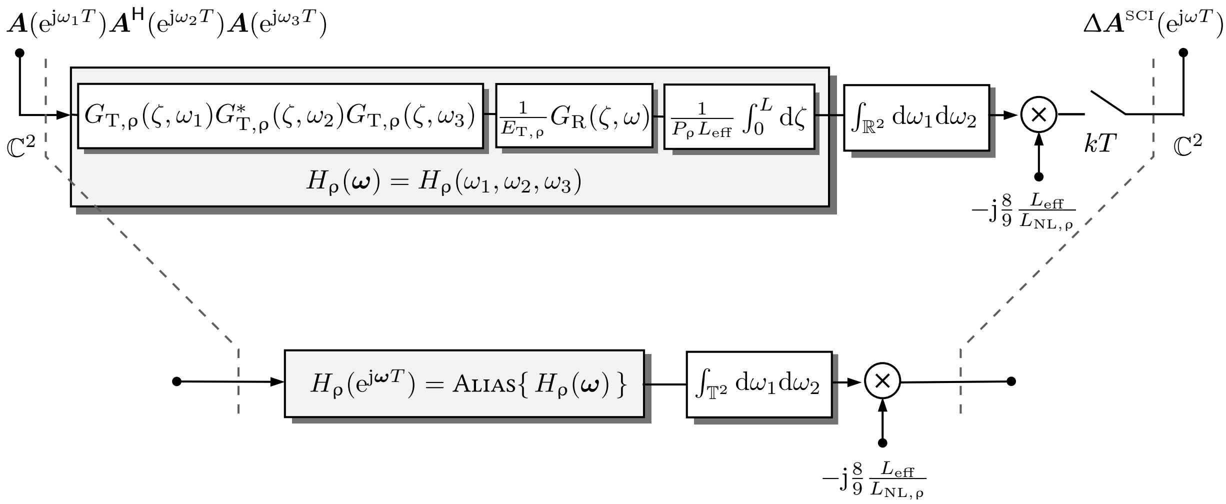


Figure 5.2: A block diagram representation of the frequency-domain continuous-time (single-channel, i.e., $\nu = \mathsf{\rho}$) perturbation model (top), and the deduced discrete-time end-to-end equivalent where T-spaced sampling is included via the *aliased* kernel representation and integration bounds that coincide with the Nyquist interval (bottom) [FFF20]. The end-to-end relation on top is the extension of Figure 4.5 including the transmit and receive filter of the probe via $G_{\mathrm{T},\mathsf{\rho}}(\zeta,\omega)$ and $G_{\mathrm{R}}(\zeta,\omega)$.

The normalization[3] by $T/E_{\mathrm{T},\nu} = 1/P_\nu$ in (5.11) is done such that $H_{\mathsf{\rho}}(\mathrm{e}^{\mathrm{j}\mathbf{0}T}) = 1$ and dimensionless. Conversely, the inverse $E_{\mathrm{T},\nu}/T = P_\nu$ is combined with the (path-average) nonlinearity coefficient $\bar{\gamma}$ to the (inverse) *nonlinear length* $1/L_{\mathrm{NL},\nu}$ from (3.74) and included in the multiplicative factor $-\frac{8}{9}\frac{L_{\mathrm{eff}}}{L_{\mathrm{NL},\nu}}$ in the block diagram. We define the (ν-dependent) *nonlinear phase shift* as

$$\phi_{\mathrm{NL},\nu} \overset{\mathrm{def}}{=} \frac{8}{9}\frac{L_{\mathrm{eff}}}{L_{\mathrm{NL},\nu}}\,, \tag{5.13}$$

which depends via $L_{\mathrm{NL},\nu}$ linearly on the launch *power* P_ν and essentially acts as a scaling factor of the nonlinear distortion's *magnitude*.

Figure 5.2 (bottom) shows the deduced discrete-time end-to-end model which now includes T-spaced sampling via the *aliased* frequency-domain kernel $H_{\mathsf{\rho}}(\mathrm{e}^{\mathrm{j}\boldsymbol{\omega}T})$. The two-fold integration over $[\omega_1,\omega_2]^{\mathsf{T}}$ is now performed over the Nyquist interval $\mathbb{T}^2$ instead of $\mathbb{R}^2$. This representation is in analogy with linear, dispersive channels as, e.g., in Figure 2.10 with $z = \mathrm{e}^{\mathrm{j}\omega T}$ [Fis02, Fig. 2.3].

In summary, we can now relate the sampled perturbation $\Delta\boldsymbol{a}[k] \circ\!\!-\!\!\bullet \Delta\boldsymbol{A}(\mathrm{e}^{\mathrm{j}\omega T})$ to the transmit sequence (here, again exemplarily for the SCI contribution) both in discrete-time and $1/T$-

---

[3]Note, that by definition the optical launch power P_ν of the ν^{th} wavelength channel is related to the pulse energy of $H_{\mathrm{T},\nu}(\omega)$ in (2.76), (2.77).

periodic frequency by

$$\Delta \boldsymbol{A}^{\text{SCI}}(\mathrm{e}^{\mathrm{j}\omega T}) = -\mathrm{j}\,\underbrace{\frac{8}{9}\frac{L_{\text{eff}}}{L_{\text{NL},\rho}}}_{\phi_{\text{NL},\rho}}\frac{T^2}{(2\pi)^2}\int_{\mathbb{T}^2}\boldsymbol{A}(\mathrm{e}^{\mathrm{j}\omega_1 T})\boldsymbol{A}^{\mathsf{H}}(\mathrm{e}^{\mathrm{j}\omega_2 T})\boldsymbol{A}(\underbrace{\mathrm{e}^{\mathrm{j}\omega_3 T}}_{\omega_3=\text{mod}_{\mathbb{T}}\{\,\omega-\omega_1+\omega_2\,\}\,\in\mathbb{T}})H_\rho(\mathrm{e}^{\mathrm{j}\boldsymbol{\omega}T})\,\mathrm{d}^2\boldsymbol{\omega} \tag{5.14}$$

$$\bullet\!-\!\circ$$

$$\Delta \boldsymbol{a}^{\text{SCI}}[k] = -\mathrm{j}\frac{8}{9}\frac{L_{\text{eff}}}{L_{\text{NL},\rho}}\sum_{\boldsymbol{\kappa}\in\mathbb{Z}^3}\boldsymbol{a}[k-\kappa_1]\boldsymbol{a}^{\mathsf{H}}[k-\kappa_2]\boldsymbol{a}[k-\kappa_3]h_\rho[\boldsymbol{\kappa}]\,. \tag{5.15}$$

The integration in (5.14) is two-fold over $[\omega_1, \omega_2]^{\mathsf{T}} \in \mathbb{T}^2$ while the time-domain summation in (5.15) is over three independent integer variables $\boldsymbol{\kappa} = [\kappa_1, \kappa_2, \kappa_3]^{\mathsf{T}} \in \mathbb{Z}^3$. This is a consequence of the time-frequency relation between convolution and element-wise multiplication, also compare with the continuous-time third-order Volterra operator in (2.63)–(2.64).

The temporal matching $\mathrm{t}_3 \overset{\text{def}}{=} t - \mathrm{t}_1 + \mathrm{t}_2$, present for the *optical* or *analog baseband* end-to-end relation in (4.63), (4.94), is now canceled in (5.15) due to the convolution with the matched filter $h_{\mathrm{T}}^*(-t)$, i.e., κ_3 does not depend on κ_1 and κ_2 in (5.15).

Note, that the frequency variable ω_3 in (5.14) still complies with the frequency matching $\omega_3 = \omega - \omega_1 + \omega_2$ where the sum on the right-hand side may lie outside the Nyquist interval $\mathbb{T}$. Due to the $1/T$-periodicity of the spectrum $\boldsymbol{A}(\mathrm{e}^{\mathrm{j}\omega T})$ and kernel $H_\rho(\mathrm{e}^{\mathrm{j}\boldsymbol{\omega}T})$, any frequency component outside $\mathbb{T}$ is effectively *folded* back into the Nyquist interval by addition of integer multiples of $2\omega_{\text{Nyq}}$. We define the related modulo function as

$$\text{mod}_{\mathbb{T}}\{\,\omega\,\} \overset{\text{def}}{=} \omega - 2\pi R_{\text{s}} \cdot \lceil \omega/(2\pi R_{\text{s}}) \rfloor\ , \tag{5.16}$$

which reduces any frequency point $\omega \in \mathbb{R}$ to the fundamental Nyquist region $\mathbb{T}$ using the rounding operation $\lceil \cdot \rfloor$. This is used to calculate the reduced frequency

$$\omega_3 = \text{mod}_{\mathbb{T}}\{\,\omega - \omega_1 + \omega_2\,\} \quad \in \mathbb{T}\,, \tag{5.17}$$

which can be interpreted as the FWM constraint (4.22) in the $1/T$-periodic frequency domain.

Example 5.1: Obtaining the $1/T$-periodic frequency-domain kernel

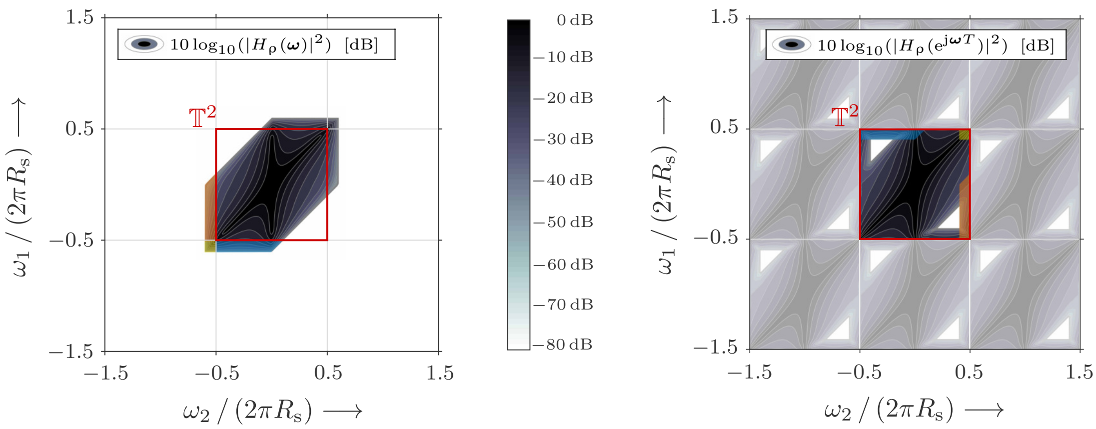


Figure 5.3: Contour plot of a 2D cut ($\omega_3 = 0$) of the 3D (single-channel, i.e., $\nu = \rho$) frequency-domain kernel $H_\rho(\boldsymbol{\omega})$ (left) [FFF20]. The parameters are the same as in Figure 4.7, and the basic pulse has RRC shape with roll-off factor $\rho = 0.2$. The non-aliased kernel on the left exhibits the well-known polygon-shape, compare, e.g., with [CBCJ14, Fig. 4] or [JA14, Fig. 2]. The projection of the Nyquist cube $\mathbb{T}^3$ into two dimensions is highlighted by the red boundaries. The discrete-time end-to-end nonlinear transfer function $H_\rho(\mathrm{e}^{\mathrm{j}\boldsymbol{\omega}T})$ (right) is obtained by aliasing the kernel $H_\rho(\boldsymbol{\omega})$ into the Nyquist cube over all three dimensions $[\omega_1, \omega_2, \omega_3]^\mathsf{T}$. The three color-coded regions indicate where the spectral components outside the Nyquist region appear after the aliasing operation. As the notation implies, $H_\rho(\mathrm{e}^{\mathrm{j}\boldsymbol{\omega}T})$ has a $1/T$-periodic structure into all three dimensions—shown here by the transparent continuations in the ω_1-ω_2-plane.

In Figure 5.3, the contour plot of a 2D cut (i.e., a 2D slice with fixed ω_3) from $H_\rho(\boldsymbol{\omega})$ before aliasing, and the corresponding 2D cut from $H_\rho(\mathrm{e}^{\mathrm{j}\boldsymbol{\omega}T})$ after aliasing to the Nyquist interval is shown. The same link parameters are used as in Example 4.1. The non-aliased kernel $H_\rho(\boldsymbol{\omega})$ on the left-side exhibits the well-known polygon-shape of the SCI contribution for $\omega_3 = 0$, compare, e.g., with the integration islands displayed in Figure 4.16 (left) for the probe channel, or [CBCJ14, Fig. 4] and [JA14, Fig. 2]. The portion of spectral support outside the Nyquist region depends on the roll-off factor, given an RRC transmit shape with matched filter receiver, and may extend to $\omega_i/(2\pi R_\mathrm{s}) = \pm 1$ in all three dimensions $i = 1, 2, 3$. In this example, a roll-off factor of $\rho = 0.2$ is chosen.

The aliased kernel on the right-side is obtained by the operation defined in (5.12) and can be visually understood as folding, see (5.16), the frequency components from outside the Nyquist interval into the Nyquist interval (indicated by the red square in Figure 5.3). For illustration, the color-coded regions display some selected locations of the spectral components before and after aliasing.

The folded spectrum $H_\rho(\mathrm{e}^{\mathrm{j}\boldsymbol{\omega}T})$ has a $1/T$-periodic structure in all three dimensions $[\omega_1, \omega_2, \omega_3]^\mathsf{T}$, indicated in Figure 5.3 by the transparent continuations. Here, only the cut $\omega_3 = 0$ is shown, but the periodicity also extends to the third dimension ω_3. This is visualized in Figure 5.4 on the following page where 2D cuts of the same kernel $H_\rho(\mathrm{e}^{\mathrm{j}\boldsymbol{\omega}T})$ for varying ω_3 are shown. In particular, the cuts at the positive and negative Nyquist frequency $\omega_3/(2\pi R_\mathrm{s}) = 0.5$ and $\omega_3/(2\pi R_\mathrm{s}) = -0.5$ are identical.

The kernel $H_\rho(\mathrm{e}^{\mathrm{j}\boldsymbol{\omega}T})$ is real-valued on the main-diagonal $\omega_1 = \omega_2$, see (5.24), and has unity magnitude except for the regions affected by aliasing, e.g., the yellow patch in Figure 5.4. The kernel exhibits a similar signature for the case $\omega_2 = \omega_3$, see Figure 5.4, which is a particular property of the intra-channel kernel with $\nu = \rho$. The gradient of kernel depends mainly on the symbol rate R_s (for a fixed dispersion coefficient $\bar{\beta}_2$), see Figure 4.8, and is directly related to the temporal width of the time-domain kernel, see Example 5.2.

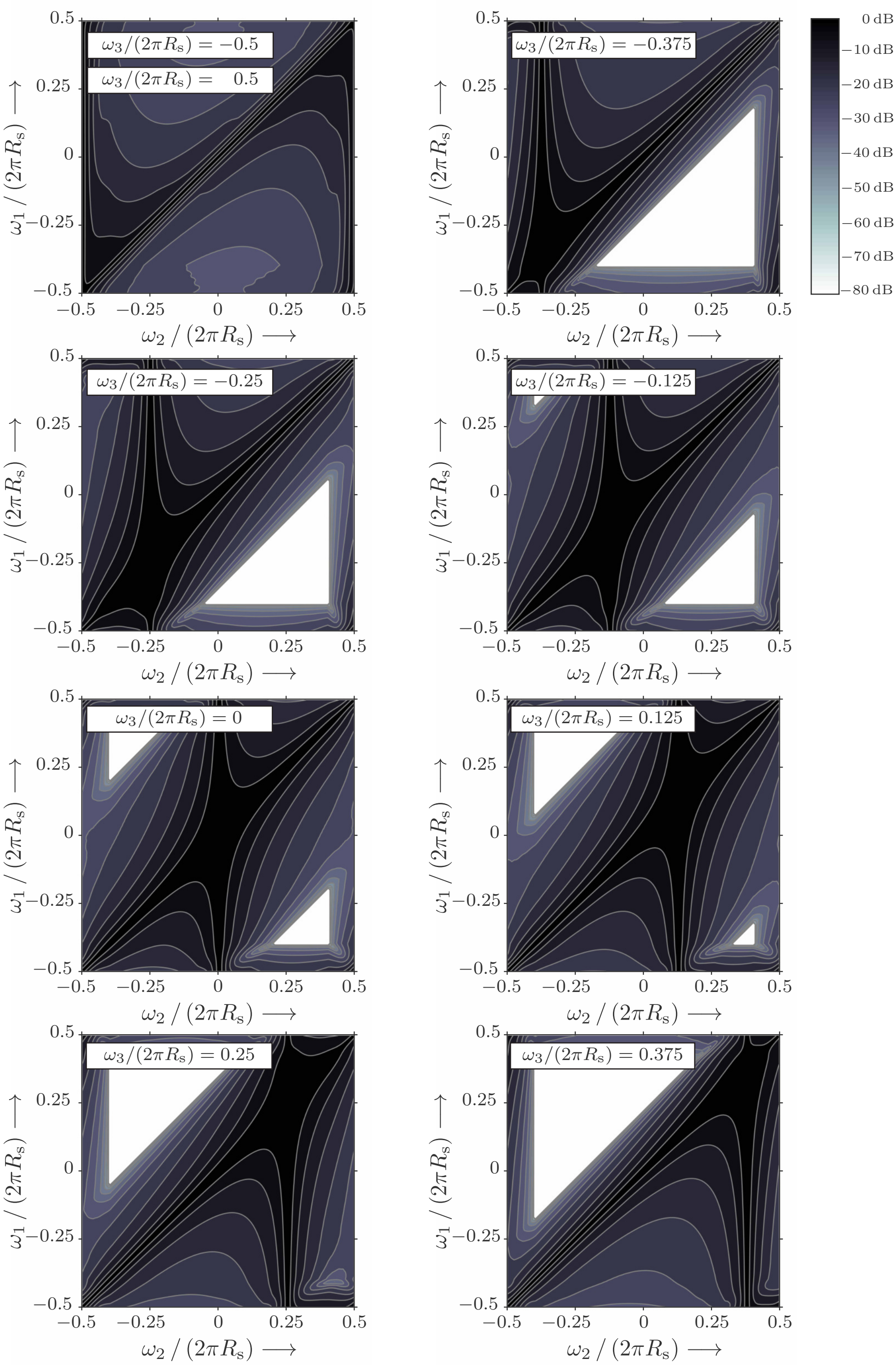


Figure 5.4: Contour plot of the 2D cut of the frequency-domain kernel $10\log_{10}(|H_\rho(\mathrm{e}^{\mathrm{j}\boldsymbol{\omega}T})|^2)$ for varying ω_3. The cut with $\omega_3/(2\pi R_\mathrm{s}) = 0$ coincides with the fundamental Nyquist region $\mathbb{T}^2$ in Figure 5.3 (right). Due to the periodicity into the ω_3-direction the cut for $\omega_3/(2\pi R_\mathrm{s}) = -0.5$ and $\omega_3/(2\pi R_\mathrm{s}) = 0.5$, i.e., the positive and negative Nyquist frequency in ω_3, are identical.

The corresponding result to (5.15) for the discrete-time, *inter*-channel end-to-end relation was published in the seminal paper by A. Mecozzi and R.-J. Essiambre in [ME12, Eq. (61), (62)] and later used in [DFMS13, Eq. (6), (7)]. A recent review of the time-domain perturbative model is part of the book chapter in [BDS+20, Sec. 9.5]. The derivation provided in the literature [ME12] is based on reasoning connected to the *time-domain* perturbation theory. The counterpart in frequency domain has not been considered before.

Based on the previous results in (4.92)–(4.94) and the considerations on the intra-channel model, see above, the XCI complement to (5.14)–(5.15) follows as

$$\begin{aligned}\Delta\boldsymbol{A}^{\text{XCI}}(\mathrm{e}^{\mathrm{j}\omega T}) = -\mathrm{j}\sum_{\nu\neq\rho}\frac{8}{9}\frac{L_{\text{eff}}}{L_{\text{NL},\nu}}\frac{T^2}{(2\pi)^2}\int_{\mathbb{T}^2}\Big(&\boldsymbol{B}_\nu(\mathrm{e}^{\mathrm{j}\omega_1 T})\boldsymbol{B}_\nu^{\mathsf{H}}(\mathrm{e}^{\mathrm{j}\omega_2 T}) + \boldsymbol{B}_\nu^{\mathsf{H}}(\mathrm{e}^{\mathrm{j}\omega_2 T})\boldsymbol{B}_\nu(\mathrm{e}^{\mathrm{j}\omega_1 T})\mathbf{I}\Big)\\ &\times\boldsymbol{A}(\mathrm{e}^{\mathrm{j}\omega_3 T})H_\nu(\mathrm{e}^{\mathrm{j}\omega T})\,\mathrm{d}^2\boldsymbol{\omega}\end{aligned} \tag{5.18}$$

$$\bullet\!\!-\!\!\circ$$

$$\begin{aligned}\Delta\boldsymbol{a}^{\text{XCI}}[k] = -\mathrm{j}\sum_{\nu\neq\rho}\frac{8}{9}\frac{L_{\text{eff}}}{L_{\text{NL},\nu}}\sum_{\boldsymbol{\kappa}\in\mathbb{Z}^3}\Big(&\boldsymbol{b}_\nu[k-\kappa_1]\boldsymbol{b}_\nu^{\mathsf{H}}[k-\kappa_2] + \boldsymbol{b}_\nu^{\mathsf{H}}[k-\kappa_2]\boldsymbol{b}_\nu[k-\kappa_1]\mathbf{I}\Big)\\ &\times\boldsymbol{a}[k-\kappa_3]h_\nu[\boldsymbol{\kappa}]\,.\end{aligned} \tag{5.19}$$

The time-domain description of the T-spaced channel model in (5.19) is equivalent to the *pulse-collision picture* (cf. [DFMS16, Eq. (3-4)] and [Dar16, Eq. (3-4)]). Compared to the derivation provided in [ME12], we arrive at the same result in (5.19) following a derivation in frequency-domain. A detailed, alternative derivation for the end-to-end relation in (5.19) is presented in Appendix A.4. Here, we take advantage of the theory on the time-domain Volterra series developed in Section 2.1.3.2. Due to the equivalence between the Volterra and perturbation theory [VSB02], we again arrive at the same result in (5.19) under a somewhat different premise. Some exemplary time- and frequency-domain, intra- and inter-channel kernels $h_\nu[\boldsymbol{\kappa}]$ and $H_\nu(\mathrm{e}^{\mathrm{j}\omega T})$ are discussed in Example 5.1 and Example 5.2.

The time-domain and *aliased* frequency-domain kernel are related by a three-dimensional (3D) DTFT according to

$$h_\nu[\boldsymbol{\kappa}] = \hat{\mathcal{F}}^{-1}_{\boldsymbol{\kappa}\leftrightarrow\boldsymbol{\omega}}\{\,H_\nu(\mathrm{e}^{\mathrm{j}\boldsymbol{\omega}T})\,\}\,, \tag{5.20}$$

where the time-domain kernel $h_\nu[\boldsymbol{\kappa}] = h_\nu[\kappa_1,\kappa_2,\kappa_3]$ can be equivalently derived using the Volterra theory from (2.65), see Appendix A.4, to obtain

$$\begin{aligned}h_\nu[\boldsymbol{\kappa}] =& \frac{1}{L_{\text{eff}}}\frac{T}{E_{\text{T},\nu}E_{\text{T},\rho}}\int_0^L\int_{-\infty}^{\infty} g_{\text{T},\nu}(\zeta,\kappa_1 T-\sigma)g^*_{\text{T},\nu}(\zeta,\kappa_2 T-\sigma) \qquad (5.21)\\ &\times g_{\text{T},\rho}(\zeta,\kappa_3 T-\sigma)g_{\text{R}}(\zeta,\sigma)\,\mathrm{d}\sigma\mathrm{d}\zeta\\ =& \frac{1}{L_{\text{eff}}}\frac{T}{E_{\text{T},\nu}E_{\text{T},\rho}}\int_0^L\exp(-\mathcal{G}(\zeta))\int_{-\infty}^{\infty} g_{\text{T},\nu}(\zeta,\kappa_1 T-\sigma)g^*_{\text{T},\nu}(\zeta,\kappa_2 T-\sigma) \qquad (5.22)\\ &\times g_{\text{T},\rho}(\zeta,\kappa_3 T-\sigma)g^*_{\text{T},\rho}(\zeta,\sigma)\,\mathrm{d}\sigma\mathrm{d}\zeta\,.\end{aligned}$$

Here, the additional integral, compared to the Volterra theory in (2.65), over ζ accounts for the parallel, spatial branches, of which one is shown in Figure 5.5. The time-domain kernel is effectively *sampled* at $\boldsymbol{\tau} = \boldsymbol{\kappa}T$, hence the aliasing operation in (5.12). As in frequency domain,

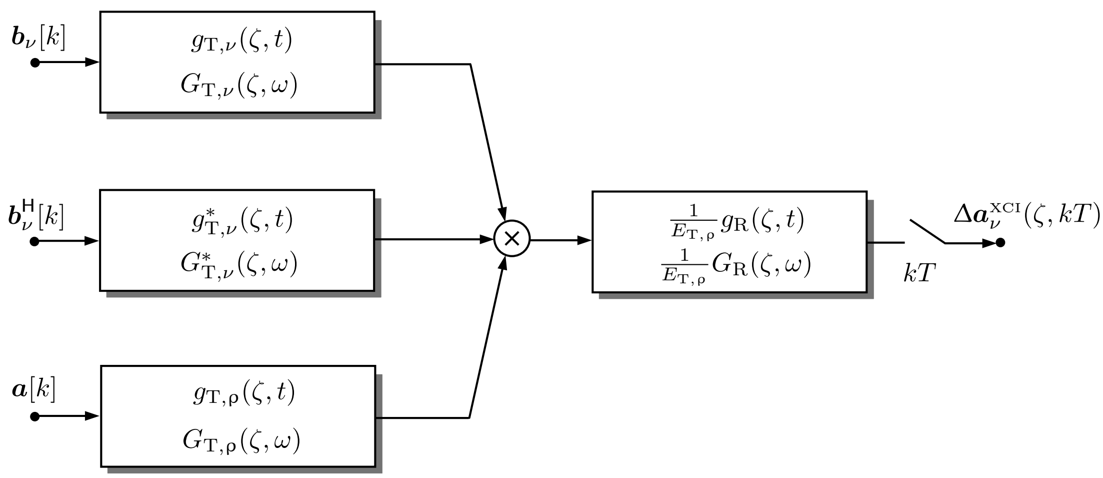


Figure 5.5: Block diagram of the basic third-order Volterra system (compare with Figure 2.1, Figure 4.6, and Figure 4.17) constituting a single spatial branch of the discrete-time end-to-end model of the ν^{th} wavelength channel. The input signals are given by the discrete-time transmit sequences of the probe $\boldsymbol{a}[k]$ and interferer $\boldsymbol{b}_\nu[k]$. The linear system acting on the interferer $g_{\mathrm{T},\nu}(\zeta,t) \circ\!\!-\!\!\bullet\, G_{\mathrm{T},\nu}(\zeta,\omega)$ already takes the relative channel separation $\Delta\omega_\nu$ into account. This corresponds to a retardation of the dipsersed basic pulse in time by the walk-off $\tau_{\mathrm{wo}} = \Delta\omega_\nu \mathcal{B}(\zeta)$ depending on the *local* amount of accumulated dispersion, see, e.g., Example 3.6.

the kernel $h_\nu[\boldsymbol{\kappa}]$ is normalized using the same pre-factor $\frac{1}{L_{\mathrm{eff}}}\frac{T}{E_{\mathrm{T},\nu}}$. Note, that the exponential $\exp(-\mathrm{j}\mathcal{B}(z)\Delta\omega_\nu^2)$ present in $g_{\mathrm{T},\nu}(z,t)$ is canceled considering the product $g_{\mathrm{T},\nu}(\zeta,\kappa_1 T - \sigma) g^*_{\mathrm{T},\nu}(\zeta,\kappa_2 T - \sigma)$ in (5.21).

Kernel Symmetry We now review the kernel properties of the aliased frequency-domain kernel which can be deduced from (5.11), (5.12) and, e.g., based on the symmetric properties of the basic third-order system shown in Figure 5.5 (top), where we assume the matched filter receiver front-end. We have

$$H_\nu(\mathrm{e}^{\mathrm{j}\boldsymbol{\omega}T}) \in \mathbb{R}, \qquad \text{if } \omega_2 = \omega_1 \;\Leftrightarrow\; \omega_3 = \omega \tag{5.23}$$

$$H_\rho(\mathrm{e}^{\mathrm{j}\boldsymbol{\omega}T}) \in \mathbb{R}, \qquad \text{if } \omega_2 = \omega_1 \;\Leftrightarrow\; \omega_3 = \omega \tag{5.24}$$
$$\vee\ \omega_2 = \omega_3 \;\Leftrightarrow\; \omega_1 = \omega\,,$$

where the two (*doubly-degenerate*[4]) cases $\omega_2 = \omega_1$ and $\omega_2 = \omega_3$ (for $\nu = \rho$) result in a purely real-valued kernel[5]. The first degenerate case with $\omega_1 = \omega_2$ corresponds to the *diagonal* of the Nyquist region in Figure 5.3, shown here for $\omega_3 = 0$, and in Figure 5.4. For the special case that the transmit pulses $h_{\mathrm{T},\rho}(t)$ and $h_{\mathrm{T},\nu}(t)$ have RRC shape with roll-off factor $\rho = 0$ (i.e., no spectral support outside the Nyquist region), we find that the folded kernel always takes the value 1 on the diagonal—independent of ω_3. In general, the elements on the diagonal are always real-valued and close to the value 1 (i.e., quasi *frequency-flat* on the ω_1-ω_2-diagonal), and the flatness depends on the amount of spectral support which is folded into the Nyquist region, cf. the yellow region in Figure 5.3. This approximation will be used later to simplify the expression for the average phase- and polarization rotation, see next section.

[4]Out of the four interacting frequencies ω, ω_1, ω_2, ω_3, two come as identical pairs due to (5.17).

[5]This corresponds to classical inter- and intra-channel XPM, i.e., with pure phase modulation, see analysis in the following section

We also find the following symmetry properties of the intra-channel kernel

$$H_\rho(\mathrm{e}^{\mathrm{j}\,[\omega_1,\omega_2,\omega_3]^\mathsf{T}T}) = H_\rho(\mathrm{e}^{\mathrm{j}\,[\omega_3,\omega_2,\omega_1]^\mathsf{T}T}) \tag{5.25}$$

$$H_\rho(\mathrm{e}^{\mathrm{j}\,[\omega_1,\omega_2,\omega_3]^\mathsf{T}T}) = H_\rho(\mathrm{e}^{\mathrm{j}\,[-\omega_1,-\omega_2,-\omega_3]^\mathsf{T}T})\,, \tag{5.26}$$

and we can conclude form (5.25) that the cut of the kernel shown in Figure 5.3 is equivalent to the cut with $\omega_1 = 0$, shown in the ω_3-ω_2-plane instead. Similarly, in Figure 5.4, ω_1 and ω_3 can be swapped without changing the picture.

We continue to review the kernel properties in discrete-time domain. Here, the properties directly follow from (5.22), and the symmetry properties of the basic third-order system shown in Figure 5.5 (bottom). We have

$$h_\nu[\kappa_1,\kappa_2,0] \in \mathbb{R}, \qquad \text{if } \kappa_2 = \kappa_1 \tag{5.27}$$

$$h_\nu[\kappa_1,\kappa_2,0] = h_\nu^*[\kappa_2,\kappa_1,0] \in \mathbb{C} \qquad \text{if } \kappa_2 \neq \kappa_1\,, \tag{5.28}$$

where the case in (5.27) is doubly-degenerate (i.e., $\kappa_1 = \kappa_2$ and $\kappa_3 = 0$ leads to two matching pulses in (5.22), which is later termed *two-pulse collision*), and the kernel is real-valued[6]. For the second case (5.28), the kernel is generally complex-valued but conjugate-symmetric w.r.t. permutation of $[\kappa_1,\kappa_2]^\mathsf{T}$. Due to the double sum over all $[\kappa_1,\kappa_2]^\mathsf{T}$ (with the constraint $\kappa_1 \neq \kappa_2$ and $\kappa_3 = 0$, later termed *three-pulse collision*) in (5.19) the overall perturbation from this subset of $[\kappa_1,\kappa_2]^\mathsf{T} \in \mathbb{Z}^2$ is still real-valued.

Additionally, for intra-channel contributions ($\nu = \rho$) we find the following symmetry properties of the kernel

$$h_\rho[\kappa_1,\kappa_2,\kappa_3] = h_\rho[\kappa_3,\kappa_2,\kappa_1] \tag{5.29}$$

$$h_\rho[\kappa_1,\kappa_2,\kappa_3] = h_\rho[-\kappa_1,-\kappa_2,-\kappa_3]\,, \tag{5.30}$$

and we identify a second degenerate case corresponding to (5.27)–(5.28) but with $\kappa_1 = 0$ for the intra-channel case, cf., e.g., the symmetric form of (5.15) w.r.t. κ_1 and κ_3.

[6]The transmit pulse-shapes $h_{\mathrm{T},\rho}(t)$ and $h_{\mathrm{T},\nu}(t)$ are assumed to be a real-valued (root) raised-cosine, and the receiver uses the matched filter front-end.

Example 5.2: Time- and frequency-domain, intra- and inter-channel kernels

This example follows up on Example 5.1. Here, we show a frequency-domain kernel that describes the nonlinear cross-talk between a single interfering wavelength channel (with channel index ν) and the probe channel (with channel index $\mathsf{\rho}$). In this setting, we assume the same system parameters as in Example 4.1. Additionally, the interferer has the same symbol rate $R_\mathrm{s} = 64\,\mathrm{GBd}$ and RRC pulse-shape with $\rho = 0.2$ as the probe channel and is spaced at a relative channel offset of $\Delta\omega_\nu/(2\pi) = 80\,\mathrm{GHz}$.

The contour plot in logarithmic scale of the aliased kernel $H_\nu(\mathrm{e}^{\mathrm{j}\boldsymbol{\omega}T})$ is shown in Figure 5.6. Again, we can observe the periodic structure of the kernel within the 3D spectral support. Similar to the intra-channel, frequency-domain kernel in Figure 5.4, the main diagonal with $\omega_1 = \omega_2$ is real-valued, see (5.23), however, the second degenerate case with $\omega_2 = \omega_3$ is missing. The kernel gradient is much steeper compared to the intra-channel kernel, suggesting that the corresponding time-domain kernel has a greater temporal extent (i.e., the system is more dispersive). With increasing channel separation the gradient will increase accordingly. This can be seen, e.g., in Figure 4.16, as the integration bounds of the shifted nonlinear transfer function $H_\mathrm{NL}(\upsilon_1, \upsilon_2)$ move into regions of stronger slopes with increasing channel separation.

Next, we discuss the corresponding time-domain kernels related to Figure 5.4 and Figure 5.6. The time-domain kernel $h_\nu[\boldsymbol{\kappa}]$ can be computed either via a two-fold integral over the spatial and temporal domain ζ and σ, see (5.21)–(5.22), or via a 3D inverse DTFT, see the correspondence in (5.20). Performing the numerical integration in either of the two former methods requires proper *sampling* of the integrand. E.g., for the DTFT method, we effectively used a DFT with 512 equidistant samples per dimension to carry out the transform. This choice results in a good match w.r.t. the two-fold integral over ζ and σ using small step-sizes $\Delta\zeta \ll L$ and $\Delta\sigma \ll T$, and finite integration bounds. Note, that using a DFT will inherently lead to cyclic artifacts if the number of samples M per dimension is too small for the system memory specific to the kernel. E.g., going to higher channel offsets $\Delta\omega_\nu/(2\pi) \gg R_\mathrm{s}$ and large $L \gg L_\mathrm{D}$ will make the DFT method computationally challenging.

Considering Figure 5.7 and Figure 5.8, the time-domain kernels $h_\mathsf{\rho}[\boldsymbol{\kappa}]$ and $h_\nu[\boldsymbol{\kappa}]$ also have a symmetric structure, see (5.27)–(5.30), and are relatively sparse compared to the frequency-domain kernels, i.e., much of the nonlinear interaction is captured by relatively few kernel coefficients. E.g., many of the perturbation-based methods for intra-channel fiber nonlinearity compensation in time domain only use the 2D cut with $\kappa_2 = \kappa_1 + \kappa_3$ in Figure 5.7, see, e.g., [GCK+14,ONO+14,ZRB+14]. In [FES+18], an improved approach is presented which takes all of the 2D cuts into account leading to better performance/complexity trade-offs.

We recap from Example 4.2 that the normalized map strength for $R_\mathrm{s} = 64\,\mathrm{GBd}$ and $\rho = 0.2$ calculates to $(1+\rho)|\mathcal{S}_{T,\mathsf{\rho}}| \approx 14$ which is in the same order as the interaction length of the *intra*-channel, time-domain kernel $h_\mathsf{\rho}[\boldsymbol{\kappa}]$ in Figure 5.7. Similarly, the interaction length of the *inter*-channel, time-domain kernel $h_\nu[\boldsymbol{\kappa}]$ can be approximated by[a]

$$(1+\rho)(|\mathcal{S}_{T,\mathsf{\rho}}| + |\mathcal{S}_{T,\nu}|) = (1+\rho)\,R_\mathrm{s}\,(2\pi R_\mathrm{s} + |\Delta\omega_\nu|)|\bar{\beta}_2|L_\mathrm{eff} \approx 32\,,$$

i.e., the combined effect of intra-channel, nonlinear memory and inter-channel walk-off over the effective length L_eff.

[a]Note, that when comparing Figure 5.7 and Figure 5.8, the coefficients are normalized by their respective center coefficient with a relative ratio of $10\log_{10}(h_\mathsf{\rho}[\mathbf{0}]^2/h_\nu[\mathbf{0}]^2) \approx 10\,\mathrm{dB}$.

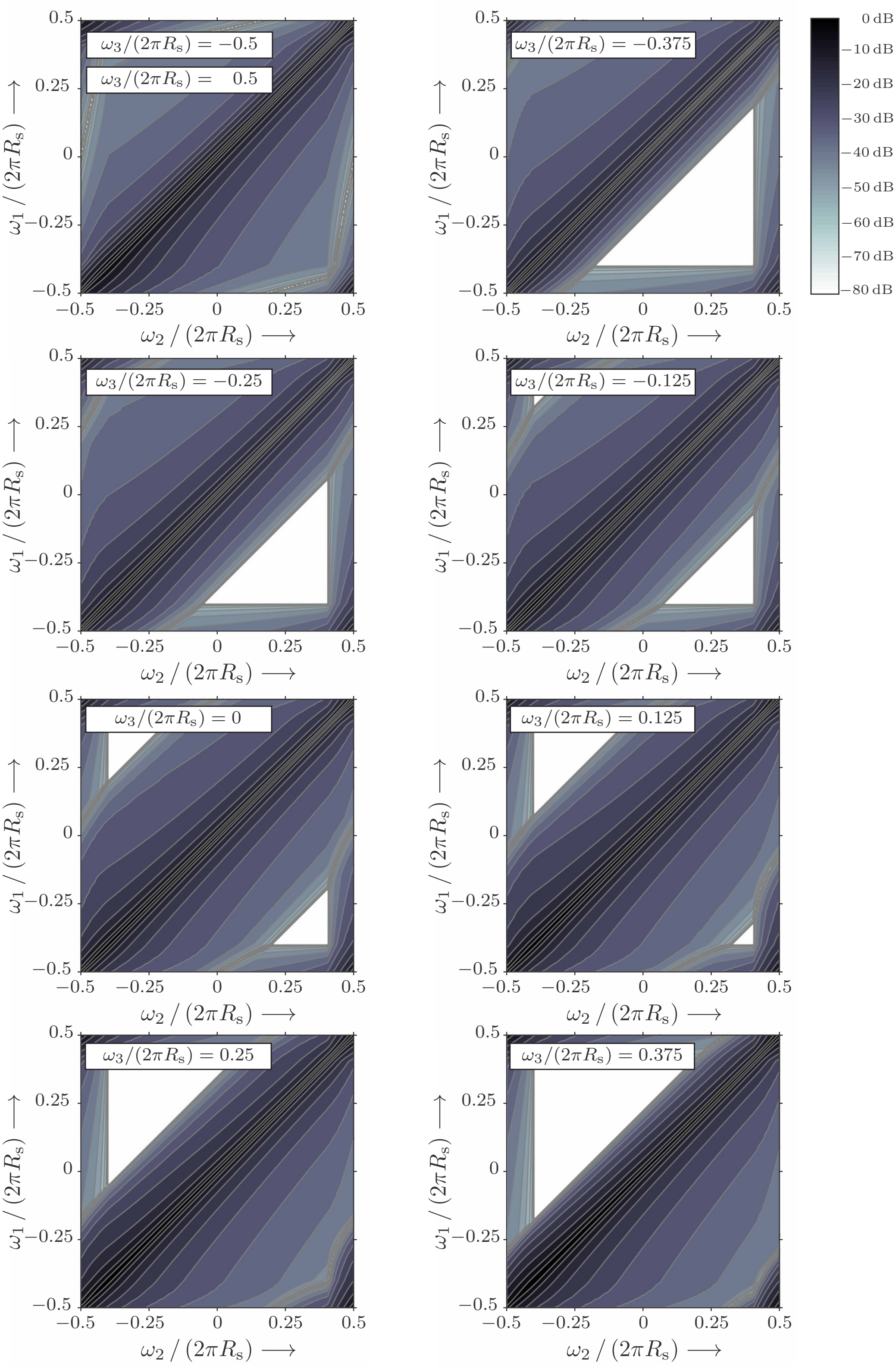


Figure 5.6: Contour plot of the 2D cuts of the frequency-domain kernel $10\log_{10}(|H_\nu(\mathrm{e}^{\mathrm{j}\boldsymbol{\omega}T})|^2)$ of the ν^{th} interferer with relative channel offset $\Delta\omega_\nu/(2\pi) = 80\,\text{GHz}$. We assume the same link parameters as in Figure 5.4, i.e., $\bar{\beta}_2 = -21\ \text{ps}^2/\text{km}$, $\mathcal{B}_0 = 0\ \text{ps}^2$, $10\log_{10}\mathrm{e}^{\bar{\alpha}} = 0.2\ \text{dB/km}$ and $L = L_{\text{sp}} = 100\ \text{km}$, see also Example 4.1. The interferer has the same symbol rate, $R_s = 64\,\text{GBd}$, and transmit shape, RRC with roll-off $\rho = 0.2$, as the probe.

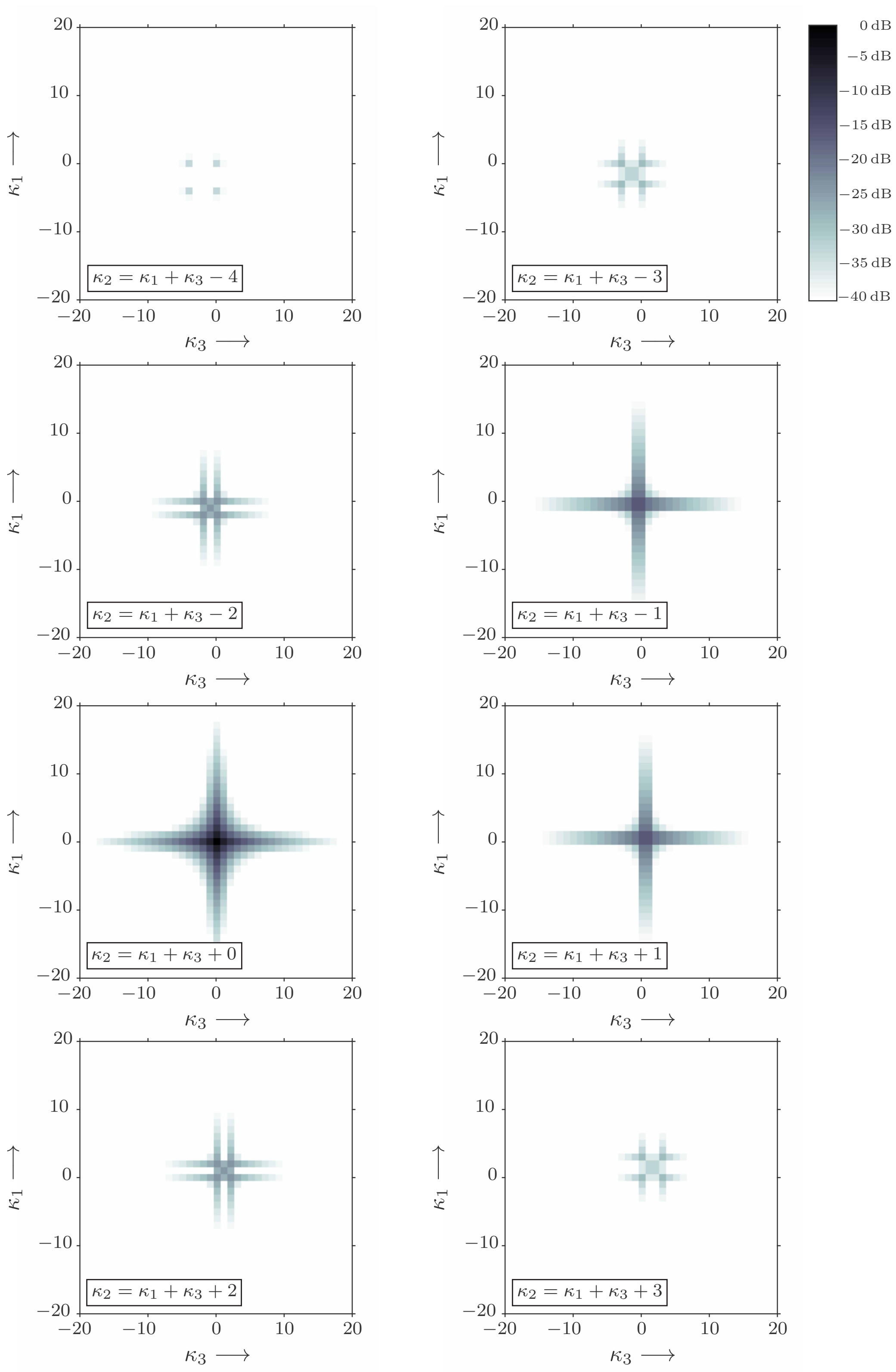


Figure 5.7: Contour plot of the 2D cut of the time-domain kernel $10\log_{10}(|h_\rho[\boldsymbol{\kappa}]|^2)$ normalized to the energy of the center coefficient $|h_\rho[\mathbf{0}]|^2 \approx 20$ for varying κ_2. The same link and signal parameters are used as in Figure 5.4, i.e., the time-domain kernel $h_\rho[\boldsymbol{\kappa}]$ can be obtained by a 3D inverse DTFT of $H_\rho(\mathrm{e}^{\mathrm{j}\boldsymbol{\omega}T})$. In the literature, often only the cut with $\kappa_2 = \kappa_1 + \kappa_3$ is shown, i.e., the case that corresponds to the temporal matching constraint (which is *not* present anymore for the discrete-time end-to-end model).

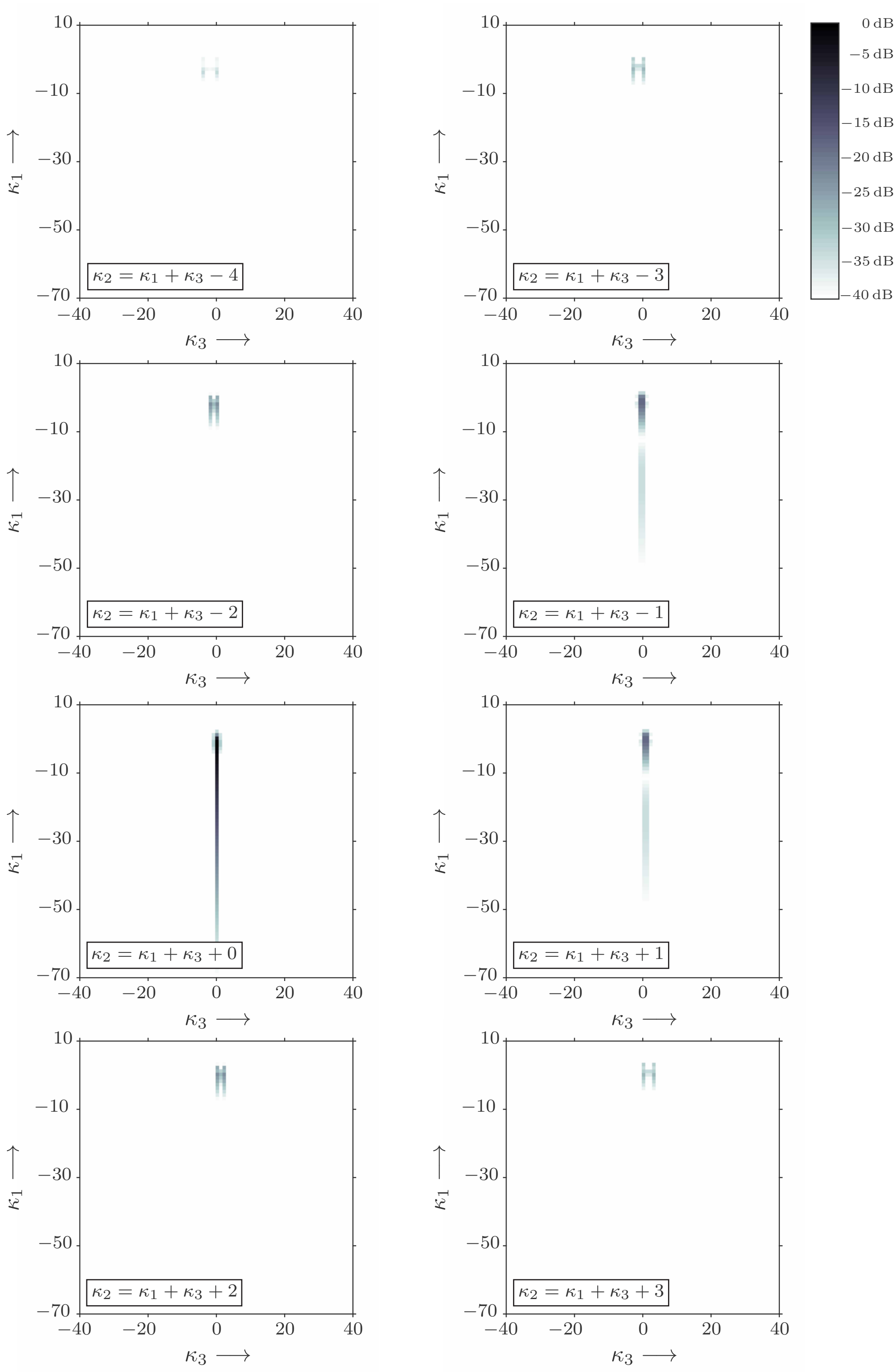


Figure 5.8: Contour plot of the 2D cut of the time-domain kernel $10\log_{10}(|h_\nu[\boldsymbol{\kappa}]|^2)$ of the ν^{th} interferer with relative channel offset $\Delta\omega_\nu/(2\pi) = 80\,\text{GHz}$, normalized to the center coefficient $|h_\nu(\mathbf{0})|^2 \approx 2$. Note, that compared to Figure 5.7 the κ_1-κ_3-axes scaling has changed. The time-domain kernel $h_\nu[\boldsymbol{\kappa}]$ can be obtained by a 3D inverse DTFT of $H_\nu(\mathrm{e}^{\mathrm{j}\boldsymbol{\omega}T})$ from Figure 5.6. The channel walk-off of about 15 symbol periods between the probe and the interferer translates into a significant temporal spread of the nonlinear response.

5.1.1 Kernel Energy and Relation to the GN-Model

Parseval's theorem applied to (5.20) yields the energy of the time- and frequency-domain coefficients of the discrete-time, third-order Volterra system

$$E_{h,\nu} \overset{\text{def}}{=} \sum_{\boldsymbol{\kappa}\in\mathbb{Z}^3} |h_\nu[\boldsymbol{\kappa}]|^2 = \left(\frac{T}{2\pi}\right)^3 \int_{\mathbb{T}^3} |H_\nu(\mathrm{e}^{\mathrm{j}\boldsymbol{\omega}T})|^2 \,\mathrm{d}^3\boldsymbol{\omega} \overset{\text{def}}{=} E_{H,\nu}\,, \tag{5.31}$$

where the right-hand side of (5.31) is closely related to the GN-model [Pog12], see discussion in Section 1.1. In contrast to (5.31), the conventional GN-model performs the integration over the *un-aliased* kernel $|H_\nu(\boldsymbol{\omega})|^2$ instead, i.e.,

$$E_{H,\nu}^{(\mathrm{GN})} \overset{\text{def}}{=} \left(\frac{T}{2\pi}\right)^3 \int_{\mathbb{R}^3} |H_\nu(\boldsymbol{\omega})|^2 \,\mathrm{d}^3\boldsymbol{\omega}\,, \tag{5.32}$$

where the integration is performed over $[\omega_1,\omega_2,\omega_3]^\mathsf{T} \in \mathbb{R}^3$. In doing so, the GN-model calculates the strength of the nonlinear perturbation *before* T-spaced sampling, on the basis of the *continuous-time* end-to-end model (see, e.g., the comprehensive derivation of the GN-model in [PBC$^+$12], or the more recent review in the textbook [BDS$^+$20]). When implemented for numerical evaluation, the GN-model from equation (5.32) is essentially discretized, i.e., *sampled* in frequency domain, and summed over the contribution from each wavelength channel[7], see the *Sampled GN-model* in Figure 1.1 (bottom). This sampling operation in frequency domain is an implicit result of the signal model assumed in the derivation of the GN-model, where the PSD of a communication signal $\boldsymbol{U}(\omega)$ is modeled as the sum of equidistant Dirac functions $\delta(\omega)$ shifted by some fixed delta in frequency, see [PBC$^+$12, Eq. (13)] or [BDS$^+$20, Eq. (9.9)].

The GN-model equation in its conventional form, see [Pog12, Eq. (1)], is often only evaluated at a single frequency $\omega = \omega_1 - \omega_2 + \omega_3$, typically at the center-frequency of the probe. This reduces the triple integral in (5.32) to a double integral over $[\omega_1,\omega_2]^\mathsf{T} \in \mathbb{R}^2$ where ω_3 must comply to the FWM constraint (4.22). Beyond that, the common pre-factor $(\frac{8}{9}\frac{L_\mathrm{eff}}{L_{\mathrm{NL},\nu}})^2$, not shown in (5.32), is included in the standard GN-model equation as it acts as a scaling factor to the power/variance of the nonlinear distortion.

There is a multitude of variations and approximations of the GN-model. A common approximation is done by, e.g., limiting the integration bounds in (5.32) to the spectral support of the respective wavelength channel, cf. the integration islands in Figure 5.4 or [CBCJ14, Fig. 4]. More details on some common approximations can be found in, e.g., [PJ17, BDS$^+$20].

The energies computed from the *Discrete-Time GN-model*, see (5.31) and Figure 1.1 (bottom), are in general not equivalent to the conventional GN-model, since the latter takes the squared magnitude *before* aliasing the kernel to the Nyquist interval, while the operations are reversed for the former Discrete-Time GN-model. The deviation between the two models depends on the pulse shapes of the involved wavelength channels, in particular, on the spectral support outside the Nyquist interval. Both models obtain the *same* result for, e.g., RRC pulses with roll-off factor $\rho = 0$. Then, the spectral support outside the Nyquist interval is zero, assuming

[7] To account for all SCI and XCI contributions in (5.32) summation over all ν is required—the GN/EGN-model in its standard form also includes MCI.

Algorithm 1: REG-PERT-FD for the SCI contribution

$\boldsymbol{a}_\lambda[k] = \mathrm{overlapSaveSplit}(\langle\, \boldsymbol{a}[k]\,\rangle, M, K)$
$k, \mu, \mu_1, \mu_2 \in \{\, 0, 1, \dots, M-1 \,\}$
$H_\rho[\mu_1, \mu_2, \mu_3] = H_\rho[\boldsymbol{\mu}] = H_\rho(\mathrm{e}^{\mathrm{j}\frac{2\pi}{M}\boldsymbol{\mu}})$
forall λ **do**
  $\boldsymbol{A}_\lambda[\mu] = \mathrm{DFT}\{\, \boldsymbol{a}_\lambda[k]\,\}$
  forall μ **do**
    $\mu_3 = \mathrm{mod}_M(\mu - \mu_1 + \mu_2)$
    $\Delta\boldsymbol{A}_\lambda^{\mathrm{SCI}}[\mu] = -\mathrm{j}\frac{\phi_{\mathrm{NL},\rho}}{M^2}\sum_{\mu_1,\mu_2} \boldsymbol{A}_\lambda[\mu_1]\boldsymbol{A}_\lambda^{\mathsf{H}}[\mu_2]\boldsymbol{A}_\lambda[\mu_3]H_\rho[\mu_1,\mu_2,\mu_3]$
    $\boldsymbol{Y}_\lambda^{\mathrm{PERT}}[\mu] = \boldsymbol{A}_\lambda[\mu] + \Delta\boldsymbol{A}_\lambda^{\mathrm{SCI}}[\mu]$
  end
  $\boldsymbol{y}_\lambda^{\mathrm{PERT}}[k] = \mathrm{DFT}^{-1}\{\, \boldsymbol{Y}_\lambda^{\mathrm{PERT}}[\mu]\,\}$
end
$\langle \boldsymbol{y}^{\mathrm{PERT}}[k]\rangle = \mathrm{overlapSaveAppend}(\boldsymbol{y}_\lambda^{\mathrm{PERT}}[k], M, K)$

the matched filter on the receive-side, and no aliasing of frequency components occurs due to T-spaced sampling. This also facilitates the connection between the GN-model and the time-domain perturbation models [ME12, DFMS16]. Due to Parseval's theorem the kernel energies in time-domain (i.e., the results formerly obtained by the time-domain perturbation models) are identical to the kernel energies in frequency-domain when the *aliased* frequency-domain kernel in (5.31) is considered.

5.1.2 Discrete-Frequency Algorithmic Implementation

The algorithmic implementation of the time-domain model from (5.15) and (5.19) is straightforward as far as the kernel coefficients are present. Here, we present an algorithmic implementation of the end-to-end model in $1/T$-periodic frequency domain from (5.14) and (5.18), which requires discretization in frequency domain. To that end, the algorithm is exemplarily derived for intra-channel (i.e., SCI) contributions corresponding to the continuous-frequency relation in (5.14), and shown at the bottom of Figure 5.2.

In order to realize the frequency-domain processing, the periodic spectrum of the transmit sequence $\boldsymbol{A}(\mathrm{e}^{\mathrm{j}\omega T})$ and the frequency-domain kernel $H_\rho(\mathrm{e}^{\mathrm{j}\omega T})$ are *discretized*, i.e., *sampled* in the transform domain. Then, the (point-wise) multiplication in frequency-domain results in a *cyclic* convolution in time domain, and we have to resort to block-wise processing using, e.g., the *overlap-and-save* method [Shy92].

Algorithm 1 realizes the regular perturbation (REG-PERT) procedure in $1/T$-periodic *discrete* frequency domain (FD). Here, the overlap-save algorithm is used to split the sequence $\langle\, \boldsymbol{a}[k]\,\rangle$ into overlapping blocks $\boldsymbol{a}_\lambda[k] \circ\!\!-\!\!\bullet\, \boldsymbol{A}_\lambda[\mu]$ of size M enumerated by the subindex $\lambda \in \mathbb{N}$. The block size is equal to the size of the DFT[8] and the overlap between successive blocks is K.

[8]The one-dimensional DFT is performed on each vector component of $\boldsymbol{a}_\lambda[k]$ and always relates the *whole* blocks of length M.

The *aliased* frequency-domain kernel is discretized to obtain the coefficients

$$H_{\mathsf{\rho}}[\mu_1, \mu_2, \mu_3] = H_{\mathsf{\rho}}[\boldsymbol{\mu}] \stackrel{\text{def}}{=} H_{\mathsf{\rho}}(\mathrm{e}^{\mathrm{j}\frac{2\pi}{M}\boldsymbol{\mu}}) \tag{5.33}$$

where M is the number of discrete-frequency samples per dimension[9].

The time- and frequency-domain picture of the *regular* perturbation approach are equivalent due to the DTFT in (5.14), (5.15) which interrelates both representations. Algorithm 1 represents a practical realization in discrete-frequency which produces the same (numerical) results as the discrete-time model as long as M and K are chosen sufficiently large for a given system scenario. Accordingly, the number of (frequency) samples M per $1/T$-cycle (and per dimension) is critical as discretization in one domain produces aliasing (and hence periodicity) in the other domain, i.e., here, the nonlinear impulse response obtained by $h_{\mathsf{\rho}}[\boldsymbol{\kappa}] = \mathrm{DFT}^{-1}\{\, H_{\mathsf{\rho}}[\boldsymbol{\mu}] \,\}$ has potential cyclic artifacts if M is chosen too small. E.g., in Figure 5.4 to 5.8, one can expect that the required number of samples M (i.e., the required frequency resolution) for the inter-channel kernel $H_\nu(\mathrm{e}^{\mathrm{j}\boldsymbol{\omega}T})$ is higher than for the intra-channel kernel $H_{\mathsf{\rho}}(\mathrm{e}^{\mathrm{j}\boldsymbol{\omega}T})$ as the *gradient* of the kernel takes higher values for $H_\nu(\mathrm{e}^{\mathrm{j}\boldsymbol{\omega}T})$. See, e.g., the ω_1-ω_2-diagonal in Figure 5.4 and 5.6 which drops off significantly faster to the edges for the inter-channel kernel compared to the intra-channel kernel. In time domain, this is reflected by an impulse response $h_\nu[\boldsymbol{\kappa}]$ with an increased temporal extent compared to $h_{\mathsf{\rho}}[\boldsymbol{\kappa}]$.

The discrete-frequency indices μ_1 and μ_2 are elements of the set

$$\mathbb{M} = \{\, 0, 1, \ldots, M-1 \,\}\,, \tag{5.34}$$

whereas μ_3 must be (modulo) *reduced*, cf. Line 7, to the same number set. This is due to the frequency matching constraint in (4.22). Due to the periodicity of both the discrete spectrum $\boldsymbol{A}_\lambda[\mu]$ and the kernel $H_{\mathsf{\rho}}[\boldsymbol{\mu}]$, the frequency index μ_3 is *folded* back into the Nyquist interval just like the modulo reduction in (5.17) for the continuous-frequency case.

Line 8 of the algorithm realizes Equation (5.14) where the (double) sum is performed over all μ_1, $\mu_2 \in \mathbb{M}$. After frequency-domain processing the blocks of *perturbed* receive symbols $\boldsymbol{Y}_\lambda^{\mathrm{PERT}}[\mu] \bullet\!\!-\!\!\circ\, \boldsymbol{y}_\lambda^{\mathrm{PERT}}[k]$ are transformed back to time domain where the $M-K$ desired output symbols of each block are *appended* to obtain the perturbed sequence $\langle\, \boldsymbol{y}^{\mathrm{PERT}}[k] \,\rangle$.

The number of coefficients can be controlled by pruning the kernel, similar to techniques already applied to oversampled VSTF models [GP13].

In terms of computational efficiency a frequency-domain implementation *can* be superior to the time-domain implementation, in particular, for cases where the number of nonlinear interacting pulses is large. This is typically the case if the system memory is large, i.e., for large map strengths $\mathcal{S}_{T,\mathsf{\rho}}$ or $\mathcal{S}_{T,\nu}$, or large relative frequency offsets $\Delta\omega_\nu$, or pulse shapes $h_{\mathrm{T}}(t)$ that extend over multiple symbol durations, e.g., a RRC shape with small roll-off factor ρ. Then, the number of coefficients of the time-domain kernel $h_\nu[\boldsymbol{\kappa}]$ exceeding a relevant energy level grows very rapidly leading to a large number of multiplications and summations. Vice-versa, we can conclude from Figure 4.8 and Figure 5.3 that for increasing system memory, the

[9]Note, that the frequency discretization of the kernel must not necessarily coincide with the transformation length M. Also, not all dimensions must be discretized with the same number of samples M.

energy of the kernel coefficients is confined in a smaller *volume* within the Nyquist cube, i.e., more coefficients can be pruned. This is in analogy with linear systems, where a large-memory system is represented by a narrow-banded transfer function. Moreover, the frequency-domain picture comprises only a double sum per frequency index μ instead of a triple sum for each k in the time-domain model—this is again in analogy with linear systems where time-domain convolution is dual to frequency-domain point-wise multiplication.

A thorough complexity analysis is, however, beyond the scope of this work, as it heavily depends on the specific application and system scenario in mind. Section 5.3 will hence focus on the validity and accuracy of the proposed model—deliberately using a very low pruning level of the coefficients to provide a benchmark performance of the discussed schemes. To that end, Section 5.3 will compare the regular discrete-time and -frequency model to the reference channel model implemented via the SSFM.

In the next section, the regular model is extended to a combined regular-logarithmic model where a subset of the perturbations are considered as multiplicative, i.e., perturbations that cause a rotation in *phase* or in the *state of polarization* (SOP), see Appendix A.1.

5.2 The First-Order Regular-Logarithmic Solution

It was already noted in [XBP02] that the (first-order) *regular* VSTF approach (or the equivalent RP method) in (4.1)–(4.2) reveals an energy-divergence problem if the optical launch power P_ν is too high—or more precisely if the nonlinear phase shift $\phi_{\mathrm{NL},\nu}$ is too large.

This issue was first addressed in the early 2000s for the analog baseband end-to-end RP solutions [XBP01, VSB02] and years later revived in the context of intra-channel fiber nonlinearity mitigation. In [VSB02, Sec. VI], the RP method is derived in a reference system rotated by the time-invariant nonlinear phase, called enhanced regular perturbation (eRP) method. It was shown in [VSB02, SB13, SB15], that the eRP model provides a significant improvement over pure RP models. A similar correction formula was proposed for VSTF methods in [XBP02]. In [FS05, SF12, SFP13], a LP model is derived which is exact in the limit of zero-dispersion links. On the other hand in DU links, as pointed out in [SB13], the LP method yields a log-normal distribution of the nonlinear distortion which is inconsistent with observations from simulations and experiments.

In the *additive-multiplicative* (A-M) model derived in [FDT$^+$12, TZF$^+$14], it turned out that a certain subset of symbol combinations in the time-domain RP model deterministically creates a perturbation oriented into the $-$j-direction from the transmit symbol $\boldsymbol{a}[k]$. Similarly, in the pulse-collision picture [DFMS13, DFMS14, DFMS16] a subset of *degenerate*[10] cross-channel pulse collisions were properly associated to distortions exhibiting a *multiplicative* nature. In the same series of contributions, these subsets of degenerate distortions were first termed *two-* and *three-pulse collisions*, i.e., symbol combinations $\boldsymbol{\kappa} \in \mathbb{Z}^3$ in (5.19) with $\kappa_3 = 0$ in our terminology. While the pulse collision picture covers mainly cross-channel effects, we will extend the discussion on separating additive and multiplicative terms also to intra-channel

[10] In the sense that not all four interacting pulses are distinct.

effects.

It is through using a first-order RP approach that a pure phase rotation is not modeled well. It is approximated only by the first two terms of the exponents Taylor series, i.e.,

$$\exp(-\mathrm{j}\phi) \approx 1 - \mathrm{j}\phi\,. \tag{5.35}$$

While multiplication with $\exp(-\mathrm{j}\phi)$ is an energy conserving transformation (i.e., the norm is invariant under phase rotation), the RP approximation is *not* energy conserving (cf. also the discussion in [TDY$^+$11, Sec. II B.] and [DFMS16, Sec. VIII]). In the context of optical transmission, already a trivial (time-invariant) common phase rotation due to nonlinear interaction is not well modeled by the RP method. E.g., using the discrete-time RP ansatz from (4.5) under a time-invariant rotation ϕ, we find

$$\boldsymbol{y}[k] = \boldsymbol{a}[k]\exp(-\mathrm{j}\phi) \approx \boldsymbol{a}[k] - \mathrm{j}\phi\,\boldsymbol{a}[k]\,, \tag{5.36}$$

i.e., the (additive) perturbation $\Delta\boldsymbol{a}[k]$ depends directly on the transmit symbol $\boldsymbol{a}[k]$, is oriented into the $-\mathrm{j}$-direction from the transmit symbol $\boldsymbol{a}[k]$ (i.e., in a 90° clock-wise rotation), and scaled by the scalar factor ϕ. The same argument can be made for *polarization rotations* in $\mathbb{C}^{2\times 2}$, see Appendix A.1. The first-order approximation of a (time-invariant) polarization rotation results in

$$\boldsymbol{y}[k] = \boldsymbol{a}[k]\exp(-\mathrm{j}\vec{\boldsymbol{s}}\cdot\vec{\boldsymbol{\sigma}}) \approx \boldsymbol{a}[k] - \mathrm{j}(\vec{\boldsymbol{s}}\cdot\vec{\boldsymbol{\sigma}})\,\boldsymbol{a}[k]\,. \tag{5.37}$$

where $\exp(\cdot)$ denotes the *matrix exponential* from (2.11). We use the common notation $\vec{\boldsymbol{s}}\cdot\vec{\boldsymbol{\sigma}} \in \mathbb{C}^{2\times 2}$ from (A.39) to express Hermitian 2×2 matrices.

Comparing $-\mathrm{j}\phi\boldsymbol{a}[k]$ from (5.36) to the result for the intra-channel perturbation $\Delta\boldsymbol{a}^{\mathrm{SCI}}[k]$ in (5.15), we may identify certain conditions (depending on the selection of $\boldsymbol{\kappa} = [\kappa_1, \kappa_2, \kappa_3]^\mathsf{T}$ and the properties of the kernel $h_\rho[\boldsymbol{\kappa}]$) for which $\Delta\boldsymbol{a}^{\mathrm{SCI}}[k]$ is deterministically proportional to $-\mathrm{j}\boldsymbol{a}[k]$ as in (5.36). The same analysis can be done for inter-channel perturbations $\boldsymbol{a}^{\mathrm{XCI}}[k]$, for polarization rotations as in (5.37), and for the ansatz in frequency domain (4.6).

The strategy in what follows is to explicitly implement those cases, see above, by a multiplication via the exponent $\exp(\cdot)$, and to exclude them from the additive term. The content and much of the terminology in the next section is based on the pulse-collision picture [DFMS16]. We add to the discussion also a time-domain description of intra-channel XPolM, and extend the approach in the following section to the derived frequency-domain view of the discrete-time model.

5.2.1 Regular-Logarithmic Model in Time Domain

In the following, the original RP solution is modified such that perturbations originating from certain (degenerate) mixing products are associated with a multiplicative perturbation. Similar to [SFM09, TZF$^+$14, DFMS16], we extend the previous RP model to a combined regular-logarithmic model. It takes the general form of[11]

$$\boldsymbol{y}[k] = \exp\left(\mathrm{j}\boldsymbol{\phi}[k] + \mathrm{j}\vec{\boldsymbol{s}}[k] \cdot \vec{\boldsymbol{\sigma}}\right)\left(\boldsymbol{a}[k] + \Delta\boldsymbol{a}[k]\right) . \tag{5.38}$$

In addition to the regular, additive perturbation $\Delta\boldsymbol{a}[k]$ we now also consider a *phase* rotation by $\exp(\mathrm{j}\boldsymbol{\phi}[k])$, where $\boldsymbol{\phi}[k]$ is a diagonal matrix in $\mathbb{R}^{2\times 2}$, and a rotation in the *state of polarization* by $\exp(\mathrm{j}\vec{\boldsymbol{s}}[k] \cdot \vec{\boldsymbol{\sigma}})$, where $\vec{\boldsymbol{s}}[k] \cdot \vec{\boldsymbol{\sigma}}$ is Hermitian in $\mathbb{C}^{2\times 2}$.

All perturbative terms combine both SCI and XCI effects, i.e., the *additive* perturbation $\Delta\boldsymbol{a}[k] \in \mathbb{C}^2$ in (5.38) is the sum of SCI and XCI contributions. The *time-dependent* phase rotation is given by $\exp(\mathrm{j}\boldsymbol{\phi}[k])$ with the diagonal matrix $\boldsymbol{\phi}[k] \in \mathbb{R}^{2\times 2}$ defined as

$$\boldsymbol{\phi}[k] \stackrel{\text{def}}{=} \phi^{\text{SCI}}[k]\,\mathbf{I} + \phi^{\text{XCI}}[k]\,\mathbf{I}\,, \tag{5.39}$$

i.e., we define a *common* phase term for both polarizations originating from intra- and inter-channel effects.

The combined effect of intra- and inter-channel XPolM is expressed by the *Pauli matrix expansion* $\vec{\boldsymbol{s}}[k]\cdot\vec{\boldsymbol{\sigma}} \in \mathbb{C}^{2\times 2}$ using (A.39), with the notation adopted from [GK00] and [WBSP09]. The expansion defines a unitary rotation in Jones space of the perturbed vector $\boldsymbol{a}[k] + \Delta\boldsymbol{a}[k]$ around the *time-dependent* Stokes vector $\vec{\boldsymbol{s}}[k]$ and is explained in more detail in the subsequent paragraphs.

SCI Contribution To discuss the SCI contribution we first introduce the following symbol sets, motivated by the symmetries of the time-domain kernel $h_\rho[\boldsymbol{\kappa}]$ in (5.27)–(5.30). We define

$$\begin{aligned}
\mathcal{K}^{\text{SCI}} &= \left\{ [\kappa_1, \kappa_2, \kappa_3]^\mathsf{T} \in \mathbb{Z}^3 \mid |h_\rho[\boldsymbol{\kappa}]/h_\rho[\mathbf{0}]|^2 > \mathit{\Gamma}^{\text{SCI}} \right\} && (5.40)\\
\mathcal{K}_\phi^{①} &\stackrel{\text{def}}{=} \left\{ \mathcal{K}^{\text{SCI}} \mid \kappa_1 = 0 \land \kappa_2 \neq 0 \land \kappa_3 \neq 0 \right\} && (5.41)\\
\mathcal{K}_\phi^{③} &\stackrel{\text{def}}{=} \left\{ \mathcal{K}^{\text{SCI}} \mid \kappa_3 = 0 \land \kappa_2 \neq 0 \land \kappa_1 \neq 0 \right\} && (5.42)\\
\mathcal{K}_\phi^{\text{SCI}} &\stackrel{\text{def}}{=} \mathcal{K}_\phi^{①} \cup \mathcal{K}_\phi^{③} \cup \left\{ \boldsymbol{\kappa} = \mathbf{0} \right\} && (5.43)\\
\mathcal{K}_\Delta^{\text{SCI}} &\stackrel{\text{def}}{=} \mathcal{K}^{\text{SCI}} \setminus \mathcal{K}_\phi^{\text{SCI}}\,, && (5.44)
\end{aligned}$$

where (5.40) defines the *base* set including all possible symbol combinations that exceed a certain energy (clipping) level $\mathit{\Gamma}^{\text{SCI}}$ normalized to the energy of the *center* tap at $\boldsymbol{\kappa} = \mathbf{0}$. In (5.41), (5.42) the joint set of degenerate two- and three-pulse collisions for SCI are defined which follow directly from the kernel properties in (5.27), (5.28) for $\kappa_3 = 0$, and (5.29), (5.30) for $\kappa_1 = 0$. The set of indices for *multiplicative* distortions $\mathcal{K}_\phi^{\text{SCI}}$ in (5.43) also includes the singular case $\boldsymbol{\kappa} = \mathbf{0}$. Then, the additive set is simply the complementary set of $\mathcal{K}_\phi^{\text{SCI}}$ w.r.t. the base set $\mathcal{K}^{\text{SCI}}$.

We start with the additive perturbation from the previous section in (5.15) which now reads

$$\Delta\boldsymbol{a}^{\text{SCI}}[k] = -\mathrm{j}\phi_{\text{NL},\rho} \sum_{\mathcal{K}_\Delta^{\text{SCI}}} \boldsymbol{a}[k-\kappa_1]\boldsymbol{a}^\mathsf{H}[k-\kappa_2]\boldsymbol{a}[k-\kappa_3]h_\rho[\boldsymbol{\kappa}], \tag{5.45}$$

[11]Note, that the order, in which the additive and multiplicative perturbation is applied, matters. We chose the same order as in the original *additive-multiplicative (A-M) model* from [TZF+14], but we have no proof that this is the optimal order of how to combine the two operations.

where the triple sum is now restricted to the set $\mathcal{K}_{\Delta}^{\text{SCI}}$ excluding all combinations which result in a multiplicative distortion, cf. (5.44).

To calculate the common phase $\phi^{\text{SCI}}[k]$ and the intra-channel Stokes rotation vector $\vec{\boldsymbol{s}}^{\,\text{SCI}}[k]$ we first analyze the expression $\boldsymbol{a}[k-\kappa_1]\boldsymbol{a}^{\mathsf{H}}[k-\kappa_2]\boldsymbol{a}[k-\kappa_3]$ from the original equation in (5.15). For the set $\mathcal{K}_{\phi}^{①}$ with $\kappa_1 = 0$ the triple product factors into the respective transmit symbol $\boldsymbol{a}[k]$ and a scalar value $\boldsymbol{a}^{\mathsf{H}}[k-\kappa_2]\boldsymbol{a}[k-\kappa_3]$. After multiplication with $h_{\rho}[0,\kappa_2,\kappa_3]$ and summation of all $\boldsymbol{\kappa} \in \mathcal{K}_{\phi}^{①}$ the perturbation is strictly imaginary-valued, cf. symmetry properties in (5.29), (5.30).

On the other hand, for $\mathcal{K}_{\phi}^{③}$ with $\kappa_3 = 0$ we have to rearrange the triple product using the matrix expansion from (A.40) to factor the expression accordingly as[12]

$$\boldsymbol{a}\boldsymbol{a}^{\mathsf{H}}\boldsymbol{a} = \frac{1}{2}\left(\boldsymbol{a}^{\mathsf{H}}\boldsymbol{a}\,\mathbf{I} + (\boldsymbol{a}^{\mathsf{H}}\vec{\boldsymbol{\sigma}}\boldsymbol{a})\cdot\vec{\boldsymbol{\sigma}}\right)\boldsymbol{a}\,, \tag{5.46}$$

where $\boldsymbol{a}$ is an element of the discrete-time transmit symbol sequence $\langle\, \boldsymbol{a}[k]\,\rangle$. The first term $\boldsymbol{a}^{\mathsf{H}}\boldsymbol{a}\,\mathbf{I}$ contributes to a common phase rotation, whereas the second term $(\boldsymbol{a}^{\mathsf{H}}\vec{\boldsymbol{\sigma}}\boldsymbol{a})\cdot\vec{\boldsymbol{\sigma}} \in \mathbb{C}^{2\times 2}$ is a traceless and Hermitian matrix such that $\exp(\mathrm{j}(\boldsymbol{a}^{\mathsf{H}}\vec{\boldsymbol{\sigma}}\boldsymbol{a})\cdot\vec{\boldsymbol{\sigma}})$ is a unitary polarization rotation[13].

The multiplicative perturbation $\exp(\mathrm{j}\phi^{\text{SCI}}[k])$ with $\phi^{\text{SCI}}[k] \in \mathbb{R}$ is then given by

$$\begin{aligned}
\phi^{\text{SCI}}[k] = &- \phi_{\text{NL},\rho} \sum_{\mathcal{K}_{\phi}^{①}} \boldsymbol{a}^{\mathsf{H}}[k-\kappa_2]\boldsymbol{a}[k-\kappa_3]\, h_{\rho}[0,\kappa_2,\kappa_3] \\
&- \frac{1}{2}\phi_{\text{NL},\rho} \sum_{\mathcal{K}_{\phi}^{③}} \boldsymbol{a}^{\mathsf{H}}[k-\kappa_2]\boldsymbol{a}[k-\kappa_1]\, h_{\rho}[\kappa_1,\kappa_2,0] \\
&- \phi_{\text{NL},\rho}\, \|\boldsymbol{a}[k]\|^2\, h_{\rho}[\mathbf{0}] \qquad (5.47)\\
= &- \frac{3}{2}\phi_{\text{NL},\rho} \sum_{\mathcal{K}_{\phi}^{③}} \boldsymbol{a}^{\mathsf{H}}[k-\kappa_2]\boldsymbol{a}[k-\kappa_1]\, h_{\rho}[\kappa_1,\kappa_2,0] \\
&- \phi_{\text{NL},\rho}\, \|\boldsymbol{a}[k]\|^2\, h_{\rho}[\mathbf{0}]\,, \qquad (5.48)
\end{aligned}$$

where we used $h_{\rho}[0,\kappa_2,\kappa_3] = h_{\rho}[\kappa_1,\kappa_2,0]$ from (5.29). Given a wide-sense stationary transmit sequence $\langle\, \boldsymbol{a}[k]\,\rangle$, the induced nonlinear phase shift has a *time-average* value $\bar{\phi}^{\text{SCI}}$, around which the instantaneous phase $\phi^{\text{SCI}}[k]$ may fluctuate (cf. also [SB15]).

The instantaneous rotation of the SOP due to the expression $\exp(\mathrm{j}\vec{\boldsymbol{s}}^{\,\text{SCI}}[k]\cdot\vec{\boldsymbol{\sigma}}) \in \mathbb{C}^{2\times 2}$ causes intra-channel XPolM [MM12]. It is given by

$$\begin{aligned}
\vec{\boldsymbol{s}}^{\,\text{SCI}}[k]\cdot\vec{\boldsymbol{\sigma}} = -\frac{1}{2}\phi_{\text{NL},\rho} \sum_{\mathcal{K}_{\phi}^{③}} \Big(& 2\,\boldsymbol{a}[k-\kappa_1]\boldsymbol{a}^{\mathsf{H}}[k-\kappa_2] \\
&- \boldsymbol{a}^{\mathsf{H}}[k-\kappa_2]\boldsymbol{a}[k-\kappa_1]\mathbf{I}\Big)\, h_{\rho}[\kappa_1,\kappa_2,0]\,,
\end{aligned} \tag{5.49}$$

where we made use of the relation in (A.33), (A.40) to rewrite the second summand on the right-hand side of (5.46) as

$$(\boldsymbol{a}^{\mathsf{H}}\vec{\boldsymbol{\sigma}}\boldsymbol{a})\cdot\vec{\boldsymbol{\sigma}} = 2\boldsymbol{a}\boldsymbol{a}^{\mathsf{H}} - \boldsymbol{a}^{\mathsf{H}}\boldsymbol{a}\mathbf{I}\,. \tag{5.50}$$

[12]Multiplication with $h_{\rho}[\boldsymbol{\kappa}]$ and summation over $\boldsymbol{\kappa} \in \mathcal{K}_{\phi}^{\text{SCI}}$ are implied.

[13]Since the Pauli expansion $\vec{\boldsymbol{u}}\cdot\vec{\boldsymbol{\sigma}}$ in (A.39) is Hermitian, the expression $\exp(\mathrm{j}\,\vec{\boldsymbol{u}}\cdot\vec{\boldsymbol{\sigma}})$ is unitary.

The physical meaning of the transformation described in (5.49) is as follows: The perturbed transmit vector $(\boldsymbol{a}[k] + \Delta\boldsymbol{a}[k])$ in (5.38) is transformed into the polarization eigenstate $\vec{\boldsymbol{s}}^{\,\mathrm{SCI}}[k]$ (i.e., into the basis defined by the eigenvectors of $\vec{\boldsymbol{s}}^{\,\mathrm{SCI}}[k] \cdot \vec{\boldsymbol{\sigma}}$). There, both vector components receive equal but opposite phase shifts and the result is transformed back to the x/y-basis of the transmit vector. In Stokes space, the operation can be understood as a precession of $(\vec{\boldsymbol{a}}[k] + \Delta\vec{\boldsymbol{a}}[k])$ around the Stokes vector $\vec{\boldsymbol{s}}^{\,\mathrm{SCI}}[k]$ by an angle equal to its length $\left\|\vec{\boldsymbol{s}}^{\,\mathrm{SCI}}[k]\right\|$, see, e.g., Example A.1. The intra-channel Stokes vector $\vec{\boldsymbol{s}}^{\,\mathrm{SCI}}[k]$ depends via the nonlinear kernel $h_\rho[\boldsymbol{\kappa}]$ on the transmit symbols within the neighborhood of $\boldsymbol{a}[k]$, i.e., depending on the memory $\mathcal{S}_{T,\rho}$ of the nonlinear interaction. Similar to the nonlinear phase shift—for a wide-sense stationary input sequence—the Stokes vector $\vec{\boldsymbol{s}}^{\,\mathrm{SCI}}[k]$ has a time-constant average value around which it fluctuates over time.

XCI Contribution The same methodology is now applied to cross-channel effects. The symbol set definitions for XCI follow from the considerations in the previous section. We find

$$\mathcal{K}_\nu^{\mathrm{XCI}} = \{\, [\kappa_1, \kappa_2, \kappa_3]^\mathsf{T} \in \mathbb{Z}^3 \mid |h_\nu[\boldsymbol{\kappa}]/h_\nu[\mathbf{0}]|^2 > \Gamma_\nu^{\mathrm{XCI}} \,\} \tag{5.51}$$

$$\begin{aligned} \mathcal{K}_{\phi,\nu}^{\mathrm{XCI}} &\overset{\text{def}}{=} \{\, \mathcal{K}_\nu^{\mathrm{XCI}} \mid \kappa_3 = 0 \wedge \kappa_2 \neq 0 \wedge \kappa_1 \neq 0 \,\} \\ &\quad \cup \{\, \boldsymbol{\kappa} = \mathbf{0} \,\} \end{aligned} \tag{5.52}$$

$$\mathcal{K}_{\Delta,\nu}^{\mathrm{XCI}} \overset{\text{def}}{=} \mathcal{K}_\nu^{\mathrm{XCI}} \setminus \mathcal{K}_{\phi,\nu}^{\mathrm{XCI}}\,, \tag{5.53}$$

where the subscript ν indicates the channel number of the respective interfering channel. For $\mathcal{K}_{\phi,\nu}^{\mathrm{XCI}}$, only the degenerate case $\kappa_3 = 0$ has to be considered due to the kernel properties of $h_\nu[\kappa_1, \kappa_2, 0]$ in (5.27), (5.28). Similar to (5.46), the expression $\boldsymbol{b}\boldsymbol{b}^\mathsf{H} + \boldsymbol{b}^\mathsf{H}\boldsymbol{b}\,\mathbf{I}$ from (5.19) is rearranged to obtain

$$\underbrace{\frac{3}{2}\begin{bmatrix} b_\mathsf{x}b_\mathsf{x}^* + b_\mathsf{y}b_\mathsf{y}^* & 0 \\ 0 & b_\mathsf{y}b_\mathsf{y}^* + b_\mathsf{x}b_\mathsf{x}^* \end{bmatrix}}_{\boldsymbol{b}^\mathsf{H}\boldsymbol{b}\,\mathbf{I}} + \underbrace{\frac{1}{2}\begin{bmatrix} b_\mathsf{x}b_\mathsf{x}^* - b_\mathsf{y}b_\mathsf{y}^* & 2b_\mathsf{x}b_\mathsf{y}^* \\ 2b_\mathsf{y}b_\mathsf{x}^* & b_\mathsf{y}b_\mathsf{y}^* - b_\mathsf{x}b_\mathsf{x}^* \end{bmatrix}}_{2\,\boldsymbol{b}\boldsymbol{b}^\mathsf{H} - \boldsymbol{b}^\mathsf{H}\boldsymbol{b}\,\mathbf{I} \,=\, (\boldsymbol{b}^\mathsf{H}\vec{\boldsymbol{\sigma}}\boldsymbol{b})\cdot\vec{\boldsymbol{\sigma}}}, \tag{5.54}$$

where the argument and subscript ν is omitted for concise notation. The multiplicative cross-channel contribution is again split into a common phase shift in both polarizations, i.e., the first summand in (5.54), and an equal but opposite phase shift in the basis given by the instantaneous Stokes vector of the ν^{th} interferer, i.e., the second summand in (5.54).

We define the total, common phase shift due to cross-channel interference as

$$\phi^{\mathrm{XCI}}[k] = -\sum_{\nu\neq\rho} \frac{3}{2}\phi_{\mathrm{NL},\nu} \sum_{\mathcal{K}_{\phi,\nu}^{\mathrm{XCI}}} \boldsymbol{b}_\nu^\mathsf{H}[k-\kappa_1]\boldsymbol{b}_\nu[k-\kappa_2]\; h_\nu[\kappa_1, \kappa_2, 0]\,, \tag{5.55}$$

which depends on the sum over all interfering channels and the sum of $\boldsymbol{b}_\nu^\mathsf{H}\boldsymbol{b}_\nu$ over $[\kappa_1, \kappa_2]^\mathsf{T}$. The effective, instantaneous cross-channel Stokes vector $\vec{\boldsymbol{s}}^{\mathrm{XCI}}[k]$ is given by

$$\begin{aligned} \vec{\boldsymbol{s}}^{\,\mathrm{XCI}}[k] \cdot \vec{\boldsymbol{\sigma}} = -\sum_{\nu\neq\rho} \frac{1}{2}\phi_{\mathrm{NL},\nu} \sum_{\mathcal{K}_{\phi,\nu}^{\mathrm{XCI}}} \Big(& 2\,\boldsymbol{b}_\nu[k-\kappa_1]\boldsymbol{b}_\nu^\mathsf{H}[k-\kappa_2] \\ & - \boldsymbol{b}_\nu^\mathsf{H}[k-\kappa_2]\boldsymbol{b}_\nu[k-\kappa_1]\mathbf{I}\Big)\, h_\nu[\kappa_1, \kappa_2, 0]\,. \end{aligned} \tag{5.56}$$

Note, that the expressions in (5.55), (5.56) include both contributions from two- and three pulse collisions (cf. [DFMS16, Eq. (10)–(13)]).

Energy of Coefficients in Discrete-Time Domain The energy of the kernel coefficients is defined for the subsets given in (5.40)–(5.44). We find for the different symbol sets

$$E_h^{\mathrm{SCI}} \stackrel{\mathrm{def}}{=} \sum_{\mathcal{K}^{\mathrm{SCI}}} |h_\rho[\boldsymbol{\kappa}]|^2 \tag{5.57}$$

$$E_{h,\Delta}^{\mathrm{SCI}} \stackrel{\mathrm{def}}{=} \sum_{\mathcal{K}_\Delta^{\mathrm{SCI}}} |h_\rho[\boldsymbol{\kappa}]|^2 \tag{5.58}$$

$$E_{h,\phi}^{\mathrm{SCI}} \stackrel{\mathrm{def}}{=} \sum_{\mathcal{K}_\phi^{\mathrm{SCI}}} |h_\rho[\boldsymbol{\kappa}]|^2 \,, \tag{5.59}$$

with the clipping factor Γ^{SCI} in (5.40) equal to zero. The energy for cross-channel effects is defined accordingly with the sets from (5.51)–(5.53). Since the subsets for additive and multiplicative effects are always disjoint we have $E_h^{\mathrm{SCI}} = E_{h,\Delta}^{\mathrm{SCI}} + E_{h,\phi}^{\mathrm{SCI}}$.

5.2.2 Regular-Logarithmic Model in Frequency Domain

The frequency-domain model is now modified such that certain contributions will be associated with multiplicative distortions. Despite the multiplicative nature of phase and polarization rotations, *time-invariant* rotations, e.g., the time-average phase rotation $\bar{\phi} = \bar{\phi}^{\mathrm{SCI}} + \bar{\phi}^{\mathrm{XCI}}$, can be straightforwardly incorporated into the frequency-domain model as they are both treated as constant pre-factors in the time- and frequency-domain representation. We will see in the next section that this already leads to significantly improved (numerical) results compared to the regular model. Note that, in contrast to the regular models, the regular-logarithmic model in time and frequency are no longer equivalent. The two approaches have a similar form, compare with (5.38), but lead to different results.

The general form of the combined regular-logarithmic model in frequency is given by

$$\boldsymbol{Y}(\mathrm{e}^{\mathrm{j}\omega T}) = \exp\left(\mathrm{j}\bar{\boldsymbol{\phi}} + \mathrm{j}\vec{\boldsymbol{S}} \cdot \vec{\boldsymbol{\sigma}}\right) \left(\boldsymbol{A}(\mathrm{e}^{\mathrm{j}\omega T}) + \Delta\boldsymbol{A}(\mathrm{e}^{\mathrm{j}\omega T})\right) \,, \tag{5.60}$$

where the phase- and polarization-term take on a *frequency*-invariant value, i.e., independent of $\mathrm{e}^{\mathrm{j}\omega T}$ (indicated here by the lack of argument for $\bar{\boldsymbol{\phi}}$ and $\vec{\boldsymbol{S}}$). Following the same terminology as before, we introduce the *average* multiplicative perturbation of the common phase term

$$\bar{\boldsymbol{\phi}} \stackrel{\mathrm{def}}{=} \bar{\phi}^{\mathrm{SCI}}\,\mathbf{I} + \bar{\phi}^{\mathrm{XCI}}\,\mathbf{I} \,, \tag{5.61}$$

as the sum of the intra-channel contribution $\bar{\phi}^{\mathrm{SCI}} \in \mathbb{R}$ and the inter-channel contribution $\bar{\phi}^{\mathrm{XCI}} \in \mathbb{R}$. Similarly, for the *average* polarization rotation we have

$$\vec{\boldsymbol{S}} \cdot \vec{\boldsymbol{\sigma}} \stackrel{\mathrm{def}}{=} \vec{\boldsymbol{S}}^{\mathrm{SCI}} \cdot \vec{\boldsymbol{\sigma}} + \vec{\boldsymbol{S}}^{\mathrm{XCI}} \cdot \vec{\boldsymbol{\sigma}} \,, \tag{5.62}$$

where $\vec{\boldsymbol{S}} \cdot \vec{\boldsymbol{\sigma}}$ is again Hermitian and traceless.

SCI Contribution The two degenerate frequency conditions in (5.24) are used in the expression (5.14) to obtain the *average*, intra-channel phase rotation. To that end, the triple product

Algorithm 2: REGLOG-PERT-FD for the SCI contribution

$\boldsymbol{a}_\lambda[k] = \text{overlapSaveSplit}(\langle\, \boldsymbol{a}[k]\,\rangle, M, K)$

$k, \mu, \mu_1, \mu_2 \in \{\, 0, 1, \ldots, M-1\,\}$

$H_\rho[\mu_1, \mu_2, \mu_3] = H_\rho[\boldsymbol{\mu}] = H_\rho(\mathrm{e}^{\mathrm{j}\frac{2\pi}{M}\boldsymbol{\mu}})$

forall λ **do**

  $\boldsymbol{A}_\lambda[\mu] = \text{DFT}\{\, \boldsymbol{a}_\lambda[k]\,\}$

  $\bar{\phi}_\lambda^{\text{SCI}} = -\frac{3}{2}\frac{\phi_{\text{NL},\rho}}{M^2}\ \sum_\mu \|\boldsymbol{A}_\lambda[\mu]\|^2$

  $\vec{\boldsymbol{S}}_\lambda^{\text{SCI}} \cdot \vec{\boldsymbol{\sigma}} = -\frac{1}{2}\frac{\phi_{\text{NL},\rho}}{M^2} \sum_\mu 2\boldsymbol{A}_\lambda[\mu]\boldsymbol{A}_\lambda^{\mathsf{H}}[\mu] - \|\boldsymbol{A}_\lambda[\mu]\|^2\,\mathbf{I}$

  forall μ **do**

    $\mu_3 = \text{mod}_M(\mu - \mu_1 + \mu_2)$

    $\mathcal{U} = \{\, [\mu_1, \mu_2]^{\mathsf{T}} \mid \mu_2 \neq \mu_1 \wedge \mu_2 \neq \mu_3\,\}$

    $\Delta\boldsymbol{A}_\lambda^{\text{SCI}}[\mu] = -\mathrm{j}\frac{\phi_{\text{NL},\rho}}{M^2}\sum_{\mathcal{U}} \boldsymbol{A}_\lambda[\mu_1]\boldsymbol{A}_\lambda^{\mathsf{H}}[\mu_2]\boldsymbol{A}_\lambda[\mu_3]H_\rho[\mu_1, \mu_2, \mu_3]$

    $\boldsymbol{Y}_\lambda^{\text{PERT}}[\mu] = \exp(\mathrm{j}\bar{\phi}_\lambda^{\text{SCI}}\mathbf{I} + \mathrm{j}\vec{\boldsymbol{S}}_\lambda^{\text{SCI}} \cdot \vec{\boldsymbol{\sigma}})(\boldsymbol{A}_\lambda[\mu] + \Delta\boldsymbol{A}_\lambda^{\text{SCI}}[\mu])$

  end

  $\boldsymbol{y}_\lambda^{\text{PERT}}[k] = \text{DFT}^{-1}\{\, \boldsymbol{Y}_\lambda^{\text{PERT}}[\mu]\,\}$

end

$\langle \boldsymbol{y}^{\text{PERT}}[k]\rangle = \text{overlapSaveAppend}(\boldsymbol{y}_\lambda^{\text{PERT}}[k], M, K)$

$\boldsymbol{A}\boldsymbol{A}^{\mathsf{H}}\boldsymbol{A}$ in (5.14) is rearranged similar to (5.46). First, the general *frequency-dependent* expression is given by

$$\phi^{\text{SCI}}(\mathrm{e}^{\mathrm{j}\omega T}) = -\ \phi_{\text{NL},\rho}\frac{T}{(2\pi)^2}\int_{\mathbb{T}} \left\|\boldsymbol{A}(\mathrm{e}^{\mathrm{j}\omega_2 T})\right\|^2 H_\rho(\mathrm{e}^{\mathrm{j}[\omega,\omega_2,\omega_2]^{\mathsf{T}}T})\,\mathrm{d}\omega_2 \\ -\frac{1}{2}\phi_{\text{NL},\rho}\frac{T}{(2\pi)^2}\int_{\mathbb{T}} \left\|\boldsymbol{A}(\mathrm{e}^{\mathrm{j}\omega_1 T})\right\|^2 H_\rho(\mathrm{e}^{\mathrm{j}[\omega_1,\omega_1,\omega]^{\mathsf{T}}T})\,\mathrm{d}\omega_1\,, \tag{5.63}$$

where the first term on the right-hand side in (5.63) corresponds to the degeneracy $\omega_2 = \omega_3 \Leftrightarrow \omega_1 = \omega$ and the second term corresponds to $\omega_2 = \omega_1 \Leftrightarrow \omega_3 = \omega$. We simplify the expression using the RRC $\rho = 0$ approximation (i.e., the kernel $H_\rho(\mathrm{e}^{\mathrm{j}\boldsymbol{\omega}T})$ on the ω_2-ω_3-diagonal and ω_1-ω_2-diagonal is equal to 1) and the symmetry property in (5.25) to obtain the average, intra-channel phase distortion

$$\bar{\phi}^{\text{SCI}} = -\ \frac{3}{2}\phi_{\text{NL},\rho}\frac{T}{(2\pi)^2}\int_{\mathbb{T}} \left\|\boldsymbol{A}(\mathrm{e}^{\mathrm{j}\omega T})\right\|^2\ \mathrm{d}\omega\,, \tag{5.64}$$

which does no longer depend on the power or dispersion profile of the transmission link (given a fixed L_{eff}).

Similarly, the average intra-channel XPolM contribution can be simplified to

$$\vec{\boldsymbol{S}}^{\text{SCI}} \cdot \vec{\boldsymbol{\sigma}} = -\frac{1}{2}\phi_{\text{NL},\rho}\frac{T}{(2\pi)^2}\int_{\mathbb{T}}\Big(2\,\boldsymbol{A}(\mathrm{e}^{\mathrm{j}\omega T})\boldsymbol{A}^{\mathsf{H}}(\mathrm{e}^{\mathrm{j}\omega T}) - \boldsymbol{A}^{\mathsf{H}}(\mathrm{e}^{\mathrm{j}\omega T})\boldsymbol{A}(\mathrm{e}^{\mathrm{j}\omega T})\mathbf{I}\Big)\,\mathrm{d}\omega\,. \tag{5.65}$$

In Algorithm 2 the required modifications to the regular perturbation model (REG-PERT) are highlighted to arrive at the regular-logarithmic perturbation model (REGLOG-PERT)—again exemplarily for the SCI contribution. Lines 6, 7 of Algorithm 2 translate (5.64), (5.65) to the discrete-frequency domain where the integral over all $\omega \in \mathbb{T}$ becomes a sum over all μ of the λ^{th} processing block. The average values, here, are always associated to the average values of the λ^{th} block. In Lines 10, 11, the double sum to obtain $\Delta \boldsymbol{A}_{\lambda}^{\text{SCI}}[\mu]$ is restricted to all combinations $\mathcal{U}$ of the discrete frequency pair $[\mu_1, \mu_2]^{\mathsf{T}}$ excluding the degenerate cases corresponding to (5.24). The perturbed receive vector $\boldsymbol{Y}_{\lambda}^{\text{PERT}}[\mu]$ is then calculated according to (5.60) before it is transformed back to the discrete-time domain.

XCI Contribution The cross-channel contributions follow from the considerations in the previous sections, and we obtain for the degenerate case in (5.23) the total, average XCI phase rotation

$$\bar{\phi}^{\text{XCI}} = -\sum_{\nu \neq \rho} \frac{3}{2} \phi_{\text{NL},\nu} \frac{T}{(2\pi)^2} \int_{\mathbb{T}} \left\| \boldsymbol{B}_{\nu}(\mathrm{e}^{\mathrm{j}\omega T}) \right\|^2 \, \mathrm{d}\omega \,, \tag{5.66}$$

and analogously for the total, average XCI Stokes vector we find

$$\begin{aligned} \vec{\boldsymbol{S}}^{\text{XCI}} \cdot \vec{\boldsymbol{\sigma}} = -\sum_{\nu \neq \rho} \frac{1}{2} \phi_{\text{NL},\nu} \frac{T}{(2\pi)^2} \int_{\mathbb{T}} \Big(& 2\, \boldsymbol{B}_{\nu}(\mathrm{e}^{\mathrm{j}\omega T}) \boldsymbol{B}_{\nu}^{\mathsf{H}}(\mathrm{e}^{\mathrm{j}\omega T}) \\ & - \boldsymbol{B}_{\nu}^{\mathsf{H}}(\mathrm{e}^{\mathrm{j}\omega T}) \boldsymbol{B}_{\nu}(\mathrm{e}^{\mathrm{j}\omega T}) \mathbf{I} \Big) \, \mathrm{d}\omega \,. \end{aligned} \tag{5.67}$$

Algorithm 3 implements the cross-channel perturbation model in $1/T$-periodic frequency domain. In contrast to the SCI implementation, see Line 10 in Algorithm 2, the double sum is now performed over the set $\mathcal{U}$ which excludes the case $\mu_2 = \mu_1$, see Line 12 in Algorithm 3, corresponding to the degenerate case (5.23).

Energy of Coefficients in Discrete-Frequency Domain With the notation of the discrete-frequency kernel from (5.33) we have according to Parseval's theorem in (5.31) the following definitions

$$E_H^{\text{SCI}} \overset{\text{def}}{=} \frac{1}{M^3} \sum_{\mathcal{U}^{\text{SCI}}} |H_{\rho}[\boldsymbol{\mu}]|^2 \tag{5.68}$$

$$E_{H,\Delta}^{\text{SCI}} \overset{\text{def}}{=} \frac{1}{M^3} \sum_{\mathcal{U}_{\Delta}^{\text{SCI}}} |H_{\rho}[\boldsymbol{\mu}]|^2 \tag{5.69}$$

$$E_{H,\phi}^{\text{SCI}} \overset{\text{def}}{=} \frac{1}{M^3} \sum_{\mathcal{U}_{\phi}^{\text{SCI}}} |H_{\rho}[\boldsymbol{\mu}]|^2 \overset{\rho=0}{\approx} \frac{2}{M}, \tag{5.70}$$

with the sets according to (5.24)

$$\mathcal{U}^{\text{SCI}} = \{\, \boldsymbol{\mu} = [\mu_1, \mu_2, \mu_3]^{\mathsf{T}} \in \{\, 0, 1, \ldots, M-1 \,\}^3 \,\} \tag{5.71}$$

$$\mathcal{U}_{\Delta}^{\text{SCI}} = \{\, \mathcal{U}^{\text{SCI}} \mid \mu_2 \neq \mu_1 \wedge \mu_2 \neq \mu_3 \,\} \tag{5.72}$$

$$\mathcal{U}_{\phi}^{\text{SCI}} = \{\, \mathcal{U}^{\text{SCI}} \mid \mu_2 = \mu_1 \vee \mu_2 = \mu_3 \,\}. \tag{5.73}$$

Note, that we have again $E_H^{\text{SCI}} = E_{H,\Delta}^{\text{SCI}} + E_{H,\phi}^{\text{SCI}}$ and due to Parseval's theorem $E_h^{\text{SCI}} = E_H^{\text{SCI}}$ for $M \to \infty$. The cardinalities of the sets are $|\mathcal{U}^{\text{SCI}}| = M^3$, $|\mathcal{U}_{\phi}^{\text{SCI}}| = 2M^2 - M$ and $|\mathcal{U}_{\Delta}^{\text{SCI}}| = |\mathcal{U}^{\text{SCI}}| - |\mathcal{U}_{\phi}^{\text{SCI}}|$. With the RRC pulse-shape and $\rho = 0$ we find again that $H_{\rho}(\boldsymbol{\mu}) = 1$ with $\boldsymbol{\mu} \in \mathcal{U}_{\phi}^{\text{SCI}}$, and with that the kernel energy is simplified to $E_{H,\phi}^{\text{SCI}} = (2M-1)/M^2 \approx 2/M$.

The cross-channel sets are defined according to (5.23) with only a single degeneracy.

Algorithm 3: REGLOG-PERT-FD for the XCI contribution of the ν^{th} wavelength channel

$\boldsymbol{a}_\lambda[k] = \text{overlapSaveSplit}(\langle\, \boldsymbol{a}[k] \,\rangle, M, K)$
$\boldsymbol{b}_\lambda[k] = \text{overlapSaveSplit}(\langle\, \boldsymbol{b}_\nu[k] \,\rangle, M, K)$
$k, \mu, \mu_1, \mu_2 \in \{\, 0, 1, \ldots, M-1 \,\}$
$H_\nu[\mu_1, \mu_2, \mu_3] = H_\nu[\boldsymbol{\mu}] = H_\nu(\mathrm{e}^{\mathrm{j}\frac{2\pi}{M}\boldsymbol{\mu}})$
forall λ **do**
  $\boldsymbol{A}_\lambda[\mu] = \mathrm{DFT}\{\, \boldsymbol{a}_\lambda[k] \,\}$
  $\boldsymbol{B}_\lambda[\mu] = \mathrm{DFT}\{\, \boldsymbol{b}_\lambda[k] \,\}$
  $\bar{\phi}_\lambda^{\text{XCI}} = -\frac{3}{2}\frac{\phi_{\text{NL},\nu}}{M^2}\ \sum_\mu \|\boldsymbol{B}_\lambda[\mu]\|^2$
  $\vec{\boldsymbol{S}}_\lambda^{\text{XCI}} \cdot \vec{\boldsymbol{\sigma}} = -\frac{1}{2}\frac{\phi_{\text{NL},\nu}}{M^2}\sum_\mu 2\boldsymbol{B}_\lambda[\mu]\boldsymbol{B}_\lambda^{\mathsf{H}}[\mu] - \|\boldsymbol{B}_\lambda[\mu]\|^2\,\mathbf{I}$
  forall μ **do**
    $\mu_3 = \mathrm{mod}_M(\mu - \mu_1 + \mu_2)$
    $\mathcal{U} = \{\, [\mu_1, \mu_2]^{\mathsf{T}} \mid \mu_2 \neq \mu_1 \,\}$
    $\Delta\boldsymbol{A}_\lambda^{\text{XCI}}[\mu] =$
    $-\mathrm{j}\frac{\phi_{\text{NL},\nu}}{M^2}\sum_{\mathcal{U}}(\boldsymbol{B}_\lambda[\mu_1]\boldsymbol{B}_\lambda^{\mathsf{H}}[\mu_2] + \boldsymbol{B}_\lambda^{\mathsf{H}}[\mu_2]\boldsymbol{B}_\lambda[\mu_1]\mathbf{I})\boldsymbol{A}_\lambda[\mu_3]H_\nu[\mu_1, \mu_2, \mu_3]$
    $\boldsymbol{Y}_\lambda^{\text{PERT}}[\mu] = \exp(\mathrm{j}\bar{\phi}_\lambda^{\text{XCI}}\mathbf{I} + \mathrm{j}\vec{\boldsymbol{S}}_\lambda^{\text{XCI}} \cdot \vec{\boldsymbol{\sigma}})(\boldsymbol{A}_\lambda[\mu] + \Delta\boldsymbol{A}_\lambda^{\text{XCI}}[\mu])$
  end
  $\boldsymbol{y}_\lambda^{\text{PERT}}[k] = \mathrm{DFT}^{-1}\{\, \boldsymbol{Y}_\lambda^{\text{PERT}}[\mu] \,\}$
end
$\langle \boldsymbol{y}^{\text{PERT}}[k] \rangle = \text{overlapSaveAppend}(\boldsymbol{y}_\lambda^{\text{PERT}}[k], M, K)$

5.3 Numerical Results

This section complements the theoretical considerations of the previous sections by numerical simulations. To this end, we compare the simulated received symbol sequence $\langle\, \boldsymbol{y}[k] \,\rangle$ obtained by the perturbation-based (PERT) end-to-end channel models to the sequence obtained by numerical evaluation via the SSFM (in the following indicated by the superscript SSFM).

Methodology The evaluated metric is the *normalized* MSE between the two T-spaced output sequences for a given input symbol sequence $\langle\, \boldsymbol{a}[k] \,\rangle$, i.e., we have

$$\sigma_{\mathrm{e}}^2 \overset{\text{def}}{=} \mathrm{E}\{\, \|\boldsymbol{y}^{\text{SSFM}} - \boldsymbol{y}^{\text{PERT}}\|^2 \,\}, \tag{5.74}$$

where the expectation takes the form of a statistical average of the received sequence over the discrete time index k. The MSE is already normalized due to the fixed variance $\sigma_a^2 = 1$ of the symbol alphabet and the receiver-side re-normalization in (3.69), such that the perturbation-free part of the received sequence has (approximately[14]) the same fixed variance as the transmit sequence.

The simulation parameters are summarized in Table 5.1. A total number of $N_{\text{SYM}} = 2^{16}$ transmit symbols $\langle\, \boldsymbol{a}[k] \,\rangle$ are randomly drawn from a PDM 64-ary QAM symbol alphabet $\mathcal{A}$

[14]In the numerical simulation via SSFM *signal depletion* takes place due to an energy transfer from signal to NLI. For simplicity and reproducibility of the results, this additional signal energy loss is not accounted for by additional receiver-side re-normalization.

Table 5.1: Simulation Parameters

<table>
<tr><td>$\boldsymbol{a}, \boldsymbol{b} \in \mathcal{A}$</td><td colspan="2">PDM 64-QAM</td></tr>
<tr><td>M</td><td colspan="2">4096 ($\equiv$ 64-QAM per polarization)</td></tr>
<tr><td>$h_{\mathrm{T},\rho}(t)$, $h_{\mathrm{T},\nu}(t)$</td><td colspan="2">$h_{\mathrm{RRC}}(t)$ with roll-off factor ρ</td></tr>
<tr><td>$\bar{\gamma}$</td><td colspan="2">$1.1\,\mathrm{W}^{-1}\mathrm{km}^{-1}$</td></tr>
<tr><td>$\bar{\beta}_2$</td><td colspan="2">$-21\,\mathrm{ps}^2/\mathrm{km}$</td></tr>
<tr><td>$\mathcal{B}_0$</td><td colspan="2">$0\,\mathrm{ps}^2$</td></tr>
<tr><td>$\mathcal{B}(z)$</td><td colspan="2">$\bar{\beta}_2 z$</td></tr>
<tr><td>$10\log_{10}\mathrm{e}^{\bar{\alpha}}$</td><td>$0\,\mathrm{dB/km}$</td><td>$0.2\,\mathrm{dB/km}$</td></tr>
<tr><td>L_{sp}</td><td>$21.71\,\mathrm{km}$</td><td>$100\,\mathrm{km}$</td></tr>
<tr><td>$\mathcal{G}(z)$</td><td>0</td><td>$-\bar{\alpha}z + \bar{\alpha}L_{\mathrm{sp}}\sum_{i=1}^{N_{\mathrm{sp}}}\delta(z - iL_{\mathrm{sp}})$</td></tr>
<tr><td>N_{SYM}</td><td colspan="2">2^{16}</td></tr>
<tr><td>M</td><td colspan="2">$\max\left(2^{\lceil \log_2 \mathcal{S}_{T,\nu} \rceil + 2}, 64\right)$</td></tr>
<tr><td>$10\log_{10}\mathit{\Gamma}$</td><td colspan="2">$-60\,\mathrm{dB}$</td></tr>
</table>

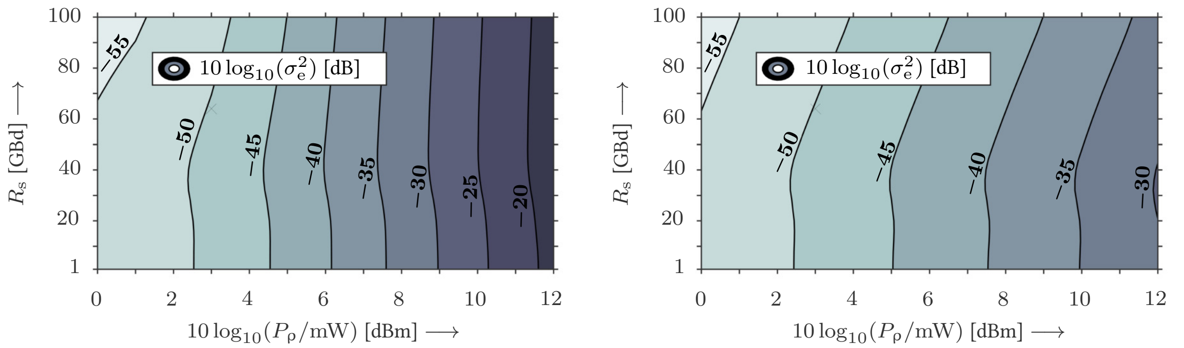
with (4D) cardinality $M = |\mathcal{A}| = 4096$, i.e., 64-QAM per polarization. The transmit pulse shape of the probe $h_{\mathrm{T},\rho}(t)$ and interferer $h_{\mathrm{T},\nu}(t)$ have an RRC shape with roll-off factor ρ, and energies $E_{\mathrm{T},\rho}$ and $E_{\mathrm{T},\nu}$ to vary the optical launch power[15] P_ρ and P_ν.

Two different optical amplification schemes are considered: ideal distributed Raman amplification (i.e., lossless transmission) and *transparent* end-of-span lumped amplification (i.e., lumped amplification where the effect of *signal-gain depletion* [Gha17, Sec. II B.] is neglected in the derivation of the perturbation model). For lumped amplification we consider homogeneous spans of SSMF with path-invariant attenuation coefficient $10\log_{10}\mathrm{e}^{\bar{\alpha}} = 0.2\ \mathrm{dB/km}$ and a span length of $L_{\mathrm{sp}} = 100\ \mathrm{km}$. In case of lossless transmission we have $10\log_{10}\mathrm{e}^{\bar{\alpha}} = 0\ \mathrm{dB/km}$ and span length $L_{\mathrm{sp}} = 21.71\ \mathrm{km}$ corresponding to the asymptotic effective length $L_{\mathrm{eff,a}} \stackrel{\mathrm{def}}{=} 1/\bar{\alpha}$ of a fictitious fiber with infinite length and attenuation $10\log_{10}\mathrm{e}^{\bar{\alpha}} = 0.2\ \mathrm{dB/km}$. In doing so, the nonlinear phase shift $\phi_{\mathrm{NL},\nu}$ from (5.13) remains approximately constant for both scenarios, and the numerical results remain comparable.

The dispersion profile $\mathcal{B}(z) = \bar{\beta}_2 z$ conforms with modern DU links, i.e., without optical in-line dispersion compensation and bulk compensation at the receiver-side (typically performed in the digital domain). Dispersion pre-compensation at the transmit-side can be incorporated via $\mathcal{B}_0$ but is not considered here. The dispersion coefficient is $\bar{\beta}_2 = -21\ \mathrm{ps}^2/\mathrm{km}$ and the nonlinearity coefficient is $\bar{\gamma} = 1.1\ \mathrm{W}^{-1}\mathrm{km}^{-1}$, both invariant over z and ω. Additive noise due to ASE and laser PN are neglected since we only focus on deterministic signal-to-signal NLI.

The numerical reference simulation is a full-vectorial field simulation implemented via the *symmetric split-step Fourier method*, see Section 3.3.6. The maximum nonlinear phase-rotation per step is $\phi_{\mathrm{NL},\Delta z}^{\max} = 3.5 \times 10^{-4}$ rad. The simulation bandwidth is $B_{\mathrm{SIM}} = 8R_{\mathrm{s}}$ for single-channel and $16R_{\mathrm{s}}$ for dual-channel transmission. All filter operations in the SSFM

[15] As of yet, signals were always treated as dimensionless entities, but by convention we will still associate the optical launch power P with units of *Watt* $[\mathrm{W}]$ and the nonlinearity coefficient γ with $[1/(\mathrm{Wm})]$, hence, e.g., the normalization in the x-axis of Figure 5.9.



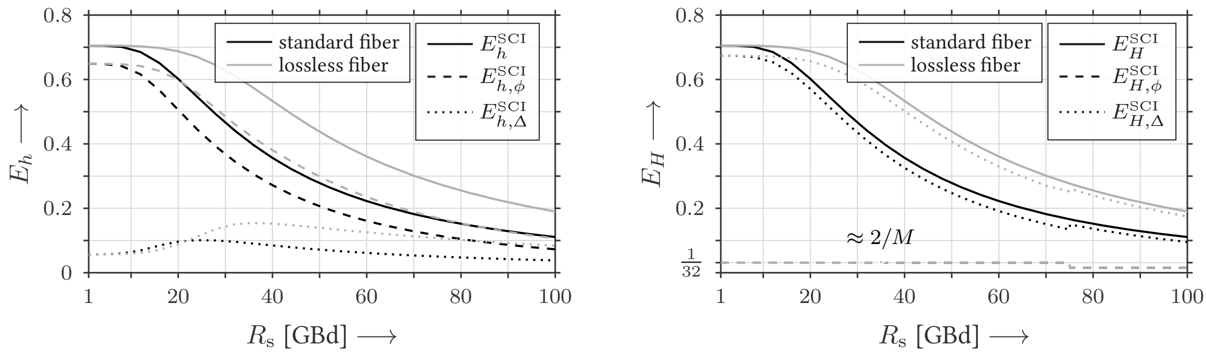
(a) REG-PERT-TD, single-channel, single-span, lossless fiber (b) REGLOG-PERT-TD, single-channel, single-span, lossless fiber

Figure 5.9: Contour plot of the normalized mean-square error $\sigma_e^2 = \mathrm{E}\{\,\|\boldsymbol{y}^{\text{SSFM}} - \boldsymbol{y}^{\text{PERT}}\|^2\,\}$ in dB between the perturbation-based (PERT) end-to-end model and the split-step Fourier method (SSFM) [FFF20]. The results are shown w.r.t. the symbol rate R_s and the optical launch power of the probe P_ρ in dBm. Parameters as in Table 5.1 with roll-off factor $\rho = 0.2$, $N_{\text{sp}} = 1$, $10\log_{10}\mathrm{e}^{\bar{\alpha}} = 0\,\text{dB/km}$ and $L_{\text{sp}} = 21.71\,\text{km}$. In (a) the regular (REG) time-domain (TD) model is carried out as in (5.15) and in (b) the regular-logarithmic (REGLOG) model is carried out as in (5.38).

reference simulation (i.e., pulse-shaping, linear step in the SSFM, linear channel matched filter) are performed at the full simulation bandwidth via fast convolution and regarding periodic boundary conditions.

Discussion of the Results In Figure 5.9 (a), we start our evaluation with the most simple scenario, i.e., single-channel, single-span, and lossless transmission. The MSE is shown in logarithmic scale $10\log_{10}(\sigma_e^2)$ in dB over the symbol rate R_s and the launch power of the probe $10\log_{10}(P_\rho/\text{mW})$ in dBm. The results are obtained from the regular (REG) perturbation-based (PERT) end-to-end channel model in discrete time-domain (TD), corresponding to (5.15). For the given effective length L_{eff} and dispersion parameter $\bar{\beta}_2$, the range of the symbol rate between $1\,\text{GBd}$ and $100\,\text{GBd}$ corresponds to a map strength $\mathcal{S}_{T,\rho}$ between 0.003 and 28.7. This amounts to virtually no memory of the intra-channel nonlinear interaction for small symbol rates (hence only very few coefficients $h_\rho[\boldsymbol{\kappa}]$ exceeding the minimum energy level of $10\log_{10}\varGamma^{\text{SCI}} = -60\,\text{dB}$) to a very broad intra-channel nonlinear memory for high symbol rates (with coefficients $h_\rho[\boldsymbol{\kappa}]$ covering a large number of symbols). Likewise, the launch power of the probe P_ρ spans a nonlinear phase shift $\phi_{\text{NL},\rho}$ from 0.02 to $0.34\,\text{rad}$. We can observe a gradual increase in σ_e^2 of about $5\,\text{dB}$ per $1.5\,\text{dBm}$ launch power in the nonlinear transmission regime. We deliberately consider a MSE $10\log_{10}\sigma_e^2 > -30\,\text{dB}$ as a *poor* match between the perturbation-based model and the full-field simulation, i.e., here for P_ρ larger than $9\,\text{dBm}$ ($\equiv 0.168\,\text{rad} \approx 10°$ in terms of average nonlinear phase shift $\phi_{\text{NL},\rho}$) independent of R_s.

In Figure 5.9 (b) the same system scenario is considered but instead of the regular model, now, the regular-logarithmic (REGLOG) model is employed according to (5.38). The gradual increase in σ_e^2 with increasing P_ρ is now considerably relaxed to about $5\,\text{dB}$ per $2.5\,\text{dBm}$ launch power. The region of poor model match with $10\log_{10}\sigma_e^2 > -30\,\text{dB}$ is now only approached for launch powers larger than $12\,\text{dBm}$. We can also observe that σ_e^2 improves with increasing symbol rate R_s, in particular for rates $R_s > 40\,\text{GBd}$. This is explained by the fact that the



(a) PERT-TD, single-channel, single-span

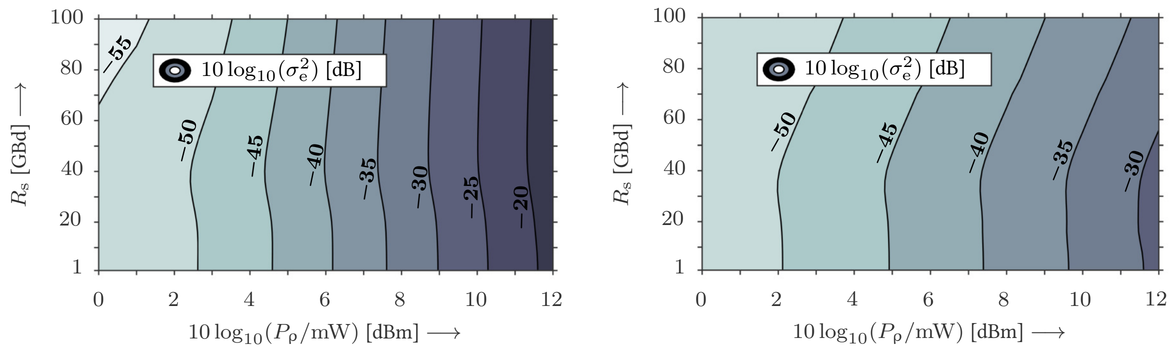
(b) PERT-FD, single-channel, single-span

Figure 5.10: Energy of the kernel coefficients in time domain E_h (a) and in frequency domain E_H (b) over the symbol rate R_s [FFF20]. The kernel coefficients are obtained from the regular-logarithmic (REGLOG) model for a single-channel ($\rho = 0.2$) over a standard single-mode fiber ($10 \log_{10} \mathrm{e}^{\bar{\alpha}} = 0.2\,\mathrm{dB/km}$ and $L_\mathrm{sp} = 100\,\mathrm{km}$) or a lossless fiber ($10 \log_{10} \mathrm{e}^{\bar{\alpha}} = 0\,\mathrm{dB/km}$ and $L_\mathrm{sp} = 21.71\,\mathrm{km}$). The subscript Δ denotes the subset of all coefficients associated with additive perturbations and the subscript ϕ denotes the subset of all coefficients with multiplicative perturbations.

kernel energy E_h^{SCI} in (5.57) depends on the symbol rate R_s such that σ_e^2 is reduced for higher symbol rates.

Figure 5.10 (a) shows the energy of the (time-domain) kernel coefficients E_h^{SCI} over R_s for a single-span SSMF with $L_\mathrm{sp} = 100\,\mathrm{km}$ and for a lossless fiber with $L_\mathrm{sp} = 21.71\,\mathrm{km}$. Generally, we see that E_h^{SCI} is constant for small R_s and then curves into a transition region towards smaller energies for increasing R_s. For transmission over SSMF this transition region is shifted to smaller R_s, e.g., E_h^{SCI} drops from 0.7 to 0.6 around $33\,\mathrm{GBd}$ for lossless transmission and at around $20\,\mathrm{GBd}$ for transmission over SSMF. We also present the kernel energies $E_{h,\Delta}^{\mathrm{SCI}}$ associated with additive perturbations, and $E_{h,\phi}^{\mathrm{SCI}}$ associated with multiplicative perturbations. For this single-span scenario, most of the energy is concentrated in $E_{h,\phi}^{\mathrm{SCI}}$, i.e., corresponding to the degenerate symbol combinations with $\kappa_1 = 0$ or $\kappa_3 = 0$ defined in (5.41)–(5.43). Interestingly, while the total energy E_h^{SCI} decreases monotonically with R_s, the additive contribution $E_{h,\Delta}^{\mathrm{SCI}}$ increases in the transition region and then decreases again for large R_s. This behavior is also visible in the results presented in Figure 5.9 (a) and (b).

Figure 5.10 (b) shows the energy of the kernel coefficients E_H^{SCI} of the frequency-domain approach for the same system scenario as in (a). The total energies are the same, i.e., $E_h^{\mathrm{SCI}} = E_H^{\mathrm{SCI}}$ (cf. Parseval's theorem), however, the majority of the energy is now contained in the regular (additive) subset of coefficients. The energy of the degenerate, i.e., multiplicative, subset of coefficients $E_{H,\phi}^{\mathrm{SCI}}$ depends on the frequency discretization (which coincides here with the transformation length M) and is approximately $2/M$. The exact value $(2M-1)/M^2$ would be achieved for $\rho = 0$. For $R_\mathrm{s} > 75.1\,\mathrm{GBd}$ we have $\mathcal{S}_{T,\rho} > 16$ and it can be seen that $E_{H,\phi}^{\mathrm{SCI}}$ drops from $1/32$ to $1/64$ and $E_{H,\Delta}^{\mathrm{SCI}}$ jumps up by an equal amount because M increases from 64 to 128 (cf. the set of simulation parameters in Table 5.1). The REGLOG frequency-domain model is hence predominantly a regular model, where only the average multiplicative effects are truly treated as such.



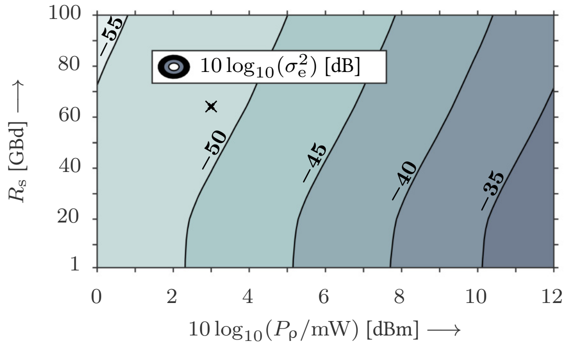
(a) REG-PERT-FD, single-channel, single-span, lossless fiber (b) REGLOG-PERT-FD, single-channel, single-span, lossless fiber

Figure 5.11: Contour plot of the normalized mean-square error σ_e^2 in dB [FFF20]. The results are shown w.r.t. the symbol rate R_s and the optical launch power of the probe P_ρ in dBm. Parameters as in Table 5.1 with roll-off factor $\rho = 0.2$, $N_{sp} = 1$, $10\log_{10} e^{\bar{\alpha}} = 0\,\mathrm{dB/km}$ and $L_{sp} = 21.71\,\mathrm{km}$. In (a) the regular (REG) frequency-domain (FD) model is carried out as in Algorithm 1 and in (b) the regular-logarithmic (REGLOG) model is carried out as in Algorithm 2.

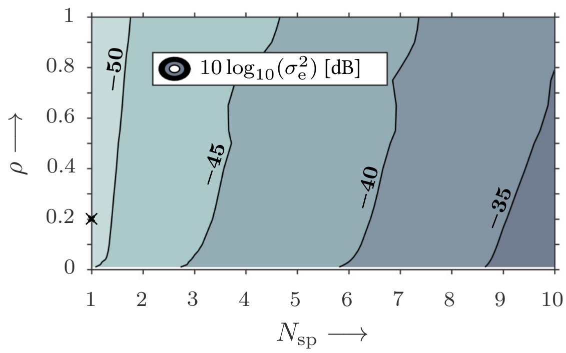
In Figure 5.11 (a) and (b), the respective results on σ_e^2 using the discrete frequency-domain (FD) model according to Algorithm 1 and 2 are shown. We can confirm our previous statement that the regular perturbation model in time and frequency are equivalent considering that the results shown in Figure 5.9 (a) and Figure 5.11 (a) are (virtually) the same. We also observe that the REGLOG-FD performs very similar to the corresponding TD model despite the fact that only *average* terms are considered as multiplicative distortions. We conclude that—in the considered system scenario—REGLOG models benefit from the fact that the *average* phase and polarization rotations are properly represented compared to pure REG models. The *time-variant* phase and polarization rotations that fluctuate around the average can to some extent also be represented by an additive perturbation without significant loss in performance. This observation is in line with the eRP method introduced in [VSB02, Sec. VI]. In the eRP view, the perturbation expansion is performed in a "SPM-rotated reference system" [SB13, SB15] where the time-average phase rotation is a priori included or a posteriori removed from the regular solution, cf. [SB13, Eq. (33)].

Figure 5.12 (a) shows σ_e^2 for a single-channel over standard single-mode fiber ($L_{sp} = 100\,\mathrm{km}$ and $10\log_{10} e^{\bar{\alpha}} = 0.2\,\mathrm{dB/km}$) and lumped end-of-span amplification. In the full-field simulation, the lumped amplifier is operated in *constant-gain* mode compensating for the exact span-loss of $20\,\mathrm{dB}$. The results over a single-span in Figure 5.12 (a) are very similar in the low symbol rate regime compared to the lossless case in Figure 5.9 (b). For R_s larger than $20\,\mathrm{GBd}$, the MSE starts to decrease at a higher rate compared to the lossless case. This is in line with the energy of the kernel coefficients E_h^{SCI} for the standard fiber shown in Figure 5.10 (a).

In Figure 5.12 (b), σ_e^2 is shown over the roll-off factor ρ and the number of spans N_{sp} for a fixed symbol rate of $R_s = 64\,\mathrm{GBd}$ and a fixed launch power of $10\log_{10}(P_\rho/\mathrm{mW}) = 3\,\mathrm{dBm}$. The black cross in Figure 5.12 (a) and (b) indicates the point with a common set of parameters. We can see a dependency on the roll-off factor ρ which is due to a dependency of E_h^{SCI} on ρ (not shown here). With increasing ρ the kernel energy E_h^{SCI} decreases and hence does σ_e^2, too.



(a) REGLOG-PERT-TD, single-channel, single-span, standard fiber



(b) REGLOG-PERT-TD, single-channel, multi-span, standard fiber

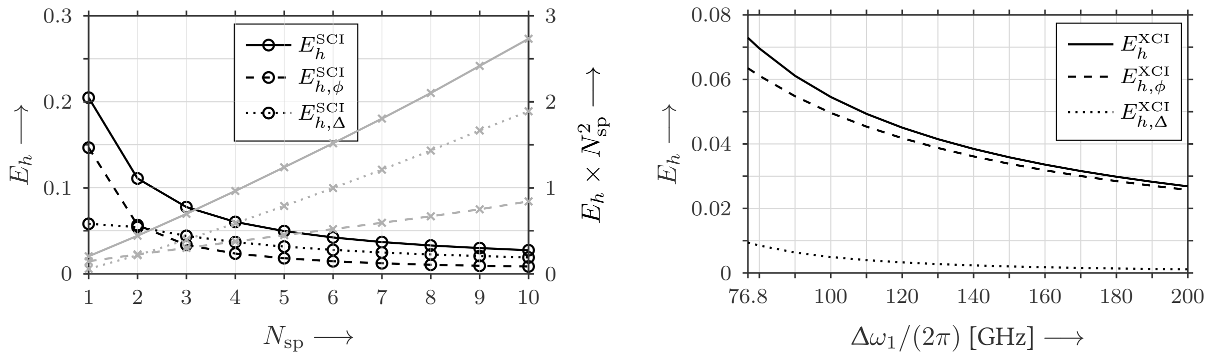
Figure 5.12: Contour plot of the normalized mean-square error σ_e^2 in dB [FFF20]. The results are obtained from the regular-logarithmic (REGLOG) time-domain (TD) model over a standard single-mode fiber ($10\log_{10} e^{\tilde{\alpha}} = 0.2\,\text{dB/km}$ and $L_{\text{sp}} = 100\,\text{km}$) with end-of-span lumped amplification. In (a) the symbol rate R_s and the optical launch power P_ρ are varied for single-span ($N_{\text{sp}} = 1$) transmission and fixed roll-off factor ($\rho = 0.2$). In (b) the roll-off factor ρ and number of spans N_{sp} are varied with fixed symbol rate ($R_s = 64\,\text{GBd}$) and fixed launch power ($10\log_{10}(P_\rho/\text{mW}) = 3\,\text{dBm}$). The black marker indicates the joint reference point with the same absolute value of $\sigma_e^2 = -51.4\,\text{dB}$ but different gradient over the sweep parameter.

The scaling laws of σ_e^2 with N_{sp} are complemented in Figure 5.13 (a) by the energy of the kernel coefficients E_h^{SCI} for the same system scenario as in Figure 5.12 (b) (with $\rho = 0.2$). It is interesting to see that (for this particular system scenario) $E_{h,\Delta}^{\text{SCI}}$ and $E_{h,\phi}^{\text{SCI}}$ intersect at $N_{\text{sp}} = 2$. We can conclude that after the second span more energy is comprised within the additive subset of coefficients than in the multiplicative one. With increasing N_{sp} the relative contribution of $E_{h,\Delta}^{\text{SCI}}$ to the total energy E_h^{SCI} is increasing. Note, while E_h^{SCI} is actually monotonically decreasing with N_{sp}, the common pre-factor $\phi_{\text{NL},\rho}$ has to be factored in as it effectively scales the nonlinear distortion. Since for heterogeneous spans we have $\phi_{\text{NL},\rho} \propto L_{\text{eff}} \propto N_{\text{sp}}$, the same traces are shown scaled by N_{sp}^2 to illustrate how the energy of the total distortion accumulates with increasing transmission length. In this respect, similar results can be obtained from the presented channel model as from the GN/extended Gaussian-noise (EGN)-model (given proper scaling with $\phi_{\text{NL},\rho}^2$ instead of just N_{sp}^2, and similarly taking all other wavelength channels into account).

In particular, the model correctly predicts the strength of the nonlinear distortion when the roll-off factor is larger than zero. Then, aliasing of frequency components from nonlinear distortions is properly included; the GN/EGN-model does not take the aliasing into account.

Additionally, qualitative statements can be derived, e.g., whether the nonlinear distortion is predominantly additive or multiplicative. From the energy spread of the kernel coefficients one can also deduce the time scale over which nonlinear distortions are still correlated.

Figure 5.14 shows the σ_e^2 for dual-channel transmission using either the REGLOG time-domain (a) or frequency-domain model (b). The transmit symbols of the interferer $\langle\, \boldsymbol{b}[k]\,\rangle$ are drawn from the same symbol set $\mathcal{A}$, i.e., 64-QAM per polarization. For both wavelength channels, the symbol rate is fixed to $R_s = 64\,\text{GBd}$ and the roll-off factor of the RRC shape is $\rho = 0.2$. The transmit power of the probe is set to $10\log_{10}(P_\rho/\text{mW}) = 0\,\text{dBm}$ while the transmit power of the interferer P_ν with channel number $\nu = 1$ is varied together with the



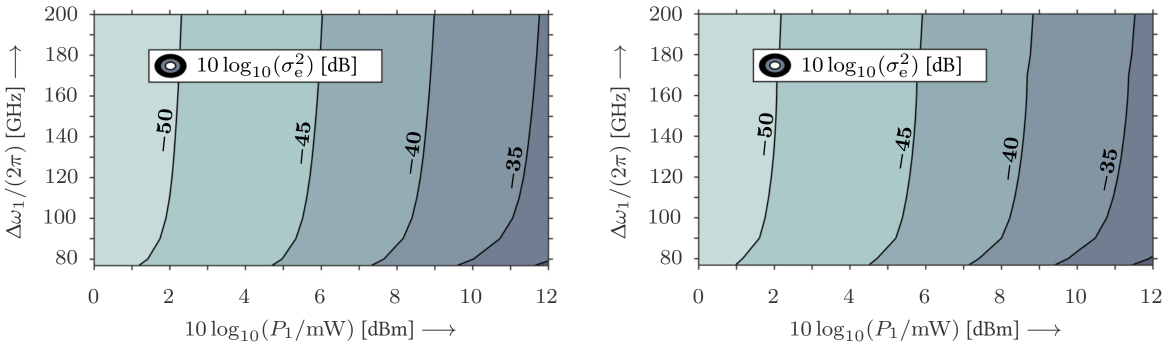
(a) PERT-TD, single-channel, multi-span, standard fiber

(b) PERT-TD, dual-channel, single-span, lossless fiber

Figure 5.13: In (a), the energy of the kernel coefficients (black lines, bullet markers, left y-axis) in time domain E_h is shown over N_{sp} spans of standard single-mode fiber ($10\log_{10}\mathrm{e}^{\bar{\alpha}} = 0.2\,\mathrm{dB/km}$ and $L_{\mathrm{sp}} = 100\,\mathrm{km}$, $\rho = 0.2$) [FFF20]. Additionally, the kernel energies are shown scaled with $N_{\mathrm{sp}}^2 \propto \phi_{\mathrm{NL},\rho}^2$ (gray lines, cross markers, right y-axis) to indicate the general growth of nonlinear distortions with increasing N_{sp} (similar to the GN-model). In (b), kernel energies E_h are shown for cross-channel interference (XCI) imposed by a single wavelength channel spaced at $\Delta\omega_1/(2\pi)\,\mathrm{GHz}$ over a single span of lossless fiber. Both probe and interferer have $R_{\mathrm{s}} = 64\,\mathrm{GBd}$ and $\rho = 0.2$.

relative frequency offset $\Delta\omega_1/(2\pi)$ ranging from $76.8\,\mathrm{GHz}$ (i.e., no guard interval with $(1+\rho)\times 64\,\mathrm{GHz}$) to $200\,\mathrm{GHz}$. In the numerical simulation via SSFM we use an ideal channel combiner and both wavelength channels co-propagate at the full simulation bandwidth $B_{\mathrm{SIM}} = 16R_{\mathrm{s}}$. In case of the end-to-end channel model both contributions from intra- and inter-channel distortions are combined into a single perturbative term (cf. (5.38) and (5.60)). The baseline error σ_{e}^2 is therefore approximately $-55\,\mathrm{dB}$ considering the respective case with $R_{\mathrm{s}} = 64\,\mathrm{GBd}$ and $P_{\rho} = 0\,\mathrm{dBm}$ in Figure 5.9 (b). It is seen that the time- and frequency-domain model perform very similar.

The dependency on the channel spacing $\Delta\omega_1$ is explained considering Figure 5.13 (b). Here, the energy of the cross-channel coefficients $h_1[\boldsymbol{\kappa}]$ is shown over $\Delta\omega_1$. Generally, with increasing $\Delta\omega_1$, E_h^{XCI} decreases and additionally the relative contribution of the degeneracy at $\kappa_3 = 0$, i.e., $E_{h,\phi}^{\mathrm{SCI}}$, is growing. Ultimately, the main distortion caused by an interferer spaced far away from the probe channel is a distortion in phase and state of polarization.



(a) REGLOG-PERT-TD, dual-channel, single-span, lossless fiber (b) REGLOG-PERT-FD, dual-channel, single-span, lossless fiber

Figure 5.14: Contour plot of the normalized mean-square error σ_e^2 in dB. The results are obtained from two co-propagating wavelength channels with PDM 64-QAM and a symbol rate of 64 GBd and roll-off factor $\rho = 0.2$ [FFF20]. The launch power of the probe is fixed at $10\log_{10}(P_\rho/\mathrm{mW}) = 0\,\mathrm{dBm}$ while the power of the interferer P_1 and the relative frequency offset $\Delta\omega_1$ are varied. In (a) the regular-logarithmic (REGLOG) time-domain (TD) model for both SCI and XCI is carried out as in (5.38) and in (b) the REGLOG frequency-domain (FD) model is carried out as in Algorithm 2 and (5.60) for both SCI and XCI.

6. Concluding Remarks

In this thesis, a comprehensive analysis of end-to-end channel models for fiber-optic transmission based on a perturbation approach is presented. The existing view on continuous- and discrete-time end-to-end channel models is described in a unified framework. A novel frequency-domain perspective that incorporates the time-discretization via an aliased kernel in frequency domain is presented. The relation between the time- and frequency-domain representation for continuous- and discrete-time Volterra systems is elucidated and we show that the kernel coefficients in both views are related by multi-dimensional (continuous or discrete-time) Fourier transforms. The energy of the *un-aliased* kernel can be directly related to the conventional GN-model. The energy of the *aliased* kernel also takes the T-spaced sampling in the receiver into account, here referred to as *Discrete-Time GN-model*.

While the pulse collision picture addresses the importance of separating additive and multiplicative terms, particularly, for *inter*-channel nonlinear interactions, a generalization to *intra*-channel nonlinear interactions is presented. An intra-channel phase distortion term and an intra-channel XPolM term are introduced and both correspond to a subset of degenerate intra-channel pulse collisions. In analogy to implementations of time-domain RP method, the implementation of the frequency-domain model is modified to also treat certain degenerate mixing products as multiplicative distortions. As a result, we have established a complete formulation of strictly *regular* (i.e., additive) models, and *regular-logarithmic* (i.e., mixed additive and multiplicative) models—both in time and in frequency domain, both for intra- and inter-channel nonlinear interference.

Derived from the frequency-domain description, a novel class of algorithms is proposed which effectively computes the end-to-end relation between transmit and receive sequences over discrete frequencies from the Nyquist interval. One potential application of the frequency-domain model can be in fiber nonlinearity compensation. Here, the model can be applied in a reverse manner at the transmitter side before pulse-shaping or on the receiver side after matched filtering. Moreover, while the time-domain implementation requires a triple summation per time-instance, the frequency-domain implementation involves only a double summation per frequency index. Similar as for linear systems, this characteristic may allow for an efficient implementation using the fast Fourier transform when the time-domain kernel comprises many coefficients.

The derived algorithms were compared to the (oversampled and inherently sequential) split-step Fourier method based on the mean-squared error between both output sequences. We show that, in particular, the regular-logarithmic models have good agreement with the split-step Fourier method over a wide range of system parameters. The presented results are further supported by a qualitative analysis involving the kernel energies to quantify the relative contributions of either additive or multiplicative distortions.

We can identify three relevant system parameters that characterize the nonlinear response: Firstly, the map strength $\mathcal{S}_{T,\rho}$ (or equivalently the ν-dependent $\mathcal{S}_{T,\nu}$) which is a measure of the temporal extent, i.e., the memory of the nonlinear interaction. Secondly, the (ν-dependent) nonlinear phase shift $\phi_{\mathrm{NL},\nu}$ that depends via $L_{\mathrm{NL},\nu}$ linearly on the launch power P_ν and essentially acts as a scaling factor to the nonlinear distortion $\Delta\boldsymbol{a}[k]$. Finally, the total kernel energy $E_{h,\nu}$ which characterizes the strength of the nonlinear interaction—independent of the launch power.

Future work may address a thorough complexity analysis of the presented models which has relevance for the application in efficient channel models or compensation algorithms.

A. Appendix

A.1 The Jones and Stokes Formalism

This appendix deals with the basics of the Jones and Stokes formalism which will be used to describe the evolution of the optical signal and the signal's state of polarization. We recap the (special) notation often used in the optical community to describe phase- and polarization rotations. The general concepts are mainly adapted from [GK00] and [Win09] and modified to integrate well with our notation. In particular, the notation developed in this appendix will be used to describe polarization rotations induced by *nonlinear* signal-to-signal interactions.

The spatial evolution of the optical baseband signal $\boldsymbol{u}(z,t)$ over a *linear* channel can generally be described by the linear, homogeneous, PDE in frequency domain given by [VB02, Eq. (6)]

$$\frac{\partial}{\partial z}\boldsymbol{U}(z,\omega) = -\mathrm{j}\,\boldsymbol{T}(z,\omega)\boldsymbol{U}(z,\omega)\,, \tag{A.1}$$

with the Fourier pair $\boldsymbol{u}(z,t) \circ\!\!-\!\!\bullet\, \boldsymbol{U}(z,\omega)$ and the transmission matrix $\boldsymbol{T}(z,\omega) \in \mathbb{C}^{2\times 2}$ describing the optical transmission medium[1].

For lossless transmission, we require that the transmission matrix $\boldsymbol{T}(z,\omega)$ does not alter the squared Euclidean norm of the signal, i.e., the energy of the signal must be preserved. In this case, signal evolution is simply modeled by unitary *rotations* in Jones space (i.e., *phase* and *polarization* rotations). As a consequence, the evolution equation in (A.1) must preserve the squared norm of the optical field envelope. Mathematically, this is obtained by imposing the following constraint to the propagation equation

$$\frac{\partial}{\partial z}\,\|\boldsymbol{U}(z,\omega)\|^2 \overset{!}{=} 0\,, \tag{A.2}$$

where we use the product rule of calculus $\frac{\partial}{\partial z}\boldsymbol{u}^{\mathsf{H}}\boldsymbol{u} = (\frac{\partial}{\partial z}\boldsymbol{u}^{\mathsf{H}})\boldsymbol{u} + \boldsymbol{u}^{\mathsf{H}}(\frac{\partial}{\partial z}\boldsymbol{u})$, and the property $(\boldsymbol{T}\boldsymbol{U})^{\mathsf{H}} = \boldsymbol{U}^{\mathsf{H}}\boldsymbol{T}^{\mathsf{H}}$ of the Hermitian conjugate to obtain

$$\mathrm{j}\,\boldsymbol{U}^{\mathsf{H}}(z,\omega)\left(\boldsymbol{T}^{\mathsf{H}}(z,\omega) - \boldsymbol{T}(z,\omega)\right)\boldsymbol{U}(z,\omega) = 0\,. \tag{A.3}$$

[1]The matrix components of $\boldsymbol{T}(z,\omega)$ have units m^{-1}. We will relate $\boldsymbol{T}(z,\omega)$ to the *common propagation constant* $\beta(z,\omega)$ and the *birefringence parameter* $\Delta\beta(z,\omega)$ which has not been formally introduced at this point.

The result in (A.3) implies that

$$\boldsymbol{T}(z,\omega) \stackrel{!}{=} \boldsymbol{T}^{\mathsf{H}}(z,\omega)\,, \tag{A.4}$$

i.e., the transmission matrix must hence be *Hermitian* for lossless transmission.

In this context, we introduce the set of *Pauli matrices* which is given by [GK00, Appx. A]

$$\boldsymbol{\sigma}_0 \stackrel{\text{def}}{=} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{\sigma}_1 \stackrel{\text{def}}{=} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \boldsymbol{\sigma}_2 \stackrel{\text{def}}{=} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \boldsymbol{\sigma}_3 \stackrel{\text{def}}{=} \begin{bmatrix} 0 & -\mathrm{j} \\ \mathrm{j} & 0 \end{bmatrix}, \tag{A.5}$$

where each of the Pauli matrices is Hermitian.

The Pauli matrices form an orthogonal basis of the complex-valued 2×2 vector space $\mathbb{C}^{2\times 2}$, i.e., any complex-valued matrix can be represented as the linear combination of the Pauli matrices with complex coefficients. Particularly, any *Hermitian* matrix can be expanded in terms of Pauli matrices with only *real*-valued coefficients.

The Pauli matrices $\boldsymbol{\sigma}_i$ with $i = 1, 2, 3$ are traceless and have a *negative one* determinant, i.e., $\mathrm{trace}(\boldsymbol{\sigma}_i) = 0$ and $\det(\boldsymbol{\sigma}_i) = -1$. The zero$^{\text{th}}$ order Pauli matrix $\boldsymbol{\sigma}_0 = \mathbf{I}$ has $\mathrm{trace}(\boldsymbol{\sigma}_0) = 2$ and $\det(\boldsymbol{\sigma}_0) = 1$, and will later be associated with a *common* (phase) rotation of both Jones vector components—unaltering the relative orientation (i.e., the polarization) of the two vector components.

The algebra generated by the Pauli matrices can be isomorphically expressed by the *quaternion*-valued algebra[2], cf. [KP04]. It is well-known, that quaternions are well suited for calculation of rotations in 3D real-valued space. Here, the coefficients of a *pure* quaternion (one where the real part is equal to zero) have a direct correspondence to the *Stokes parameters* [Sto51] (associated with $\boldsymbol{\sigma}_1$, $\boldsymbol{\sigma}_2$, $\boldsymbol{\sigma}_3$) of the *birefringence vector* which defines the polarization rotation axis.

We will now use the Pauli matrix formalism applied to Hermitian matrices in $\mathbb{C}^{2\times 2}$ to model polarization rotations, as common in optical communication [GK00]. For notational convenience it is often good practice (in the *physics* and *optics* community) to arrange the Pauli matrices into the, so-termed, *Pauli vector* (sometimes also *spin vector*) which is defined as

$$\vec{\boldsymbol{\sigma}} \stackrel{\text{def}}{=} [\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, \boldsymbol{\sigma}_3]^{\mathsf{T}} \in \mathbb{C}^{2\times 2} \otimes \mathbb{R}^3\,, \tag{A.6}$$

where each of the three "*vector components*" is a 2×2 Pauli matrix. Strictly speaking, the Pauli *vector* is not a vector in the original sense but rather an *operator* which spans $\mathbb{C}^{2\times 2} \otimes \mathbb{R}^3$ in tensor space, such that a mapping $(\mathbb{R}^3) \cdot (\mathbb{C}^{2\times 2} \otimes \mathbb{R}^3) \mapsto \mathbb{C}^{2\times 2}$ is induced by the *dot product*. This is explained in the following.

Using the Pauli vector notation, we can express any Hermitian transmission matrix $\boldsymbol{T}(z,\omega)$ as a linear combination of Pauli matrices[3] [AW05, Eq. (3.116)]

$$\boldsymbol{T}(z,\omega) = T_0(z,\omega)\boldsymbol{\sigma}_0 + T_1(z,\omega)\boldsymbol{\sigma}_1 + T_2(z,\omega)\boldsymbol{\sigma}_2 + T_3(z,\omega)\boldsymbol{\sigma}_3 \tag{A.7}$$

$$= T_0(z,\omega)\mathbf{I} + \vec{\boldsymbol{T}}(z,\omega) \cdot \vec{\boldsymbol{\sigma}}\,, \tag{A.8}$$

[2] In particular, the Pauli matrices $\mathrm{j}\boldsymbol{\sigma}_1, \mathrm{j}\boldsymbol{\sigma}_2, \mathrm{j}\boldsymbol{\sigma}_3$ (i.e., multiplied by the imaginary unit to make them *skew*-Hermitian) can be mapped (after reordering) to the *imaginary* units $\mathrm{i}, \mathrm{j}, \mathrm{k}$ of the *Hamilton quaternions* [CS99]

[3] In analogy, a Hamilton quaternion has also four real-valued coefficients, just like the expansion in (A.8).

with four real-valued coefficients $T_i(z,\omega) \in \mathbb{R}$ with $i = 0,1,2,3$, which fully parameterize the (lossless) transmission medium. The second term on the right-hand side of (A.8) is termed *Pauli matrix decomposition.* Here, the dot product should be interpreted as a point-wise multiplication of the three-element vector

$$\vec{\boldsymbol{T}}(z,\omega) \stackrel{\text{def}}{=} [T_1(z,\omega), T_2(z,\omega), T_3(z,\omega)]^\mathsf{T} \in \mathbb{R}^3\,, \tag{A.9}$$

with the Pauli vector $\vec{\boldsymbol{\sigma}}$ from (A.6). The vector $\vec{\boldsymbol{T}}(z,\omega)$ is typically referred to as local and frequency-dependent *birefringence vector* [VB02, Eq. (6)] represented in *Stokes space.*

The three unit basis vectors in Stokes space are defined as

$$\vec{\mathbf{S}}_1 \stackrel{\text{def}}{=} [1,0,0]^\mathsf{T}, \qquad \vec{\mathbf{S}}_2 \stackrel{\text{def}}{=} [0,1,0]^\mathsf{T}, \qquad \vec{\mathbf{S}}_3 \stackrel{\text{def}}{=} [0,0,1]^\mathsf{T}\,, \tag{A.10}$$

where we use *roman* typesetting.

Using the expansion in (A.8), the transmission matrix $\boldsymbol{T}(z,\omega)$ can be rewritten in terms of Stokes parameters $\vec{\boldsymbol{T}}(z,\omega)$ and $T_0(z,\omega)$ to obtain the explicit expression [GK00, Eq. (A.8)]

$$\boldsymbol{T}(z,\omega) = \begin{bmatrix} T_0(z,\omega) + T_1(z,\omega) & T_2(z,\omega) - \mathrm{j}T_3(z,\omega) \\ T_2(z,\omega) + \mathrm{j}T_3(z,\omega) & T_0(z,\omega) - T_1(z,\omega) \end{bmatrix} \in \mathbb{C}^{2\times 2}\,. \tag{A.11}$$

In optical communication, the parameter $T_0(z,\omega)$ will be associated to the local- and frequency-dependent *common propagation constant*

$$\beta(z,\omega) \stackrel{\text{def}}{=} T_0(z,\omega) \in \mathbb{R}\,, \tag{A.12}$$

which has no dependency on the polarization state (or more precisely, it characterizes the *average* propagation constant in both polarization states after transformation to the local birefringence coordinate system, and considering that $\boldsymbol{T}(z,\omega)$ is a random process w.r.t. z and ω). We now focus on polarization-dependent effects.

The norm of the birefringence vector $\|\vec{\boldsymbol{T}}(z,\omega)\|$ is related to the local and frequency-dependent *birefringence parameter* by

$$\Delta\beta(z,\omega) = \beta_+(z,\omega) - \beta_-(z,\omega) \stackrel{\text{def}}{=} \|\vec{\boldsymbol{T}}(z,\omega)\| = \det(\vec{\boldsymbol{T}}(z,\omega)\cdot\vec{\boldsymbol{\sigma}})^{\frac{1}{2}} \in \mathbb{R}\,, \tag{A.13}$$

which quantifies the *difference* in the propagation constant between a locally *fast* (denoted by the $+$ subscript) and *slow* (denoted by the $-$ subscript) propagating polarization state (aka. *eigenmode* or *eigenstate*).

Similar to the common propagation constant $\beta(z,\omega)$, the birefringence parameter $\Delta\beta(z,\omega)$ can be developed into a Taylor series w.r.t. ω by[4]

$$\Delta\beta(z,\omega) = \underbrace{\Delta\beta_0(z)}_{\textit{local}\text{ birefringence}} + \underbrace{\Delta\beta_1(z)\,\omega}_{\textit{local}\text{ first-order PMD}} + \dots\,, \tag{A.14}$$

[4] We assume that the local birefringence eigenstate has the same orientation as the frequency-dependent PMD eigenstate, which is justified according to [Men99, P. 12].

with the Taylor coefficients

$$\Delta\beta_n(z) \stackrel{\text{def}}{=} \frac{\partial^n \Delta\beta(z,\omega)}{\partial\omega^n}\Big|_{\omega=0}, \quad n \in \mathbb{N}\,. \tag{A.15}$$

Note, that the Taylor series expansion is performed in the ECB around $\omega = 0$ since the $\boldsymbol{U}(z,\omega)$ is already the baseband signal.

The frequency-independent contribution (i.e., zero$^{\text{th}}$-order term in the Taylor series) of $\Delta\beta(z,\omega)$ w.r.t. ω is associated with the *local birefringence*. The first-order term (i.e., linear frequency-dependency in the Taylor series) is related to the local *first-order* polarization-mode dispersion (PMD). In the following, we will see that the first-order PMD term which depends only linearly on ω will result in a differential group delay (DGD) between the two polarization eigenstates. The *local birefringence coordinate transformation*, discussed in Section 3.1, aims to remove the zero$^{\text{th}}$-order term $\Delta\beta_0(z)$ from the propagation equation such that only first- and higher-order PMD terms are considered in the baseband model.

If the Stokes vector $\vec{\boldsymbol{T}}(z,\omega) = [T_1(z,\omega), T_2(z,\omega), T_3(z,\omega)]^\mathsf{T}$ is normalized to unit length such that $\|\vec{\boldsymbol{T}}(z,\omega)\| \stackrel{!}{=} 1$, it is termed state of polarization (SOP). We use the *tilde* to refer to normalized Stokes vectors, i.e., the associated SOP of the birefringence vector $\vec{\boldsymbol{T}}(z,\omega)$ is given as

$$\tilde{\boldsymbol{T}}(z,\omega) = [\tilde{T}_1(z,\omega), \tilde{T}_2(z,\omega), \tilde{T}_3(z,\omega)]^\mathsf{T} \stackrel{\text{def}}{=} \vec{\boldsymbol{T}}(z,\omega)/\|\vec{\boldsymbol{T}}(z,\omega)\| \quad \in \mathbb{R}^3\,, \tag{A.16}$$

such that $\|\tilde{\boldsymbol{T}}\| = \det(\tilde{\boldsymbol{T}} \cdot \vec{\boldsymbol{\sigma}})^{\frac{1}{2}} = \sqrt{\tilde{T}_1^2 + \tilde{T}_2^2 + \tilde{T}_3^2} = 1$. An SOP is often visualized as a point on a three-dimensional sphere, the so-called *Poincaré sphere*, spanned by the three unit Stokes vectors in (A.10).

The expanded Jones matrix w.r.t. the normalized Stokes vector $\tilde{\boldsymbol{T}}(z,\omega) \cdot \vec{\boldsymbol{\sigma}}$ defines the local orientation of the fast and slow polarization states (aka. *eigenmode structure*) along the transmission link. Due to the normalization, the birefringence matrix $\tilde{\boldsymbol{T}}(z,\omega) \cdot \vec{\boldsymbol{\sigma}}$ is unitary. In turn, the birefringence parameter $\Delta\beta(z,\omega) \stackrel{\text{def}}{=} \|\vec{\boldsymbol{T}}(z,\omega)\|$ quantifies the local amount of phase mismatch accumulated per unit length between the fast and slow polarization axis.

The general, Hermitian transmission matrix in Jones space is now expressed in terms of propagation constant, birefringence parameter, and (normalized) birefringence matrix via

$$\boldsymbol{T}(z,\omega) = \beta(z,\omega)\mathbf{I} + \Delta\beta(z,\omega)\; \tilde{\boldsymbol{T}}(z,\omega) \cdot \vec{\boldsymbol{\sigma}} \quad \in \mathbb{C}^{2\times 2}\,. \tag{A.17}$$

Solution for Lumped Elements in Jones Space To simplify the analysis, we now assume discrete optical elements with homogeneous properties in spatial direction.

If we assume $\boldsymbol{T}(z,\omega)$ to be invariant under z, denoted by $\boldsymbol{T}_{[1]}(\omega)$ to express a *single* short fiber segment[5], then the PDE in (A.1) simplifies to

$$\frac{\partial}{\partial z}\boldsymbol{U}(z,\omega) = -\mathrm{j}\,\boldsymbol{T}_{[1]}(\omega)\boldsymbol{U}(z,\omega)\,, \tag{A.18}$$

[5] A bracketed sub- or superscript index denotes a (labeled/numbered) discrete fiber segment invariant under z. The superscript is used for scalar variables to avoid confusion with vector elements or Taylor coefficients. Notation borrowed from [Win09].

which can be solved using the *matrix exponential* from (2.11) as [GK00, Eq. (A.12)]

$$\boldsymbol{U}(z,\omega) = \exp(-\mathrm{j}\,\boldsymbol{T}_{[1]}(\omega)\,z)\;\boldsymbol{U}(0,\omega) \tag{A.19}$$

$$= \underbrace{\exp(-\mathrm{j}\,\beta^{[1]}(\omega)\,\mathbf{I}\,z)}_{\substack{\text{polarization-\underline{in}dependent}\\ \text{phase rotation}}}\;\underbrace{\exp(-\mathrm{j}\Delta\beta^{[1]}(\omega)\,\tilde{\boldsymbol{T}}_{[1]}(\omega)\cdot\vec{\boldsymbol{\sigma}}\,z)}_{\substack{\text{polarization-dependent}\\ \text{phase rotation}}}\;\boldsymbol{U}(0,\omega)\,, \tag{A.20}$$

where $\boldsymbol{U}(0,\omega)$ is the optical signal at the *input* of the short fiber segment. The *linear channel transfer function* of a homogeneous fiber segment of length z is defined as

$$\boldsymbol{H}_{[1]}(z,\omega) \stackrel{\text{def}}{=} \exp(-\mathrm{j}\,\boldsymbol{T}_{[1]}(\omega)\,z) \in \mathbb{C}^{2\times 2}\,. \tag{A.21}$$

Since $\boldsymbol{T}_{[1]}(\omega)$ is by assumption *Hermitian*, we find that $\boldsymbol{H}_{[1]}(z,\omega)$ is *unitary* [AW05, Sec. 3.4].

The birefringence matrix $\exp(-\mathrm{j}\,\Delta\beta^{[1]}(\omega)\tilde{\boldsymbol{T}}_{[1]}(\omega)\cdot\vec{\boldsymbol{\sigma}}z)$ belongs to the group of *special*[6] *unitary transformations*, $\mathsf{SU}(2)$, and is isomorphic to the *unit* quaternion-valued algebra [KP04, Hal15], i.e., quaternion-valued rotations with *unit* magnitude (aka. *versors*) maintaining the norm during the transformation.

The *retardation angle* $\theta^{[1]}(z,\omega)$ of a single piece-wise constant fiber segment of length z is defined as

$$\frac{\theta^{[1]}(z,\omega)}{2} \stackrel{\text{def}}{=} \|\vec{\boldsymbol{T}}_{[1]}(\omega)\|\,z \stackrel{\text{def}}{=} \Delta\beta^{[1]}(\omega)\,z = \underbrace{\Delta\beta_0^{[1]}}_{\textit{lumped}\text{ birefringence}} z \quad + \quad \underbrace{\Delta\beta_1^{[1]}\,\omega}_{\textit{lumped}\text{ first-order PMD}} z \quad + \ldots\;, \tag{A.22}$$

which measures the (frequency-dependent) differential phase shift applied to the signal in the eigenstate of the birefringent element, see below. Here, the terminology *lumped* is used to highlight that this type of birefringence and PMD is caused by a lumped (i.e., discrete) optical element rather than a z-continuous medium. The transition to a continuous medium is done by cascading lumped fiber segments and letting the length of each segment become infinitesimal small, see discussion on the general solution below.

We can now use an extension of Euler's formula[7] [VB02, Eq. (1)] [Kar14, Eq. (8)] to give an explicit expression for the polarization rotation matrix in Jones space as

$$\exp\left(-\mathrm{j}\,\frac{\theta^{[1]}(z,\omega)}{2}\,\tilde{\boldsymbol{T}}_{[1]}(\omega)\cdot\vec{\boldsymbol{\sigma}}\right) = \cos\left(\frac{\theta^{[1]}(z,\omega)}{2}\right)\mathbf{I} - \mathrm{j}\sin\left(\frac{\theta^{[1]}(z,\omega)}{2}\right)\tilde{\boldsymbol{T}}_{[1]}(\omega)\cdot\vec{\boldsymbol{\sigma}}\,. \tag{A.23}$$

Since the birefringence matrix $\tilde{\boldsymbol{T}}_{[1]}(\omega)\cdot\vec{\boldsymbol{\sigma}}$ is Hermitian, it can also be diagonalized using the *similarity transform* $\boldsymbol{V}^{\mathsf{H}} = \boldsymbol{V}^{-1}$ with

$$-\,\mathrm{j}\,\frac{\theta^{[1]}(z,\omega)}{2}\,\tilde{\boldsymbol{T}}_{[1]}(\omega)\cdot\vec{\boldsymbol{\sigma}} = \boldsymbol{V}_{[1]}(\omega)\begin{bmatrix}-\mathrm{j}\frac{\theta^{[1]}(z,\omega)}{2} & 0\\ 0 & \mathrm{j}\frac{\theta^{[1]}(z,\omega)}{2}\end{bmatrix}\boldsymbol{V}_{[1]}^{\mathsf{H}}(\omega)\,, \tag{A.24}$$

[6]The set of all *unitary* transformations $\{\,\boldsymbol{M}\in\mathbb{C}^{2\times2}\mid \boldsymbol{M}^{-1}=\boldsymbol{M}^{\mathsf{H}}\,\}$ is given by the group $\mathsf{U}(2)$, the subset with *determinant* $\det(\boldsymbol{M})=+1$ is called *special* unitary group $\mathsf{SU}(2)$.

[7]An equivalent expression of Euler's formula can also be given for quaternions.

where the eigenvalues $\pm\mathrm{j}\,\theta^{[1]}(z,\omega)/2$ are the elements of the diagonal matrix, and the columns of $\boldsymbol{V}_{[1]}(\omega)$ are the corresponding eigenvectors. We can alternatively recover the rotation matrix decomposed in terms of eigenvectors using the relation from (2.13) for diagonalizable matrices. We find

$$\exp\left(-\mathrm{j}\,\frac{\theta^{[1]}(z,\omega)}{2}\,\tilde{\boldsymbol{T}}_{[1]}(\omega)\cdot\vec{\boldsymbol{\sigma}}\right)=\boldsymbol{V}_{[1]}(\omega)\begin{bmatrix}\exp\left(-\mathrm{j}\frac{\theta^{[1]}(z,\omega)}{2}\right) & 0\\ 0 & \exp\left(\mathrm{j}\frac{\theta^{[1]}(z,\omega)}{2}\right)\end{bmatrix}\boldsymbol{V}^{\mathsf{H}}_{[1]}(\omega)\,, \tag{A.25}$$

with the unitary matrix describing the eigenmode structure of the birefringent element

$$\boldsymbol{V}_{[1]}(\omega)\stackrel{\text{def}}{=}\frac{1}{\sqrt{2+2\tilde{T}_1^{[1]}(\omega)}}\begin{bmatrix}1+\tilde{T}_1^{[1]}(\omega) & \tilde{T}_2^{[1]}(\omega)-\mathrm{j}\tilde{T}_3^{[1]}(\omega)\\ \tilde{T}_2^{[1]}(\omega)+\mathrm{j}\tilde{T}_3^{[1]}(\omega) & 1+\tilde{T}_1^{[1]}(\omega)\end{bmatrix}\,, \tag{A.26}$$

where we used (A.11) with $\det(\boldsymbol{V}_{[1]}(\omega))\stackrel{!}{=}1$.

The physical understanding of (A.25) is as follows. The optical field envelope $\boldsymbol{U}(0,\omega)$ is transformed from the x-y polarization state into the polarization eigenstate defined by $\tilde{\boldsymbol{T}}_{[1]}(\omega)$ using the similarity transform $\boldsymbol{V}^{\mathsf{H}}_{[1]}(\omega)=\boldsymbol{V}^{-1}_{[1]}(\omega)$. There, both polarization states experience equal but opposite phase shifts by $\pm\theta^{[1]}(z,\omega)/2$, after which the signal is transformed back into the original x-y polarization state by $\boldsymbol{V}_{[1]}(\omega)$. The imposed differential phase shift $\theta^{[1]}(z,\omega)$ and the eigenmode structure $\boldsymbol{V}_{[1]}(\omega)$ are in general frequency-dependent.

If only the zero$^{\text{th}}$-order term in the Taylor expansion of $\Delta\beta^{[1]}(\omega)$ is considered (e.g., a birefringent device like a *waveplate*) and $\boldsymbol{V}_{[1]}(\omega)$ is invariant under ω, then the channel matrix becomes periodic in z under the condition $\theta^{[1]}(L_{\mathrm{b}}^{[1]},\omega)/2=\Delta\beta_0^{[1]}L_{\mathrm{b}}^{[1]}\stackrel{!}{=}2\pi\ \forall\,\omega$, where we define the *beat length* of a lumped, birefringent element as

$$L_{\mathrm{b}}^{[1]}\stackrel{\text{def}}{=}\frac{2\pi}{\Delta\beta_0^{[1]}}\,. \tag{A.27}$$

The beat length $L_{\mathrm{b}}^{[1]}$ measures the distance after which the signal $\boldsymbol{U}(z,\omega)$ arrives at its original polarization state after traveling in a z-invariant, frequency-flat birefringent fiber segment with birefringence parameter $\Delta\beta_0^{[1]}$.

If instead the first-order PMD term is considered, i.e., $\theta^{[1]}(z,\omega)/2=\Delta\beta_1^{[1]}\,\omega z$, a linear phase response proportional to the length of the segment z is applied on the diagonal matrix in (A.25). The linear phase response with opposite signs in frequency domain corresponds to a differential group delay in time domain between the polarization eigenstates. Hence, if first or higher-order terms are considered in the Taylor expansion of $\Delta\beta^{[1]}(\omega)$, the polarization eigenstates of a (single, discrete) fiber segment will disperse, which is known as *polarization-mode dispersion* due to the frequency-dependency of the birefringent material.

A different type of PMD is caused by the longitudinal variation of the birefringence parameter. E.g., a cascade of discrete, birefringent segments with a random orientation of the polarization eigenstates and (only) linear frequency-dependence of the birefringence parameter, will also result in dispersion of the polarization states.

General Solution in Jones Space If $\boldsymbol{T}(z,\omega)$ depends on z, i.e., it can not be treated as a discrete element, the solution to the PDE in (A.1) is not straightforward. This is apparent since concatenated piece-wise constant transformations $\boldsymbol{T}_{[i]}(\omega)$ with $i = 1,2,3,\ldots$ do not commute (cf. discussion in [GK00, Sec. 7] and [Win09, Sec. 2.2.2]). The solution can only be given by the *symbolic* expression [Fey51]

$$\boldsymbol{U}(z,\omega) = \exp\left(-\mathrm{j}\int_0^z \boldsymbol{T}(\zeta,\omega)\,\mathrm{d}\zeta\right)\,\boldsymbol{U}(0,\omega)\,, \tag{A.28}$$

where the dependence of $\boldsymbol{T}(z,\omega)$ on the spatial parameter z also indicates the *order* of the *operators*, here $\boldsymbol{T}(\zeta,\omega)$, under the integral [Fey51, Kor02].

We define the *local birefringence coordinate transformation* as

$$\boldsymbol{H}_{\mathrm{BIR}}(z) \overset{\text{def}}{=} \exp\left(-\mathrm{j}\int_0^z \Delta\beta_0(\zeta)\,\tilde{\boldsymbol{T}}(\zeta,0)\cdot\vec{\boldsymbol{\sigma}}\,\mathrm{d}\zeta\right)\,, \tag{A.29}$$

which includes only the zero$^{\text{th}}$-order term in the Taylor expansion of $\Delta\beta(z,\omega)$. To model the optical field envelope $\boldsymbol{U}(z,\omega)$ in the ECB domain, the transformation $\boldsymbol{H}^{\mathsf{H}}_{\mathrm{BIR}}(z)$ is applied to the original signal in order to remove the (rapid) motion of the signal's polarization state due to pure (i.e., zero$^{\text{th}}$-order) birefringence. We say that the ECB signal is modeled in the *local birefringence coordinate system*. The transformation does not alter the characteristics of, e.g., the optical noise process or the nonlinear signal-to-signal interference.

Similarly, also higher-order terms, e.g., PMD can be removed from the (linear) propagation equation by an appropriate transformation of the coordinates. Then, the nonlinear source term in the PDE must be scaled properly, see, e.g., [MM06, Sec. II B].

System State of Polarization We summarize that, similar to the previous chapter, the baseband signal at the input $\boldsymbol{U}(0,\omega) \in \mathbb{C}^2$ is transformed using the (now z-dependent) channel matrix $\boldsymbol{H}(z,\omega) \in \mathbb{C}^{2\times 2}$ when propagating over a linear channel. If not stated otherwise, we assume that the transform $\boldsymbol{H}(z,\omega)$ maintains the *norm* of the signal, i.e., it is a unitary transformation with $\boldsymbol{H}^{-1}(z,\omega) = \boldsymbol{H}^{\mathsf{H}}(z,\omega)$.

The channel matrix is fully parameterized using the four independent real-valued coefficients

$$T_0(z,\omega) \in \mathbb{R}, \qquad \vec{\boldsymbol{T}}(z,\omega) = [T_1(z,\omega), T_2(z,\omega), T_3(z,\omega)]^{\mathsf{T}} \in \mathbb{R}^3\,. \tag{A.30}$$

Here, the first coefficient $T_0(z,\omega)$ is equivalent to the propagation constant $\beta(z,\omega)$ and describes a common, frequency-dependent phase rotation imposed on both vector components in Jones space. The latter one is the Stokes vector $\vec{\boldsymbol{T}}(z,\omega)$ which defines the orientation of a (local) eigenstate in Jones space via the Pauli matrix decomposition of the normalized vector $\tilde{\boldsymbol{T}}(z,\omega)$ (i.e., the *system's state of polarization*). The magnitude of the differential phase rotation $\theta(z,\omega)$ is given by the birefringence parameter $\Delta\beta(z,\omega) = \|\vec{\boldsymbol{T}}(z,\omega)\|$ and the length of the birefringent element z.

In the ECB model, we use the Taylor expansion of the propagation constant to remove the zero$^{\text{th}}$ and first-order common phase $\beta_0(z)$ and $\beta_1(z)$, as well as the zero$^{\text{th}}$-order differential phase $\Delta\beta_0(z)$ from the propagation equation by a suited transformation of the coordinate system.

Signal State of Polarization Up to now, only the *system* was associated with a state of polarization. In order to develop a better understanding of polarization rotations, we will now introduce the notion of the *signal's state of polarization.*

The optical signal in Jones space $\boldsymbol{u}(z,t) = [u_\mathsf{x}(z,t), u_\mathsf{y}(z,t)]^\mathsf{T}$ can also be parameterized via Stokes parameters. The associated signal polarization describes the partition of the signal's energy (or equivalently *intensity*) in the x- and y-component and the phase difference between the two vector components[8]. The information about the common (i.e., *mean*) phase of the two vector components is lost in the Stokes space representation. The notion of polarization originates from *observable states*, i.e., measurable intensities [Sto51], proportional to the squared magnitude of the electric field in the x- and y-component. The set of coefficients parametrizing the signal polarization are collected in the *signal's Stokes vector*

$$\vec{\boldsymbol{u}}(z,t) = [u_1(z,t), u_2(z,t), u_3(z,t)]^\mathsf{T} \in \mathbb{R}^3 \tag{A.31}$$

$$\circ\!\!-\!\!\bullet$$

$$\vec{\boldsymbol{U}}(z,\omega) = [U_1(z,\omega), U_2(z,\omega), U_3(z,\omega)]^\mathsf{T} \in \mathbb{R}^3 \,. \tag{A.32}$$

The relation between Jones and Stokes space can be established by the concise (symbolic) expression [Fan54]

$$\vec{\boldsymbol{u}} = \boldsymbol{u}^\mathsf{H} \vec{\boldsymbol{\sigma}} \boldsymbol{u}\,, \tag{A.33}$$

to denote the element-wise operation $u_i = \boldsymbol{u}^\mathsf{H} \boldsymbol{\sigma}_i \boldsymbol{u}$ for all Stokes vector components $i = 1, 2, 3$. Here and in the following, the spatial and temporal dependency $\vec{\boldsymbol{u}}(z,t)$ is implied. The same analysis also applies to the frequency-domain Stokes vector $\vec{\boldsymbol{U}}(z,\omega)$.

Using (A.33), we recover with the zero$^{\text{th}}$ order Pauli matrix $\boldsymbol{\sigma}_0$ the squared norm $\|\boldsymbol{u}\|^2$ and define the zero$^{\text{th}}$ order *Stokes parameter* of the signal as

$$u_0 \overset{\text{def}}{=} \boldsymbol{u}^\mathsf{H} \boldsymbol{\sigma}_0 \boldsymbol{u} = \|\boldsymbol{u}\|^2 = |u_\mathsf{x}|^2 + |u_\mathsf{y}|^2 = u_\mathsf{x} u_\mathsf{x}^* + u_\mathsf{y} u_\mathsf{y}^* \;\in \mathbb{R}\,. \tag{A.34}$$

The remaining Stokes parameters are given accordingly. The first Stokes parameter is given by

$$u_1 \overset{\text{def}}{=} \boldsymbol{u}^\mathsf{H} \boldsymbol{\sigma}_1 \boldsymbol{u} = \boldsymbol{u}^\mathsf{H} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \boldsymbol{u} = |u_\mathsf{x}|^2 - |u_\mathsf{y}|^2 = u_\mathsf{x} u_\mathsf{x}^* - u_\mathsf{y} u_\mathsf{y}^* \;\in \mathbb{R}\,, \tag{A.35}$$

being a measure of whether the signal energy is concentrated more in the x- or in the y-polarization. The second Stokes parameter is given by

$$u_2 \overset{\text{def}}{=} \boldsymbol{u}^\mathsf{H} \boldsymbol{\sigma}_2 \boldsymbol{u} = \boldsymbol{u}^\mathsf{H} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \boldsymbol{u} = u_\mathsf{x} u_\mathsf{y}^* + u_\mathsf{x}^* u_\mathsf{y} = 2\,\mathrm{Re}\{u_\mathsf{x} u_\mathsf{y}^*\} \;\in \mathbb{R}\,, \tag{A.36}$$

being a measure of $+45°$- over $-45°$-polarized light, and the third Stokes parameter

$$u_3 \overset{\text{def}}{=} \boldsymbol{u}^\mathsf{H} \boldsymbol{\sigma}_3 \boldsymbol{u} = \boldsymbol{u}^\mathsf{H} \begin{bmatrix} 0 & -\mathrm{j} \\ \mathrm{j} & 0 \end{bmatrix} \boldsymbol{u} = \mathrm{j}(u_\mathsf{x} u_\mathsf{y}^* - u_\mathsf{x}^* u_\mathsf{y}) = -2\,\mathrm{Im}\{u_\mathsf{x} u_\mathsf{y}^*\} \;\in \mathbb{R}\,, \tag{A.37}$$

[8] The (time-resolved) *Stokes parameter* $\vec{\boldsymbol{u}}(z,t)$ will be defined to be *local* and *instantaneous*, i.e., as a function of z and t, in direct correspondence with the Jones vector $\boldsymbol{u}(z,t)$, i.e., without any averaging/expectation. Similarly, the (frequency-resolved) signal Stokes parameter $\vec{\boldsymbol{U}}(z,\omega)$ are given in terms of z and ω.

being a measure of *right* over *left* circularly-polarized light. We will again denote normalized Stokes parameters by the tilde, i.e.,

$$\tilde{\boldsymbol{u}} = [\tilde{u}_1, \tilde{u}_2, \tilde{u}_3]^\mathsf{T} \iff u_0 = \|\vec{\boldsymbol{u}}\| = \|\boldsymbol{u}\|^2 = 1\,. \tag{A.38}$$

The Stokes vector $\vec{\boldsymbol{u}}$ can also be expanded to obtain the equivalent Jones matrix description using the dot product with the Pauli vector. Then, the complex-valued 2×2 matrix is obtained using the reciprocal relation w.r.t. (A.33) by

$$\vec{\boldsymbol{u}}\cdot\vec{\boldsymbol{\sigma}} = u_1\boldsymbol{\sigma}_1 + u_2\boldsymbol{\sigma}_2 + u_3\boldsymbol{\sigma}_3 = \begin{bmatrix} u_1 & u_2 - \mathrm{j}u_3 \\ u_2 + \mathrm{j}u_3 & -u_1 \end{bmatrix} = \begin{bmatrix} u_\mathsf{x}u_\mathsf{x}^* - u_\mathsf{y}u_\mathsf{y}^* & 2u_\mathsf{x}u_\mathsf{y}^* \\ 2u_\mathsf{x}^*u_\mathsf{y} & u_\mathsf{y}u_\mathsf{y}^* - u_\mathsf{x}u_\mathsf{x}^* \end{bmatrix}\,. \tag{A.39}$$

In Section 3.3.5, we discuss a nonlinear process that causes a polarization rotation induced by the state of the signal itself. We will, e.g., recover expressions similar to (A.39) in the Manakov equation leading to rotations of the signal Stokes vector evaluated, e.g., at frequency $\vec{\boldsymbol{U}}(z,\omega_0)$ around the signals Stokes vector at another frequency at, e.g., $\vec{\boldsymbol{U}}(z,\omega_1)$.

We also have the useful equality [GK00, Eq. (3.9)] [Fan54]

$$\boldsymbol{u}\boldsymbol{u}^\mathsf{H} = \frac{1}{2}\left(\boldsymbol{u}^\mathsf{H}\boldsymbol{u}\,\mathbf{I} + \vec{\boldsymbol{u}}\cdot\vec{\boldsymbol{\sigma}}\right) = \begin{bmatrix} u_\mathsf{x}u_\mathsf{x}^* & u_\mathsf{x}u_\mathsf{y}^* \\ u_\mathsf{x}^*u_\mathsf{y} & u_\mathsf{y}u_\mathsf{y}^* \end{bmatrix}\,, \tag{A.40}$$

to relate the matrix $\boldsymbol{u}\boldsymbol{u}^\mathsf{H}$ (aka. *coherency matrix* [Kar14, Eq. (52)]) to the signal intensity $\|\boldsymbol{u}\|^2 = \boldsymbol{u}^\mathsf{H}\boldsymbol{u}$ and the Stokes parameters $\vec{\boldsymbol{u}}$.

Polarization Rotations in Stokes Space Birefringence and PMD in Jones space can be best understood by the eigenvalue decomposition of the rotation matrix in (A.25), where a differential phase shift is applied to the signal rotated to the eigenstate of the birefringent optical element. In Stokes space, those effects have a direct geometrical interpretation, see below.

We will again use the simplified case of a lossless and discrete, optical element with homogeneous, spatial properties. In this case, the propagation equation takes the form

$$\frac{\partial}{\partial z}\boldsymbol{U}(z,\omega) = -\mathrm{j}\,\boldsymbol{T}_{[1]}(\omega)\,\boldsymbol{U}(z,\omega) \tag{A.41}$$

$$= -\mathrm{j}\,\frac{\theta^{[1]}(z,\omega)}{2}\tilde{\boldsymbol{T}}_{[1]}(\omega)\cdot\vec{\boldsymbol{\sigma}}\;\boldsymbol{U}(z,\omega)\,, \tag{A.42}$$

where we use the Pauli matrix decomposition of the rotation matrix $\boldsymbol{T}_{[1]}(\omega)$.

We will now use the relation $\vec{\boldsymbol{u}} = \boldsymbol{u}^\mathsf{H}\vec{\boldsymbol{\sigma}}\boldsymbol{u}$ from (A.33) to translate the propagation equation of the signal in Jones space to the propagation equation of the corresponding Stokes vector. The evolution of the Stokes vector is described by

$$\frac{\partial}{\partial z}\vec{\boldsymbol{U}}(z,\omega) = \boldsymbol{U}^\mathsf{H}(z,\omega)\;\mathrm{j}\left(\boldsymbol{T}_{[1]}^\mathsf{H}(\omega)\vec{\boldsymbol{\sigma}} - \vec{\boldsymbol{\sigma}}\boldsymbol{T}_{[1]}(\omega)\right)\;\boldsymbol{U}(z,\omega) \tag{A.43}$$

$$= \boldsymbol{U}^\mathsf{H}(z,\omega)\;\mathrm{j}\frac{\theta^{[1]}(z,\omega)}{2}\left((\tilde{\boldsymbol{T}}_{[1]}(\omega)\cdot\vec{\boldsymbol{\sigma}})\,\vec{\boldsymbol{\sigma}} - \vec{\boldsymbol{\sigma}}(\tilde{\boldsymbol{T}}_{[1]}(\omega)\cdot\vec{\boldsymbol{\sigma}})\right)\;\boldsymbol{U}(z,\omega)\,, \tag{A.44}$$

where we use the product rule of calculus $\frac{\partial}{\partial z}\vec{\boldsymbol{u}} = (\frac{\partial}{\partial z}\boldsymbol{u}^\mathsf{H})\vec{\boldsymbol{\sigma}}\boldsymbol{u} + \boldsymbol{u}^\mathsf{H}\vec{\boldsymbol{\sigma}}(\frac{\partial}{\partial z}\boldsymbol{u})$, and the two properties $(\boldsymbol{T}\boldsymbol{U})^\mathsf{H} = \boldsymbol{U}^\mathsf{H}\boldsymbol{T}^\mathsf{H}$ and $\boldsymbol{T}^\mathsf{H} = \boldsymbol{T}$ of the Hermitian conjugate.

We now define the *cross-product operator*[9] [GK00, Eq. (4.8)]

$$\tilde{\boldsymbol{T}}\times \stackrel{\text{def}}{=} \begin{bmatrix} 0 & -\tilde{T}_3 & \tilde{T}_2 \\ \tilde{T}_3 & 0 & -\tilde{T}_1 \\ -\tilde{T}_2 & \tilde{T}_1 & 0 \end{bmatrix} = \left[\tilde{T}\times_{n,m}\right] \in \mathbb{R}^{3\times 3}, \quad n = 1,2,3, \quad m = 1,2,3\,, \tag{A.45}$$

where the components $\tilde{T}_i$ with $i = 1,2,3$ are the elements of the normalized birefringence vector $\tilde{\boldsymbol{T}} = \vec{\boldsymbol{T}}/\|\vec{\boldsymbol{T}}\|$. This allows us to formulate the spin vector rules [GK00, Eqs. (A.3), (A.4), and (A.13)] as

$$\vec{\boldsymbol{\sigma}}(\tilde{\boldsymbol{T}} \cdot \vec{\boldsymbol{\sigma}}) = \tilde{\boldsymbol{T}}\,\mathbf{I} + \mathrm{j}\,\tilde{\boldsymbol{T}}\times \vec{\boldsymbol{\sigma}} \tag{A.46}$$

$$(\tilde{\boldsymbol{T}} \cdot \vec{\boldsymbol{\sigma}})\vec{\boldsymbol{\sigma}} = \tilde{\boldsymbol{T}}\,\mathbf{I} - \mathrm{j}\,\tilde{\boldsymbol{T}}\times \vec{\boldsymbol{\sigma}} \tag{A.47}$$

$$\boldsymbol{U}^{\mathsf{H}}\;(\tilde{\boldsymbol{T}}\times \vec{\boldsymbol{\sigma}})\;\boldsymbol{U} = \tilde{\boldsymbol{T}}\times \vec{\boldsymbol{U}}\,, \tag{A.48}$$

which are required to relate the rotations in complex-valued Jones space to rotations in real-valued Stokes space (both with three degrees of freedom). The spin vector rules are again symbolic expressions which are interpreted element-wise w.r.t. the Pauli vector $\vec{\boldsymbol{\sigma}} = [\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, \boldsymbol{\sigma}_3]^{\mathsf{T}}$. E.g., the spin vector rule in (A.46) can be read as

$$\boldsymbol{\sigma}_i(\tilde{T}_1\boldsymbol{\sigma}_1 + \tilde{T}_2\boldsymbol{\sigma}_2 + \tilde{T}_3\boldsymbol{\sigma}_3) = \tilde{T}_i\,\mathbf{I} + \mathrm{j}\,(\tilde{T}\times_{i,1}\,\boldsymbol{\sigma}_1 + \tilde{T}\times_{i,2}\,\boldsymbol{\sigma}_2 + \tilde{T}\times_{i,3}\,\boldsymbol{\sigma}_3)\,, \tag{A.49}$$

with $i = 1,2,3$. Here, the operator $\tilde{\boldsymbol{T}}\times$ induces a tensor multiplication $(\mathbb{R}^{3\times 3})(\mathbb{C}^{2\times 2} \otimes \mathbb{R}^3) \mapsto \mathbb{C}^{2\times 2} \otimes \mathbb{R}^3$ where each component of the resulting object is given by a sum of the Pauli matrices weighted with the elements of the cross-product operator.

Using the spin vector rules, the Stokes space propagation equation of the signal SOP can be written as [GK00, Eq. (6.8)]

$$\frac{\partial}{\partial z}\vec{\boldsymbol{U}}(z,\omega) = \theta^{[1]}(z,\omega)\,\tilde{\boldsymbol{T}}_{[1]}(\omega)\times\,\vec{\boldsymbol{U}}(z,\omega)\,, \tag{A.50}$$

and integration over z yields

$$\vec{\boldsymbol{U}}(z,\omega) = \exp\left(\theta^{[1]}(z,\omega)\,\tilde{\boldsymbol{T}}_{[1]}(\omega)\times\right)\;\vec{\boldsymbol{U}}(0,\omega)\,. \tag{A.51}$$

where $\exp(\cdot)$ denotes the matrix exponential from (2.11). The geometrical interpretation of the latter is as follows. The Stokes vector of the signal at the input $\vec{\boldsymbol{U}}(0,\omega)$ and the unit length Stokes vector $\tilde{\boldsymbol{T}}_{[1]}(\omega)$ are both pictured as vectors in Stokes space. In particular, a unit length (i.e., normalized) Stokes vector is a point on the Poincaré sphere. The evolution of the signal's Stokes vector is described by a rotation on a circle around the birefringence vector $\tilde{\boldsymbol{T}}_{[1]}(\omega)$, while the arc length of the rotation is given by the angle $\theta^{[1]}(z,\omega)$. An illustrative example is given in the following.

[9]The notation $\tilde{\boldsymbol{T}}\times$ of the real-valued 3×3 operator is common in optics and kept here to be consistent with the literature.

Example A.1: Birefringence in Jones and Stokes space

In this academic example, we consider a discrete optical element with homogeneous properties in the z-direction. In particular, we will consider a birefringent device, a so-termed *half-wave plate*, which induces a differential phase shift of $\theta^{[1]}(z_0,\omega)=\pi$ to the input signal $U(0,\omega)$ after it has been rotated to its eigenbasis characterized by the Stokes vector $\vec{\boldsymbol{T}}_{[1]}(\omega)$.
We start with the Stokes vector of the half waveplate, arbitrarily set to

$$\vec{\boldsymbol{T}}_{[1]}(\omega)=[T_1^{[1]}(\omega),T_2^{[1]}(\omega),T_3^{[1]}(\omega)]=[1,2,3]^{\mathsf{T}}\in\mathbb{R}^3,\quad\forall\,\omega\,,$$

which is frequency-flat for all ω, i.e., we only consider lumped birefringence without any first- or higher-order PMD terms (e.g., compare with (A.22)). The coefficients $T_i^{[1]}(\omega)$ with $i=1,2,3$ have units m^{-1}. We also assume that the propagation constant is equal to zero, i.e., $T_0^{[1]}(\omega)=\beta^{[1]}(\omega)=0$. The spatial extent of the half-wave plate is fixed to $z_1=0.42\,\mathrm{m}$, which is required to accumulate the differential phase shift of π, see below.
The norm of $\vec{\boldsymbol{T}}_{[1]}$ calculates to $\|\vec{\boldsymbol{T}}_{[1]}\|=\Delta\beta_0^{[1]}=\sqrt{1^2+2^2+3^2}\,\mathrm{m}^{-1}=\sqrt{14}\,\mathrm{m}^{-1}$. We calculate the normalized Stokes parameter, i.e., the SOP, characterizing the orientation of the half-wave plate as

$$\tilde{\boldsymbol{T}}_{[1]}(\omega)=[\tilde{T}_1^{[1]}(\omega),\tilde{T}_2^{[1]}(\omega),\tilde{T}_3^{[1]}(\omega)]=[0.27,0.54,0.80]^{\mathsf{T}}\in\mathbb{R}^3,\quad\forall\omega\,.$$

The optical field envelope at the input $z=0$ is arbitrarily set to

$$\boldsymbol{U}(0,\omega)=[1+\exp(-\mathrm{j}\pi/3),\ \exp(\mathrm{j}\pi/3)]^{\mathsf{T}}\ \delta(\omega-\omega_1)\in\mathbb{C}^2\,,$$

corresponding to a continuous wave (CW) with frequency $\omega_1=2\pi f_1$ modeled in the ECB. The SOP of the optical field envelope at the input is obtained via (A.33) as

$$\tilde{\boldsymbol{U}}(0,\omega)=[\tilde{U}_1(0,\omega),\tilde{U}_2(0,\omega),\tilde{U}_3(0,\omega)]=[0.50,0,0.87]^{\mathsf{T}}\ \delta(\omega-\omega_1)\in\mathbb{R}^3\,,$$

where the polarization state of the signal is in a mixed state containing x-polarized light (into the $+\vec{\mathbf{S}}_1$ direction) and right-circular polarized light (into the $+\vec{\mathbf{S}}_3$ direction). Both the SOP of the birefringent half-wave plate and the SOP of the input signal are depicted on the Poincaré sphere shown in Fig. A.1.
Using (A.22), we calculate the retardation angle to

$$\theta^{[1]}(z_1,\omega)=2\|\vec{\boldsymbol{T}}_{[1]}(\omega)\|\,z_1\approx\pi,\quad\forall\omega\,,$$

which gives the π phase shift inherent to the half-wave plate.
The channel matrix represents the transmission characteristic of the half-wave plate (ignoring attenuation and dispersion). It computes to

$$\begin{aligned}\boldsymbol{H}_{[1]}(z_1,\omega)&=\exp\left(-\mathrm{j}\,\frac{\theta^{[1]}(z_1,\omega)}{2}\,\tilde{\boldsymbol{T}}_{[1]}(\omega)\cdot\vec{\boldsymbol{\sigma}}\right)\\&=\cos\left(\frac{\theta^{[1]}(z_1,\omega)}{2}\right)\mathbf{I}-\mathrm{j}\sin\left(\frac{\theta^{[1]}(z_1,\omega)}{2}\right)\tilde{\boldsymbol{T}}_{[1]}(\omega)\cdot\vec{\boldsymbol{\sigma}}\\&=\begin{bmatrix}-0.27\mathrm{j}&-0.80-0.54\mathrm{j}\\0.80-0.54\mathrm{j}&0.27\mathrm{j}\end{bmatrix}.\end{aligned}$$

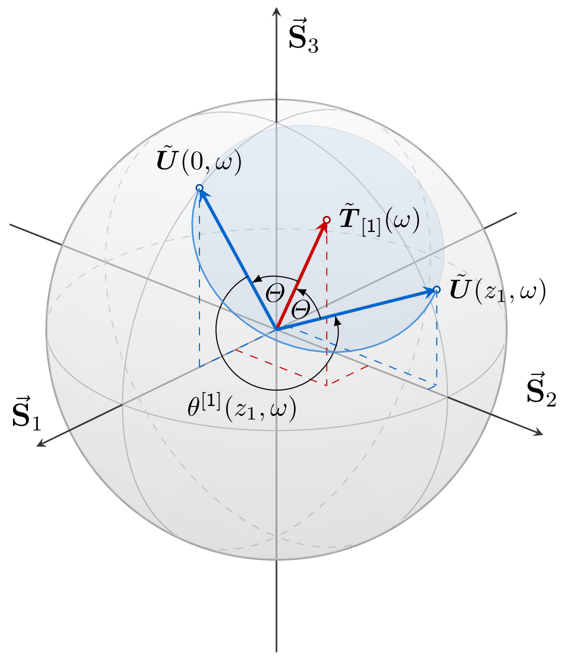


Figure A.1: Rotation of the input signal's state of polarization $\tilde{\boldsymbol{U}}(0,\omega)$ around the birefringent optical element defined by the Stokes vector $\tilde{\boldsymbol{T}}_{[1]}(\omega)$. The evolution traces out a circle on the Poincaré sphere with arc length equal to the retardation angle $\theta^{[1]}(z_1,\omega) \approx \pi$. In Jones space, this angle is equivalent to the differential phase applied to the signal rotated to the eigenstate of the polarizer, here, a *half-wave plate* due to the π phase shift. In this particular case, the signal is misaligned w.r.t. the half-wave plate by an angle of $\Theta \approx 34.1°$.

Using the singular value decomposition from (A.25), we can also express the channel matrix as

$$\begin{aligned}\boldsymbol{V}_{[1]}(\omega) &\begin{bmatrix} \exp\left(-\mathrm{j}\,\frac{\theta^{[1]}(z_1,\omega)}{2}\right) & 0 \\ 0 & \exp\left(+\mathrm{j}\,\frac{\theta^{[1]}(z_1,\omega)}{2}\right)\end{bmatrix} \boldsymbol{V}^{\mathsf{H}}_{[1]}(\omega) \\ &= \begin{bmatrix} 0.80 & -0.34{+}0.50\mathrm{j} \\ 0.34{+}0.50\mathrm{j} & 0.80 \end{bmatrix} \begin{bmatrix} \mathrm{e}^{-\mathrm{j}\pi/2} & 0 \\ 0 & \mathrm{e}^{+\mathrm{j}\pi/2}\end{bmatrix} \begin{bmatrix} 0.80 & 0.34{-}0.50\mathrm{j} \\ -0.34{-}0.50\mathrm{j} & 0.80 \end{bmatrix},\end{aligned}$$

where the eigenvalues, i.e., phase shifts with opposite signs are the components of the diagonal matrix. The first column of the similarity transform $\boldsymbol{V}_{[1]}(\omega)$ describes the eigenstate of the half-wave plate in Jones space aligned with the slow axis (i.e., corresponding to the first eigenvalue), while the second column describes the orthogonal eigenstate aligned with the fast axis of the wave plate.

The signal's Stokes parameter at the output calculates to

$$\tilde{\boldsymbol{U}}(z_1,\omega) = \boldsymbol{H}_{[1]}(z_1,\omega)\tilde{\boldsymbol{U}}(0,\omega) \;= [-0.06, 0.89, 0.46]^{\mathsf{T}}\ \delta(\omega-\omega_1)\,,$$

which is also shown in Fig. A.1. It can be seen that the Stokes vector of the input signal $\tilde{\boldsymbol{U}}(0,\omega)$ evolves on a circle in a right-hand rotation around $\tilde{\boldsymbol{T}}_{[1]}(\omega)$. While propagation through the half-wave plate the differential phase shift $\theta^{[1]}(z,\omega)$ accumulates until at z_1 a phase shift of π is present.

The misalignment angle between the Stokes vector of the signal and the birefringent element is computed as

$$\Theta(\omega) = \mathrm{acos}(\tilde{\boldsymbol{T}}_{[1]}(\omega) \cdot \tilde{\boldsymbol{U}}(0,\omega)) = \mathrm{acos}(\tilde{\boldsymbol{T}}_{[1]}(\omega) \cdot \tilde{\boldsymbol{U}}(z_1,\omega)) \approx 34.1^\circ \; \delta(\omega-\omega_1)\,.$$

Using the Stokes formalism, the rotation of the signal SOP around $\tilde{\boldsymbol{T}}_{[1]}(\omega)$ can also be directly performed in the 3D real-valued Stokes space. The cross-product operator from (A.45) is given by

$$\tilde{\boldsymbol{T}}_{[1]}(\omega)\times = \begin{bmatrix} 0 & -0.80 & 0.54 \\ 0.80 & 0 & -0.27 \\ -0.54 & 0.27 & 0 \end{bmatrix}, \quad \forall\omega\,,$$

and the Stokes vector of the output signal follows by

$$\begin{aligned}\tilde{\boldsymbol{U}}(z_1,\omega) &= \exp\left(\theta^{[1]}(z_1,\omega)\; \tilde{\boldsymbol{T}}_{[1]}(\omega)\times\right) \tilde{\boldsymbol{U}}(0,\omega) \\ &= \begin{bmatrix} -0.86 & 0.29 & 0.43 \\ 0.29 & -0.43 & 0.86 \\ 0.43 & 0.86 & 0.29 \end{bmatrix} \begin{bmatrix} 0.5 \\ 0 \\ 0.87 \end{bmatrix} \delta(\omega-\omega_1) = \begin{bmatrix} -0.06 \\ 0.89 \\ 0.46 \end{bmatrix} \delta(\omega-\omega_1)\,,\end{aligned}$$

where the same result is obtained as before using the Jones formalism.

A.2 Proof of the Fourier Relation in (4.62), (4.63)

In this appendix we compute the Fourier transform of $\Delta\boldsymbol{s}(t)$ in (4.63) similar to Ablowitz *et al.* in [AH02a, Appx.].

We start our derivation by expressing the optical field envelope $\boldsymbol{u}(0,t)$ by its inverse Fourier transform of $\boldsymbol{U}(0,\omega)$ to obtain[10]

$$\begin{aligned}
\Delta\boldsymbol{s}(t) =& -\mathrm{j}\bar{\gamma}\frac{8}{9}L_{\text{eff}}\int_{\mathbb{R}^2} h_{\text{NL}}(\tau_1,\tau_2) \\
&\times \boldsymbol{u}(0,t+\tau_1)\boldsymbol{u}^{\mathsf{H}}(0,t+\tau_1+\tau_2)\boldsymbol{u}(t+\tau_2)\,\mathrm{d}^2\boldsymbol{\tau} \\
=& -\mathrm{j}\bar{\gamma}\frac{8}{9}L_{\text{eff}}\frac{1}{(2\pi)^3}\iint_{-\infty}^{+\infty}\mathrm{d}\tau_1\mathrm{d}\tau_2\; h_{\text{NL}}(\tau_1,\tau_2) \\
&\times \int_{-\infty}^{\infty}\mathrm{d}\omega_3\,\boldsymbol{U}(0,\omega_3)\exp(\mathrm{j}\omega_3\tau_1) \\
&\times \int_{-\infty}^{\infty}\mathrm{d}\omega_2\,\boldsymbol{U}^{\mathsf{H}}(0,\omega_2)\exp(-\mathrm{j}\omega_2(\tau_1+\tau_2)) \\
&\times \int_{-\infty}^{\infty}\mathrm{d}\omega_1\,\boldsymbol{U}(0,\omega_1)\exp(\mathrm{j}\omega_1\tau_2) \\
&\times \exp(\mathrm{j}(\omega_3-\omega_2+\omega_1)t)\,.
\end{aligned} \tag{A.52}$$

The Fourier transform of the former expression yields

$$\begin{aligned}
\Delta\boldsymbol{S}(\omega) =& -\mathrm{j}\bar{\gamma}\frac{8}{9}L_{\text{eff}}\frac{1}{(2\pi)^3}\iiint_{-\infty}^{+\infty}\mathrm{d}t\,\mathrm{d}\tau_1\mathrm{d}\tau_2\; h_{\text{NL}}(\tau_1,\tau_2) \\
&\times \int_{-\infty}^{\infty}\mathrm{d}\omega_3\,\boldsymbol{U}(0,\omega_3)\exp(\mathrm{j}\omega_3\tau_1) \\
&\times \int_{-\infty}^{\infty}\mathrm{d}\omega_2\,\boldsymbol{U}^{\mathsf{H}}(0,\omega_2)\exp(-\mathrm{j}\omega_2(\tau_1+\tau_2)) \\
&\times \int_{-\infty}^{\infty}\mathrm{d}\omega_1\,\boldsymbol{U}(0,\omega_1)\exp(\mathrm{j}\omega_1\tau_2) \\
&\times \exp(\mathrm{j}(\omega_3-\omega_2+\omega_1-\omega)t)\,.
\end{aligned} \tag{A.53}$$

We now use the identity $\int_{-\infty}^{\infty}\exp(\mathrm{j}(\omega_3-\omega_2+\omega_1-\omega)t)\,\mathrm{d}t = 2\pi\,\delta(\omega_3-\omega_2+\omega_1-\omega)$ which a manifestation of the *frequency matching* constraint in (4.22). Applying the *sifting property* of the dirac function $\int_{\mathbb{R}} f(y)\delta(x-y)\,\mathrm{d}y = f(x)$, we obtain

$$\begin{aligned}
\Delta\boldsymbol{S}(\omega) =& -\mathrm{j}\bar{\gamma}\frac{8}{9}L_{\text{eff}}\frac{1}{(2\pi)^2}\iint_{-\infty}^{+\infty}\mathrm{d}\tau_1\mathrm{d}\tau_2\; h_{\text{NL}}(\tau_1,\tau_2) \\
&\times \boldsymbol{U}(0,\omega-\omega_1+\omega_2)\exp(\mathrm{j}(\omega-\omega_1+\omega_2)\tau_1) \\
&\times \int_{-\infty}^{\infty}\mathrm{d}\omega_2\,\boldsymbol{U}^{\mathsf{H}}(0,\omega_2)\exp(-\mathrm{j}\omega_2(\tau_1+\tau_2)) \\
&\times \int_{-\infty}^{\infty}\mathrm{d}\omega_1\,\boldsymbol{U}(0,\omega_1)\exp(\mathrm{j}\omega_1\tau_2)\,.
\end{aligned} \tag{A.54}$$

[10]Here, we switch to the *prefix notation* $\int \mathrm{d}x\, f(x)$ commonly used in physics.

After re-arranging the order of integration, we have

$$\begin{aligned}\Delta\boldsymbol{S}(\omega) = &-\mathrm{j}\bar{\gamma}\frac{8}{9}L_{\mathrm{eff}}\frac{1}{(2\pi)^2}\iint_{-\infty}^{+\infty}\mathrm{d}\omega_1\mathrm{d}\omega_2 \\ &\times \boldsymbol{U}(0,\omega-\omega_1+\omega_2)\boldsymbol{U}^{\mathsf{H}}(0,\omega_2)\boldsymbol{U}(0,\omega_1) \\ &\times \iint_{-\infty}^{\infty}\mathrm{d}\tau_1\mathrm{d}\tau_2\, h_{\mathrm{NL}}(\tau_1,\tau_2)\exp(\mathrm{j}\omega_1\tau_2) \\ &\times \exp(-\mathrm{j}\omega_2(\tau_1+\tau_2))\exp(\mathrm{j}(\omega-\omega_1+\omega_2)\tau_1)\,.\end{aligned} \tag{A.55}$$

And finally a change of variables with $\upsilon_1 = \omega_1 - \omega$ and $\upsilon_2 = \omega_2 - \omega_1$ yields

$$\begin{aligned}\Delta\boldsymbol{S}(\omega) = &-\mathrm{j}\bar{\gamma}\frac{8}{9}L_{\mathrm{eff}}\frac{1}{(2\pi)^2}\iint_{-\infty}^{+\infty}\mathrm{d}\upsilon_1\mathrm{d}\upsilon_2 \\ &\times \boldsymbol{U}(0,\omega+\upsilon_2)\boldsymbol{U}^{\mathsf{H}}(0,\omega+\upsilon_1+\upsilon_2)\boldsymbol{U}(0,\omega+\upsilon_1) \\ &\times \underbrace{\iint_{-\infty}^{\infty}\mathrm{d}\tau_1\mathrm{d}\tau_2\, h_{\mathrm{NL}}(\tau_1,\tau_2)\exp(-\mathrm{j}\upsilon_1\tau_1-\mathrm{j}\upsilon_2\tau_2)}_{H_{\mathrm{NL}}(\upsilon_1,\upsilon_2)=\mathcal{F}\{\,h_{\mathrm{NL}}(\tau_1,\tau_2)\,\}}\,,\end{aligned} \tag{A.56}$$

which is equivalent to the expression in (4.62). □

A.3 Alternative Derivation of the (Analog) Baseband End-to-End RP Method

In this appendix, we provide the solution to the received baseband perturbation $\Delta\boldsymbol{s}(t)$ in (4.63) using an alternative derivation based on the Volterra theory in *time domain*, see Section 2.1.3.2.

We start from the *parallel fiber model* shown in Figure A.2 which is a modified version of the known representation in Figure 4.2. Here, we assumed that the receiver performs chromatic dispersion compensation using the channel matched filter $H_{\mathrm{C}}^{*}(L,\omega)$ according to the optical receiver frond-end such that the received perturbation in baseband is given by $\Delta\boldsymbol{S}(\omega) = H_{\mathrm{C}}^{*}(L,\omega)\Delta\boldsymbol{U}(L,\omega)$.

The parallel fiber model can be understood as a nonlinear system where the output $\Delta\boldsymbol{s}(t)$ is given as the *sum* (in the limit $\zeta \to 0$, as the *integral*) of independent parallel branches, each a realization of a *basic third-order nonlinear system*[11] [Sch80, Fig. 4.4-1]. We will first derive the transfer characteristic of the basic third-order system which relates the optical signal at the input $\boldsymbol{u}(0,t)$ to the perturbation $\boldsymbol{u}(\zeta,t)$ associated with the position ζ along the link, i.e., one particular branch realization. In a second step, the solution of a single basic block is generalized to the continuum over all parallel branches.

We start with the output of the basic third-order system $\Delta\boldsymbol{u}(\zeta,t)$, i.e., the perturbation in the optical domain generated by the *local* nonlinear interaction $\left(\|\boldsymbol{u}_{\mathrm{LIN}}(\zeta,t)\|^{2}\,\boldsymbol{u}_{\mathrm{LIN}}(\zeta,t)\right)$. A block diagram of the basic third-order system is shown in Figure A.3. We can express $\Delta\boldsymbol{u}(\zeta,t)$ as the output of the *linear* system using the relation $H_{\mathrm{C}}^{-1}(\zeta,\omega) \bullet\!\!-\!\!\circ \frac{1}{2\pi|\mathcal{B}(\zeta)|} h_{\mathrm{C}}^{-1}(\zeta,t)$ from (3.66). We use the auxiliary variable σ to express the convolution as

$$\begin{aligned}\Delta\boldsymbol{u}(\zeta,t) &= \mathcal{F}^{-1}\{\,H_{\mathrm{C}}^{-1}(\zeta,\omega)\,\} * \left(\|\boldsymbol{u}_{\mathrm{LIN}}(\zeta,t)\|^{2}\,\boldsymbol{u}_{\mathrm{LIN}}(\zeta,t)\right) \\ &= \int_{-\infty}^{\infty} \frac{1}{2\pi|\mathcal{B}(\zeta)|} h_{\mathrm{C}}^{-1}(\zeta,\sigma)\,\boldsymbol{u}_{\mathrm{LIN}}(\zeta,t-\sigma)\boldsymbol{u}_{\mathrm{LIN}}^{\mathsf{H}}(\zeta,t-\sigma)\boldsymbol{u}_{\mathrm{LIN}}(\zeta,t-\sigma)\,\mathrm{d}\sigma\,.\end{aligned} \tag{A.57}$$

In the next step, the local signals at ζ are expressed as the input signals convolved with the

[11] Here, we use the terminology frequently used in *Volterra* theory. The third-order Volterra kernel corresponds to the first-order regular perturbation approach, cf. [VB02]

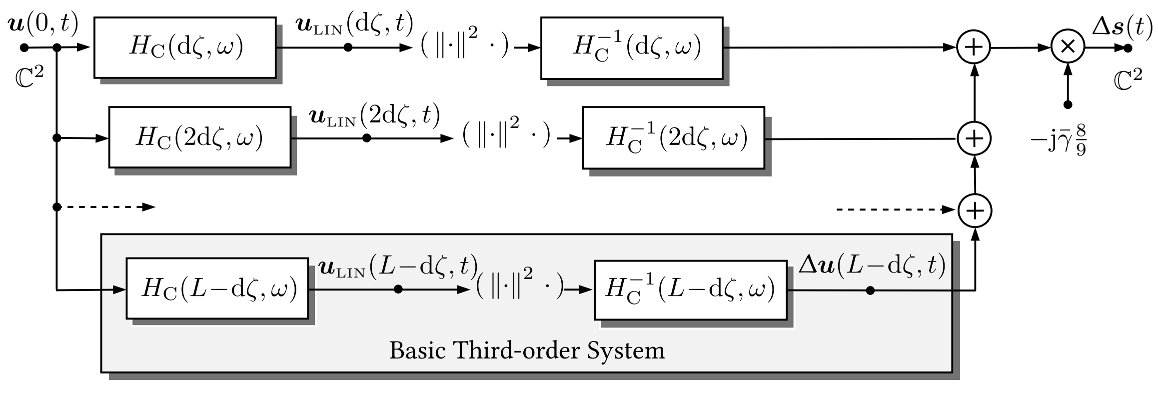


Figure A.2: The *parallel* fiber model consisting of *basic third-order systems* in each of the parallel branches. In the limit, the sum over all branches (associated to a position ζ along the link) converges to the continuous integral.

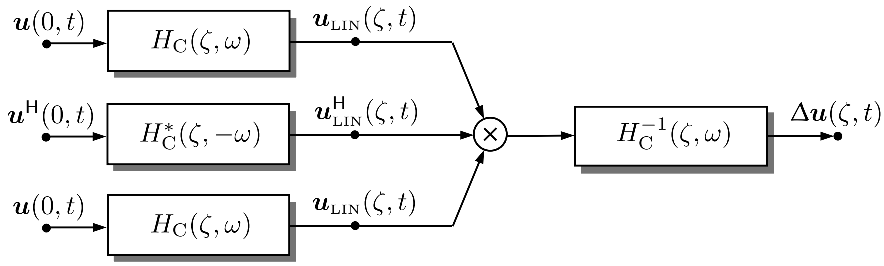


Figure A.3: Block diagram of a *basic third-order system*—a memory-less 3^{rd}-order nonlinearity, sandwiched between two linear, dispersive systems (here, the nonlinear function $(\|\cdot\|^2\,\cdot)$ in between the linear channel response and its inverse).

channel impulse response using the auxiliary variables t_1, t_2, t_3 for each of the inputs, cf. Figure A.3. We find

$$\begin{aligned}\Delta\boldsymbol{u}(\zeta,t) = &\int_{-\infty}^{\infty} \mathrm{d}\sigma\ \frac{1}{2\pi|\mathcal{B}(\zeta)|} h_\mathrm{C}^{-1}(\zeta,\sigma) \\ &\times \int_{-\infty}^{\infty} h_\mathrm{C}(\zeta,\mathrm{t}_1)\boldsymbol{u}(0,t-\sigma-\mathrm{t}_1)\ \mathrm{dt}_1 \\ &\times \int_{-\infty}^{\infty} h^*_\mathrm{C}(\zeta,\mathrm{t}_3)\boldsymbol{u}^\mathsf{H}(0,t-\sigma-\mathrm{t}_3)\ \mathrm{dt}_3 \\ &\times \int_{-\infty}^{\infty} h_\mathrm{C}(\zeta,\mathrm{t}_2)\boldsymbol{u}(0,t-\sigma-\mathrm{t}_2)\ \mathrm{dt}_2\,,\end{aligned} \tag{A.58}$$

where we used $(x * y)^* = x^* * y^*$ and $h^*_\mathrm{C}(z,t) \circ\!\!-\!\!\bullet\ H^*_\mathrm{C}(z,-\omega)$. Then, we use the following substitution of variables

$$\sigma + \mathrm{t}_1 = \tau_1, \quad \mathrm{dt}_1 = \mathrm{d}\tau_1 \tag{A.59}$$

$$\sigma + \mathrm{t}_2 = \tau_2, \quad \mathrm{dt}_2 = \mathrm{d}\tau_2 \tag{A.60}$$

$$\sigma + \mathrm{t}_3 = \tau_3, \quad \mathrm{dt}_3 = \mathrm{d}\tau_3\,, \tag{A.61}$$

and obtain using $h_\mathrm{C}(z,t) = h_\mathrm{C}(z,-t)$ the following intermediate result

$$\begin{aligned}\Delta\boldsymbol{u}(\zeta,t) = &\int_{-\infty}^{\infty} \mathrm{d}\sigma\ \frac{1}{2\pi|\mathcal{B}(\zeta)|} h_\mathrm{C}^{-1}(\zeta,\sigma) \\ &\times \int_{-\infty}^{\infty} h_\mathrm{C}(\zeta,\tau_1-\sigma)\boldsymbol{u}(0,t-\tau_1)\ \mathrm{d}\tau_1 \\ &\times \int_{-\infty}^{\infty} h^*_\mathrm{C}(\zeta,\tau_3-\sigma)\boldsymbol{u}^\mathsf{H}(0,t-\tau_3)\ \mathrm{d}\tau_3 \\ &\times \int_{-\infty}^{\infty} h_\mathrm{C}(\zeta,\tau_2-\sigma)\boldsymbol{u}(0,t-\tau_2)\ \mathrm{d}\tau_2\,.\end{aligned} \tag{A.62}$$

The order of integration is re-arranged to isolate all terms independent of t, and we obtain the following representation of the third-order Volterra system

$$\begin{aligned}\Delta\boldsymbol{u}(\zeta,t) = &\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}\tau_1\mathrm{d}\tau_2\mathrm{d}\tau_3\ \boldsymbol{u}(0,t-\tau_1)\boldsymbol{u}^\mathsf{H}(0,t-\tau_3)\boldsymbol{u}(0,t-\tau_2) \\ &\times \underbrace{\int_{-\infty}^{\infty} h_\mathrm{C}(\zeta,\tau_1-\sigma)h^*_\mathrm{C}(\zeta,\tau_3-\sigma)h_\mathrm{C}(\zeta,\tau_2-\sigma)\frac{1}{2\pi|\mathcal{B}(\zeta)|}h_\mathrm{C}^{-1}(\zeta,\sigma)\,\mathrm{d}\sigma}_{h_3(\tau_1,\tau_2,\tau_3)}\,.\end{aligned} \tag{A.63}$$

In the Volterra theory, the term $h_3(\boldsymbol{\tau}) = h_3(\tau_1, \tau_2, \tau_3)$ is the (time-invariant) *basic third-order Volterra kernel* from (2.65). We recover the direct dependence of the kernel on the four involved LTI systems according to Figure 2.1 and Figure A.3.

We will now show that $h_3(\tau_1, \tau_2, \tau_3)$ is equal to the time-domain kernel $h_{\mathrm{NL}}(\tau_1, \tau_2)$ and we derive the *temporal matching constraint* $\tau_3 = \tau_1 + \tau_2$. We use the definition of the linear channel impulse response $h_{\mathrm{C}}(z, t)$ from (3.64) in the expression of the third-order kernel $h_3(\tau_1, \tau_2, \tau_3)$ to arrive at

$$\begin{aligned}
&\int_{-\infty}^{\infty} \frac{1}{2\pi|\mathcal{B}(\zeta)|} h_{\mathrm{C}}^{-1}(\zeta, \sigma) h_{\mathrm{C}}(\zeta, \tau_1 - \sigma) h_{\mathrm{C}}^{*}(\zeta, \tau_3 - \sigma) h_{\mathrm{C}}(\zeta, \tau_2 - \sigma)\, \mathrm{d}\sigma \qquad \text{(A.64)}\\
=&\int_{-\infty}^{\infty} \frac{1}{(2\pi|\mathcal{B}(\zeta)|)^2} \exp\left(\mathcal{G}(\zeta) + \mathrm{j}\frac{1}{2\mathcal{B}(\zeta)}\left(-\sigma^2 + (\tau_1 - \sigma)^2 - (\tau_3 - \sigma)^2 + (\tau_2 - \sigma)^2\right)\right) \mathrm{d}\sigma\\
=&\int_{-\infty}^{\infty} \frac{1}{(2\pi|\mathcal{B}(\zeta)|)^2} \exp\left(\mathcal{G}(\zeta) + \mathrm{j}\frac{1}{2\mathcal{B}(\zeta)}\left(\tau_1^2 - 2\tau_1\sigma - \tau_3^2 + 2\tau_3\sigma + \tau_2^2 - 2\tau_2\sigma\right)\right) \mathrm{d}\sigma\\
=&\frac{1}{(2\pi|\mathcal{B}(\zeta)|)^2} \exp\left(\mathcal{G}(\zeta) + \mathrm{j}\frac{1}{2\mathcal{B}(\zeta)}\left(\tau_1^2 - \tau_3^2 + \tau_2^2\right)\right)\\
&\qquad \times \int_{-\infty}^{\infty} \exp\left(\mathrm{j}\frac{1}{\mathcal{B}(\zeta)}\left(-\tau_1\sigma + \tau_3\sigma - \tau_2\sigma\right)\right) \mathrm{d}\sigma,
\end{aligned}$$

where we used $\sqrt{\mathrm{j}} = (1/\sqrt{\mathrm{j}})^*$ and $\sqrt{\mathcal{B}(\zeta)} \cdot \left(\sqrt{\mathcal{B}(\zeta)}\right)^* = |\mathcal{B}(\zeta)|$.

Next, we use the identity of the *dirac impulse* together with the *scaling* and *shifting* property of the transform to rewrite the integral over σ as

$$\begin{aligned}
\int_{-\infty}^{\infty} \exp\left(\mathrm{j}\frac{1}{\mathcal{B}(\zeta)}\left(\tau_3 - \tau_1 - \tau_2\right)\sigma\right) \mathrm{d}\sigma &= 2\pi\; \delta\left(\frac{1}{\mathcal{B}(\zeta)}(\tau_3 - \tau_1 - \tau_2)\right) \qquad \text{(A.65)}\\
&= 2\pi|\mathcal{B}(\zeta)|\; \delta(\tau_3 - \tau_1 - \tau_2)\,.
\end{aligned}$$

Then, we use the *sifting property* $\int_{\mathbb{R}} f(y)\delta(x - y)\mathrm{d}y = f(x)$ of the dirac impulse to rewrite $h_3(\tau_1, \tau_2, \tau_3)$ in terms of two variables τ_1 and τ_2. At the same time, the outer integral in (A.63) over τ_3 collapses and we remain with a double integral over τ_1 and τ_2.

We find the 2D Volterra kernel belonging to the third-order basic system as

$$\begin{aligned}
h_3(\tau_1, \tau_2) &= \frac{1}{2\pi|\mathcal{B}(\zeta)|} \exp\left(\mathcal{G}(\zeta) + \mathrm{j}\frac{1}{2\mathcal{B}(\zeta)}\left(\tau_1^2 - \tau_3^2 + \tau_2^2\right)\right) \delta(\tau_3 - \tau_1 - \tau_2) \qquad \text{(A.66)}\\
&= \frac{1}{2\pi|\mathcal{B}(\zeta)|} \exp\left(\mathcal{G}(\zeta) + \mathrm{j}\frac{1}{2\mathcal{B}(\zeta)}\left(\tau_1^2 - (\tau_1 + \tau_2)^2 + \tau_2^2\right)\right)\\
&= \frac{1}{2\pi|\mathcal{B}(\zeta)|} \exp\left(\mathcal{G}(\zeta) - \mathrm{j}\frac{\tau_1\tau_2}{\mathcal{B}(\zeta)}\right),
\end{aligned}$$

which is a *symmetric* kernel in τ_1 and τ_2. Hence, the input/output relation of the basic system can be expressed as

$$\begin{aligned}
\Delta\boldsymbol{u}(\zeta, t) = &\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}\tau_1\mathrm{d}\tau_2\, \boldsymbol{u}(0, t - \tau_1)\boldsymbol{u}^{\mathsf{H}}(0, t - \tau_1 - \tau_2)\boldsymbol{u}(0, t - \tau_2) \qquad \text{(A.67)}\\
&\times \frac{1}{2\pi|\mathcal{B}(\zeta)|} \exp\left(\mathcal{G}(\zeta) - \mathrm{j}\frac{\tau_1\tau_2}{\mathcal{B}(\zeta)}\right).
\end{aligned}$$

In [Sch80], it is shown that the sum of M parallel *basic third-order systems* $h_3^{(m)}(\tau_1, \tau_2, \tau_3)$ with $m \in \{1, 2, \ldots, M\}$ can be expressed via a *general* third-order kernel as

$$h_3(\tau_1, \tau_2, \tau_3) = \sum_{m=1}^{M} h_3^{(m)}(\tau_1, \tau_2, \tau_3)\,. \tag{A.68}$$

In our case, the sum over all (continuous) positions ζ along the link converges to an integral as $\zeta \to 0$ and we recover the (normalized) nonlinear impulse response

$$h_{\mathrm{NL}}(\tau_1, \tau_2) = \frac{1}{L_{\mathrm{eff}}} \int_0^L \frac{1}{2\pi|\mathcal{B}(\zeta)|} \exp\left(\mathcal{G}(\zeta) - \mathrm{j}\frac{\tau_1\tau_2}{\mathcal{B}(\zeta)}\right) \mathrm{d}\zeta\,, \tag{A.69}$$

where we used the effective length L_{eff} as a normalization constant. Including the constant multiplier $-\mathrm{j}\bar{\gamma}\frac{8}{9}$, cf. Figure A.2, we recover the end-to-end relation

$$\Delta\boldsymbol{s}(t) = -\mathrm{j}\bar{\gamma}\frac{8}{9}L_{\mathrm{eff}} \int_{\mathbb{R}^2} \boldsymbol{u}(0, t-\tau_1)\boldsymbol{u}^{\mathsf{H}}(0, t-\tau_1-\tau_2)\boldsymbol{u}(0, t-\tau_2)\, h_{\mathrm{NL}}(\tau_1, \tau_2)\, \mathrm{d}^2\boldsymbol{\tau}\,, \tag{A.70}$$

which is equivalent to (4.63). $\square$

A.4 Alternative Derivation of the Discrete-Time End-to-End RP Method

In this appendix, we provide a time-domain derivation of the *pulse-collision picture* analogous to the derivation in A.3, i.e., using the theory on time-domain Volterra series from Section 2.1.3.2. We first start with a derivation of the self-channel interference (SCI) term, and continue in the second part with the cross-channel interference (XCI) term.

Intra-channel Ansatz In Figure A.4, the modified parallel fiber model is shown to establish the end-to-end relation of the transmit sequence and received perturbation (exemplarily for SCI effects, i.e., the nonlinear perturbation is generated by the *linearly* propagating probe signal alone). The perturbation ansatz for SCI effects is based on the Manakov equation in (3.81) considering only the probe signal $\boldsymbol{u}_\rho(z,t)$ in the source term $\boldsymbol{w}(z,t)$, see (3.75).

The generation of the *local* nonlinear interference is nested between the now z-dependent *dispersed* and *attenuated* transmit pulse $G_{\mathrm{T},\nu}(z,\omega)$ and receive-pulse $G_{\mathrm{R}}(z,\omega)$ from (5.7), (5.9), see Figure A.4. This can be understood as the parallel fiber model used in the previous section, cf. Figure A.2, however now, the transmitter and receiver frond-end are included into each of the parallel branches using the *linear* property of the modulation and demodulation operation. Also, for ease of notation, sampling at interval T is moved to the parallel branches and we denote, by abuse of notation, the sampled discrete-time perturbation at ζ as $\Delta\boldsymbol{a}^{\mathrm{SCI}}(\zeta,kT)$.

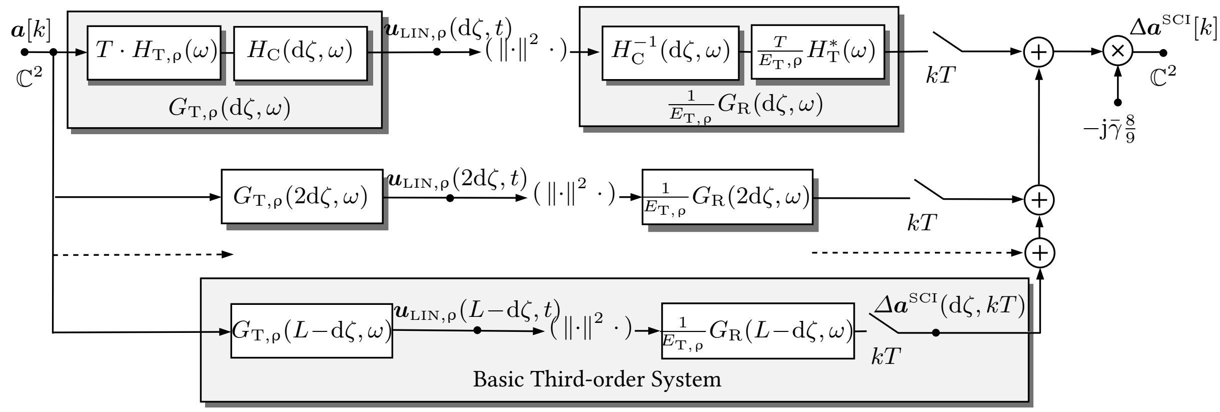


Figure A.4: The *discrete-time* end-to-end parallel fiber model consisting of *basic third-order systems* in each of the parallel branches. Here, the local generation of nonlinear interference is nested between the transmit pulse $G_{\mathrm{T},\nu}(\zeta,\omega)$ with $\nu=\rho$ and receive-pulse $G_{\mathrm{R}}(\zeta,\omega)$ associated with position ζ.

Note, that we can write the receive filter $G_{\mathrm{R}}(z,\omega)$ expressed via the complex conjugate of the transmit pulse $G_{\mathrm{T},\rho}(z,\omega)$ where we used that $\frac{1}{2\pi|\mathcal{B}(z)|}h_{\mathrm{C}}^{-1}(z,t)=\exp(-\mathcal{G}(z))\ h_{\mathrm{C}}^{*}(z,t)$. The pre-factor $1/E_{\mathrm{T},\rho}$ in the receiver frond-end is a scaling operation to set the (noiseless part of the) receive sequence to unity variance of the transmit sequence σ_a^2, i.e., to undo the scaling of the signal $\boldsymbol{u}(0,t)$ to a designated transmit power P_ρ.

The basic third-order system for SCI effects is shown in Figure A.5. We start with the output of the basic third-order system $\Delta\boldsymbol{a}^{\mathrm{SCI}}(\zeta,kT)$, i.e., the sampled perturbation generated in the optical domain by the *local* nonlinear interaction $\Big(\|\boldsymbol{u}_{\mathrm{LIN},\rho}(\zeta,t)\|^2\,\boldsymbol{u}_{\mathrm{LIN},\rho}(\zeta,t)\Big)$.

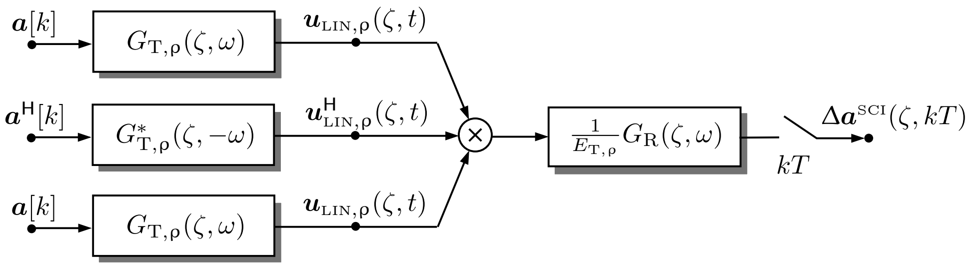


Figure A.5: Block diagram of the *basic third-order system* corresponding to a single spatial branch at position ζ of the *intra*-channel end-to-end system.

We can express $\Delta\boldsymbol{a}^{\mathrm{SCI}}(\zeta,kT)$ as the output of the sampler and the *linear* system $G_{\mathrm{R}}(\zeta,\omega)$. As in the previous derivations, we use the auxiliary variable σ to express the convolution. We yield

$$\begin{aligned}\Delta\boldsymbol{a}^{\mathrm{SCI}}(\zeta,t)\Big|_{t=kT} &= \frac{1}{E_{\mathrm{T},\rho}}\mathrm{e}^{-\mathcal{G}(\zeta)}\ \mathcal{F}^{-1}\{\,G^{*}_{\mathrm{T},\rho}(\zeta,\omega)\,\} * \left(\|\boldsymbol{u}_{\mathrm{LIN},\rho}(\zeta,t)\|^{2}\,\boldsymbol{u}_{\mathrm{LIN},\rho}(\zeta,t)\right)\Big|_{t=kT} \\ &= \frac{1}{E_{\mathrm{T},\rho}}\mathrm{e}^{-\mathcal{G}(\zeta)}\int_{-\infty}^{\infty} g^{*}_{\mathrm{T},\rho}(\zeta,\sigma) \\ &\qquad \times\boldsymbol{u}_{\mathrm{LIN},\rho}(\zeta,t-\sigma)\boldsymbol{u}^{\mathsf{H}}_{\mathrm{LIN},\rho}(\zeta,t-\sigma)\boldsymbol{u}_{\mathrm{LIN},\rho}(\zeta,t-\sigma)\,\mathrm{d}\sigma\Big|_{t=kT}.\end{aligned} \tag{A.71}$$

In the next step, the local signals at ζ are expressed as the input sequence $\boldsymbol{a}[k]$ convolved with the channel impulse response using the auxiliary variables $\upkappa_1$, $\upkappa_2$, $\upkappa_3$ for each of the inputs, cf. Figure A.3. We find

$$\begin{aligned}\Delta\boldsymbol{a}^{\mathrm{SCI}}(\zeta,kT) =&\frac{1}{E_{\mathrm{T},\rho}}\mathrm{e}^{-\mathcal{G}(\zeta)}\int_{-\infty}^{\infty}\mathrm{d}\sigma\ \ g^{*}_{\mathrm{T},\rho}(\zeta,\sigma) \\ &\times\sum_{\upkappa_1\in\mathbb{Z}}\boldsymbol{a}[\upkappa_1]\ g_{\mathrm{T},\rho}(\zeta,kT-\sigma-\upkappa_1 T) \\ &\times\sum_{\upkappa_2\in\mathbb{Z}}\boldsymbol{a}^{\mathsf{H}}[\upkappa_2]\ g^{*}_{\mathrm{T},\rho}(\zeta,kT-\sigma-\upkappa_2 T) \\ &\times\sum_{\upkappa_3\in\mathbb{Z}}\boldsymbol{a}[\upkappa_3]\ g_{\mathrm{T},\rho}(\zeta,kT-\sigma-\upkappa_3 T)\,.\end{aligned} \tag{A.72}$$

Then, we use the substitution of variables

$$\upkappa_1 = k-\kappa_1 \tag{A.73}$$

$$\upkappa_2 = k-\kappa_2 \tag{A.74}$$

$$\upkappa_3 = k-\kappa_3\,, \tag{A.75}$$

and obtain the following result

$$\Delta \boldsymbol{a}^{\mathrm{SCI}}(\zeta, kT) = \frac{1}{E_{\mathrm{T},\rho}} \mathrm{e}^{-\mathcal{G}(\zeta)} \int_{-\infty}^{\infty} \mathrm{d}\sigma \;\; g^*_{\mathrm{T},\rho}(\zeta, \sigma) \\ \times \sum_{\kappa_1 \in \mathbb{Z}} \boldsymbol{a}[k-\kappa_1] \; g_{\mathrm{T},\rho}(\zeta, \kappa_1 T - \sigma) \\ \times \sum_{\kappa_2 \in \mathbb{Z}} \boldsymbol{a}^{\mathsf{H}}[k-\kappa_2] \; g^*_{\mathrm{T},\rho}(\zeta, \kappa_2 T - \sigma) \\ \times \sum_{\kappa_3 \in \mathbb{Z}} \boldsymbol{a}[k-\kappa_3] \; g_{\mathrm{T},\rho}(\zeta, \kappa_3 T - \sigma). \tag{A.76}$$

The order of integration is re-arranged to isolate all terms independent of k, and we obtain the following representation of the third-order Volterra system

$$\Delta \boldsymbol{a}^{\mathrm{SCI}}(\zeta, kT) = \sum_{\kappa_1 \in \mathbb{Z}} \sum_{\kappa_2 \in \mathbb{Z}} \sum_{\kappa_3 \in \mathbb{Z}} \boldsymbol{a}[k-\kappa_1] \boldsymbol{a}^{\mathsf{H}}[k-\kappa_2] \boldsymbol{a}[k-\kappa_3] \\ \times \underbrace{\frac{1}{E_{\mathrm{T},\rho}} \mathrm{e}^{-\mathcal{G}(\zeta)} \int_{-\infty}^{\infty} g_{\mathrm{T},\rho}(\zeta, \kappa_1 T - \sigma) g^*_{\mathrm{T},\rho}(\zeta, \kappa_2 T - \sigma) g_{\mathrm{T},\rho}(\zeta, \kappa_3 T - \sigma) g^*_{\mathrm{T},\rho}(\zeta, \sigma) \, \mathrm{d}\sigma}_{h_{\rho,3}[\kappa_1, \kappa_2, \kappa_3]} . \tag{A.77}$$

The term $h_{\rho,3}[\kappa_1, \kappa_2, \kappa_3]$ is the *discrete-time* Volterra kernel of the (intra-channel) *basic third-order system* [Sch80].

The sum over all (continuous) positions ζ along the link converges to an integral as $\zeta \to 0$ and we recover the normalized, intra-channel nonlinear impulse response

$$h_\rho[\boldsymbol{\kappa}] = \frac{1}{P_\rho L_{\mathrm{eff}}} \int_0^L \frac{\mathrm{e}^{-\mathcal{G}(\zeta)}}{E_{\mathrm{T},\rho}} \int_{-\infty}^{\infty} g_{\mathrm{T},\rho}(\zeta, \kappa_1 T - \sigma) g^*_{\mathrm{T},\rho}(\zeta, \kappa_2 T - \sigma) \\ \times g_{\mathrm{T}}(\zeta, \kappa_3 T - \sigma) g^*_{\mathrm{T},\rho}(\zeta, \sigma) \; \mathrm{d}\sigma \mathrm{d}\zeta \,, \tag{A.78}$$

where we use the pre-factor $P_\rho L_{\mathrm{eff}}$ as a normalization constant. Including the constant multiplier $-\mathrm{j}\bar{\gamma}\frac{8}{9}$, cf. Figure A.4, and using $L_{\mathrm{NL},\rho} = 1/(\bar{\gamma} P_\rho)$, we recover the end-to-end relation

$$\Delta \boldsymbol{a}^{\mathrm{SCI}}[k] = -\mathrm{j}\frac{8}{9} \frac{L_{\mathrm{eff}}}{L_{\mathrm{NL},\rho}} \sum_{\boldsymbol{\kappa} \in \mathbb{Z}^3} \boldsymbol{a}[k-\kappa_1] \boldsymbol{a}^{\mathsf{H}}[k-\kappa_2] \boldsymbol{a}[k-\kappa_3] h_\rho[\boldsymbol{\kappa}], \tag{A.79}$$

which is equivalent to (5.15). □

Inter-channel Ansatz The corresponding proof for inter-channel effects, i.e., cross-channel interference (XCI), is similar as before. The relevant source term for the inter-channel perturbation ansatz is given by considering all XCI terms in (3.81). For this proof, we will only consider the first XPM term in (3.81) and the extension of the proof to the second XPM and XPolM term is straightforward.

The basic third-order system for SCI effects is shown in Figure A.6. We assume a path-constant chromatic dispersion coefficient along the link, i.e., $\beta_2(z) = \bar{\beta}_2$. We use the retarded time of the ν^{th} interferer defined in (3.59) as

$$t_\nu = t - \tau_{\mathrm{wo}}(z, \Delta\omega_\nu) = t - \Delta\omega_\nu \bar{\beta}_2 z, \tag{A.80}$$

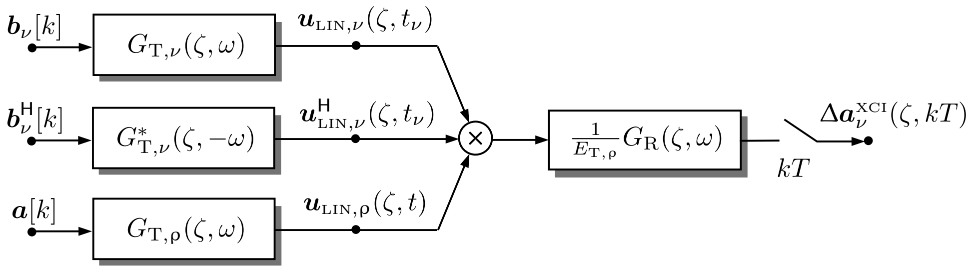


Figure A.6: Block diagram of the *basic third-order system* corresponding to a single spatial branch at position ζ of the *inter*-channel end-to-end system.

given for the constant dispersion coefficient $\bar{\beta}_2$.

We can now rewrite (A.72) to express the XCI perturbation of the basic third-order system as

$$\begin{aligned}\Delta\boldsymbol{a}^{\mathrm{XCI}}_\nu(\zeta,kT) =& \frac{1}{E_{\mathrm{T},\rho}}\mathrm{e}^{-\mathcal{G}(\zeta)}\int_{-\infty}^{\infty} g^*_{\mathrm{T},\rho}(\zeta,\sigma)\left\|\boldsymbol{u}_{\mathrm{LIN},\nu}(\zeta,t_\nu-\sigma)\right\|^2\boldsymbol{u}_{\mathrm{LIN},\rho}(\zeta,t-\sigma)\mathrm{d}\sigma\,\Big|_{t=kT}\\ =& \frac{1}{E_{\mathrm{T},\rho}}\mathrm{e}^{-\mathcal{G}(\zeta)}\int_{-\infty}^{\infty}\mathrm{d}\sigma\;\; g^*_{\mathrm{T},\rho}(\zeta,\sigma)\\ &\times\sum_{\kappa_1\in\mathbb{Z}}\boldsymbol{b}^{\mathsf{H}}_\nu[\kappa_2]\;g_{\mathrm{T},\nu}(\zeta,kT-\sigma-\kappa_2 T)\\ &\times\sum_{\kappa_2\in\mathbb{Z}}\boldsymbol{b}_\nu[\kappa_1]\;g^*_{\mathrm{T},\nu}(\zeta,kT-\sigma-\kappa_1 T)\\ &\times\sum_{\kappa_3\in\mathbb{Z}}\boldsymbol{a}[\kappa_3]\;g_{\mathrm{T},\rho}(\zeta,kT-\sigma-\kappa_3 T)\,,\end{aligned}\tag{A.81}$$

where the dispersed transmit pulse $g_{\mathrm{T},\nu}(\zeta,t)$ already includes the time-retardation given by the walk-off $\tau_{\mathrm{wo}}(z,\Delta\omega_\nu)$, see definition in (5.8).

Following the same steps as for SCI effects, we find after integrating the kernel over all local positions ζ as

$$\begin{aligned}h_\nu[\boldsymbol{\kappa}] =& \frac{1}{P_\nu L_{\mathrm{eff}}}\int_0^L\frac{\mathrm{e}^{-\mathcal{G}(\zeta)}}{E_{\mathrm{T},\rho}}\int_{-\infty}^{\infty}g_{\mathrm{T},\nu}(\zeta,\kappa_1T-\sigma)g^*_{\mathrm{T},\nu}(\zeta,\kappa_2T-\sigma)\\ &\times g_{\mathrm{T},\rho}(\zeta,\kappa_3T-\sigma)g^*_{\mathrm{T},\rho}(\zeta,\sigma)\;\mathrm{d}\sigma\mathrm{d}\zeta\,,\end{aligned}\tag{A.82}$$

where we made again use of the normalization pre-factor $P_\nu L_{\mathrm{eff}}$. The result recovers (5.21) using $g_{\mathrm{R}}(\zeta,t)=\exp(-\mathcal{G}(\zeta))g^*_{\mathrm{T},\rho}(\zeta,t)$ similar as in the original source [DFMS16, (4)].

Taking into account the second XPM and XPolM term in (3.81), we have the

$$\Delta\boldsymbol{a}^{\mathrm{XCI}}_\nu[k] = -\mathrm{j}\frac{8}{9}\frac{L_{\mathrm{eff}}}{L_{\mathrm{NL},\nu}}\sum_{\boldsymbol{\kappa}\in\mathbb{Z}^3}\Big(\boldsymbol{b}_\nu[k-\kappa_1]\boldsymbol{b}^{\mathsf{H}}_\nu[k-\kappa_2]+\boldsymbol{b}^{\mathsf{H}}_\nu[k-\kappa_2]\boldsymbol{b}_\nu[k-\kappa_1]\mathbf{I}\Big)\boldsymbol{a}[k-\kappa_3]h_\nu[\boldsymbol{\kappa}]\,.\tag{A.83}$$

which is equivalent to a single contribution ν in (5.19). □

B. Notation

B.1 Abbreviations

Acronym	Meaning
1D	one-dimensional
2D	two-dimensional
3D	three-dimensional
4D	four-dimensional
A/D	analog-to-digital
ACF	autocorrelation function
ADC	analog-to-digital converter
ASE	amplified spontaneous emission
ASK	amplitude-shift keying
AWGN	additive white Gaussian noise
BER	bit error ratio
BIBO	bounded-input bounded-output
BICM	bit-interleaved coded modulation
BRGC	binary-reflected Gray code
BRGL	binary-reflected Gray labeling
CDF	cumulative distribution function
CSI	channel state information
CD	chromatic dispersion
CM	coded modulation

Acronym	Meaning
CW	continuous wave
D/A	digital-to-analog
DAC	digital-to-analog converter
DEC	decoder
DFT	discrete Fourier transform
DGD	differential group delay
DP	dual-polarization
DSP	digital signal processing
DTFT	discrete-time Fourier transform
DU	dispersion-uncompensated
E/O	electrical-to-optical
ECB	equivalent complex baseband
EDFA	erbium-doped fiber amplifier
EGN	extended Gaussian-noise
ENC	encoder
eRP	enhanced regular perturbation
FEC	forward error correction
FFT	fast Fourier transform
FIR	finite impulse response
FWM	four-wave mixing
GN	Gaussian-noise
GVD	group velocity dispersion
IDRA	ideal distributed Raman amplification
IIR	infinite impulse response
LTI	linear time-invariant
PAF	phase array factor
PDL	polarization dependent loss
IQ	inphase-quadrature
ISI	inter-symbol interference
LDPC	low-density parity-check
LLR	log likelihood ratio
LMS	least mean square

Acronym	Meaning
LO	local-oscillator
LP	logarithmic perturbation
MCI	multi-channel interference
MIMO	multiple-input/multiple-output
MSE	mean-squared error
NLIN	nonlinear interference noise
NLC	nonlinearity compensation
NLI	nonlinear interference
NLPN	nonlinear phase noise
NLSE	nonlinear Schrödinger equation
NTI	nonlinear time-invariant
O/E	optical-to-electrical
OA	optical amplification
ODE	ordinary differential equation
OSNR	optical signal-to-noise ratio
PAM	pulse-amplitude modulation
PDE	partial differential equation
PDF	probability density function
PDM	polarization-division multiplex
PMD	polarization-mode dispersion
PMF	probability mass function
PN	phase noise
PRBS	pseudo random bit sequence
PSD	power spectral density
PWDD	power-weighted dispersion distribution
QAM	quadrature amplitude modulation
QPSK	qarternary phase-shift keying
RBW	resolution bandwidth
RC	raised cosine
RP	regular perturbation
RLP	regular-logarithmic perturbation
RRC	root-raised cosine

Acronym	Meaning
RV	random variable
SCI	self-channel interference
SER	symbol error ratio
SNR	signal-to-noise ratio
SOP	state of polarization
SPM	self-phase modulation
SP	set partitioning
SSFM	split-step Fourier method
SSMF	standard single-mode fiber
VSTF	Volterra series transfer function
WDM	wavelength-division multiplexing
XCI	cross-channel interference
XPM	cross-phase modulation
XPolM	cross-polarization modulation
iXPM	intra-channel cross-phase modulation

B.2 Mathematical Operators

Operator	Meaning
$\mathrm{E}\{\cdot\}$	expectation
$(\cdot)^*$	complex conjugation
$(\cdot)^{\mathsf{T}}$	transposition of a matrix/vector
$(\cdot)^{\mathsf{H}}$	conjugate transposition of matrix/vector
$(\cdot)^{-1}$	inverse
$\lvert \cdot \rvert$	absolute value of a scalar
$\lVert \boldsymbol{x} \rVert_p$	p-norm of a vector $\boldsymbol{x}$ (for $p = 2$, subscript is omitted)
$\langle \cdot \rangle$	sequence
$\stackrel{\text{def}}{=}$	equal per definition
$x(t) * y(t)$	linear convolution
$\arg\{\,\cdot\,\} \in (-\pi, \pi]$	argument of a complex variable
$X(\omega) = \mathcal{F}\{\, x(t)\,\}$	Fourier transform
$x(t) = \mathcal{F}^{-1}\{\, X(\omega)\,\}$	inverse Fourier transform
$\partial_x^n f(x, y)$	n^{th} partial derivative of $f(x, y)$ w.r.t. x
$\mathrm{d}^n\boldsymbol{x} \stackrel{\text{def}}{=} \mathrm{d}x_1 \mathrm{d}x_2 \ldots \mathrm{d}x_n$	multi-dimensional differential
$\mathrm{Re}\{\cdot\}$	real part of a complex variable
$\mathrm{Im}\{\cdot\}$	imaginary part of a complex variable

B.3 Mathematical Symbols

Variables, vectors, matrices and sets.

Variable	Unit	Meaning
$a[k] \in \mathbb{C}$		discrete-time data symbol
$\boldsymbol{a}[k] \in \mathbb{C}^2$		column vector of discrete-time data symbols
$\Delta\boldsymbol{a}[k]$		column vector of discrete-time (additive) perturbation
$A(\mathrm{e}^{\mathrm{j}\omega T})$		frequency-domain data symbol
$\mathcal{A}$		set of data symbols, i.e., signal constellation
$\alpha(z)$	m^{-1}	attenuation coefficient
$\bar{\alpha}$	m^{-1}	path-average attenuation coefficient
$\mathcal{M}$		set of bijective mappings
BER		bit error ratio
B_ν	Hz	spectral bandwidth of the ν^{th} wavelength signal
B_{WDM}	Hz	spectral bandwidth of the full WDM signal
B_{SIM}	Hz	simulation bandwidth
$\beta(z,\omega)$	m^{-1}	space- and frequency-dependent propagation constant
$\beta_n(z)$	$\mathrm{s}^n\mathrm{m}^{-1}$	n^{th} partial derivative of $\partial_\omega^n\beta(z,\omega)$ in the vicinity of ω_0
$\bar{\beta}_2$	$\mathrm{s}^2\mathrm{m}^{-1}$	path-average dispersion coefficient
$\mathcal{B}(z)$	s^2	dispersion profile, $\mathcal{B}(z) = \int_0^z \beta_2(\zeta)\mathrm{d}\zeta$
c	m/s	speed of light, $\mathrm{c} = 299792458\ \mathrm{m/s}$
C		Shannon capacity
$\mathcal{C}$		code (set of codewords)
D		dimensionality of the symbol alphabet, i.e., $\mathcal{A} \subset \mathbb{R}^D$
$\delta(t)$	s^{-1}	Dirac function (in continuous time)
e		Euler number
E_{b}	s	energy per information bit
E_{s}	s	energy per (4D) channel symbol
E_{T}	s	energy of the basic pulse $g_{\mathrm{T}}(t)$
$f \in \mathbb{R}$	Hz	continuous frequency
$\mathsf{f}_{\mathsf{x}}(x)$		probability density function of a random variable x
$g(z)$	m^{-1}	gain coefficient
$g_{\mathrm{T}}(t)$		transmit filter impulse shape

Variable	Unit	Meaning
$\mathcal{G}(z)$		normalized logarithmic power profile
$\gamma(z)$	m^{-1}	fiber nonlinearity coefficient (normalized by 1-Watt)
h	Js	Planck's constant, $\mathrm{h} = 6.62606896 \times 10^{-34}$ Js
$h(t)$	s^{-1}	continuous-time impulse repsonse
$H(\omega)$		transfer function
$h_{\mathrm{T}}(t)$	s^{-1}	transmit filter impulse response
$h_{\mathrm{R}}(t)$	s^{-1}	receive filter impulse response
$h_{\mathrm{C}}(z,t)$	s^{-1}	linear channel impulse response
$h_{\mathrm{CD}}(z,t)$	s^{-1}	chromatic dispersion impulse response
$h_{\mathrm{NL}}(\tau_1,\tau_2)$	s^{-2}	(continuous-time) nonlinear impulse repsonse
$H_{\mathrm{NL}}(\upsilon_1,\upsilon_2)$		nonlinear transfer function
$H_\nu(\mathrm{e}^{\mathrm{j}\boldsymbol{\omega}T})$		frequency-periodic nonlinear transfer function
j		imaginary unit
$k \in \mathbb{Z}$		index of discrete-time data symbols
$\boldsymbol{k} \in \mathbb{Z}^n$		n-dimensional column vector of discrete-time indices
$\kappa \in \mathbb{Z}$		(difference) index of discrete-time data symbols
$\boldsymbol{\kappa} \in \mathbb{Z}^n$		n-dimensional column vector of difference indices
$\mathcal{K}$		set of difference indices $\boldsymbol{\kappa}$
L	m	total transmission length of the link
L_{sp}	m	span length
L_{eff}	m	effective length of the transmission link
$L_{\mathrm{eff,a}}$	m	asymptotic effective length
L_{NL}	m	nonlinear length
L_{D}	m	dispersion length
$L_{\mathrm{wo},\nu}$	m	walk-off length, w.r.t. the ν^{th} co-propagating channel
$\lambda = \mathrm{c}/f$	m	wavelength
λ_0	m	center wavelength
m		number of signal points, cardinality of $\mathcal{A}$
M		length of DFT processing blocks
μ		discrete frequency index $\mu \in \{0,\dots,M-1\}$
$\boldsymbol{\mu}$		column vector of discrete frequency indices
$\mathcal{U}$		set of discrete frequency indices $\boldsymbol{\mu}$

Variable	Unit	Meaning
$n(t)$		noise sample
N_0		(one-sided) noise power spectral density (ECB signal)
N_{sp}		number of spans
N_{ch}		number of wavelength channels
ν		wavelength channel index
$\mathsf{\rho}$		channel index of the *probe channel*, s.t. $\omega_{\mathsf{\rho}} = \omega_0$
$\omega = 2\pi f$	Hz	angular frequency
$\boldsymbol{\omega} \in \mathbb{R}^n$	Hz^n	n-dimensional angular frequency vector
ω_0	Hz	center frequency of the signaling regime of interest
$\omega_1, \omega_2, \omega_3$	Hz	auxiliary frequency variable with $\omega_3 \overset{\mathrm{def}}{=} \omega - \omega_1 + \omega_2$
$\Delta\omega_\nu$	Hz	frequency offset of the ν^{th} interferer w.r.t. ω_0
ω_{Nyq}	Hz	Nyquist frequency
P_ν		signal power (per wavelength channel ν)
$\mathcal{P}(z)$		normalized power profile, i.e., $\mathcal{P}(0) = \mathcal{P}(L) = 1$
$\phi_{\mathrm{NL},\nu}$	rad	nonlinear phase shift (per wavelength channel ν)
$\mathsf{q}[\kappa]$		binary source symbols (uncoded)
$\mathbf{q}[k]$		row vector of binary source symbols
$\hat{\mathsf{q}}[\kappa]$		estimate on binary source symbol
$\hat{\mathbf{q}}[k]$		row vector of estimates $\hat{q}$
$\mathcal{Q}^2$		$\mathcal{Q}^2$-factor, i.e., $\mathcal{Q}^2 \overset{\mathrm{def}}{=} 20\log_{10}(\sqrt{2}\mathrm{erfc}^{-1}(2\,\mathsf{BER}))$
$r(t) \in \mathbb{C}$		electrical receive signal
$\boldsymbol{r}(t) \in \mathbb{C}^2$		column vector of electrical receive signal
$R(\omega)$	s	frequency-domain electrical receive signal
R_{s}	s^{-1}	symbol rate
R_{m}		number of bits per data symbol, i.e., *rate* of modulation
$R_{\mathcal{C}}$		code rate
ρ		roll-off factor of the transmit pulse shape
$s(t) \in \mathbb{C}$		electrical transmit signal
$\boldsymbol{s}(t) \in \mathbb{C}^2$		column vector of electrical transmit signal
$S(\omega)$	s	frequency-domain electrical transmit signal
$\mathcal{S}_{\mathsf{\rho}}$	s^2	map strength w.r.t the probe channel
$\mathcal{S}_{T,\mathsf{\rho}}$		dimensionless map strength normalized by $(2\pi R_{\mathrm{s}})^2$

Variable	Unit	Meaning
$\mathcal{S}_{T,\nu}$		dimensionless map strength w.r.t ν^{th} wavelength channel
σ_x		variance of random variable x
σ_a^2		variance of the (4D) discrete data symbols
σ_n^2		variance of the (4D) noise process
$\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, \boldsymbol{\sigma}_3$		Pauli matrices
$\vec{\boldsymbol{\sigma}}$		Pauli vector
$t \in \mathbb{R}$	s	continuous time (relative to the probe's time base)
$\boldsymbol{t} \in \mathbb{R}^n$	s^n	n-dimensional continuous time vector
$\mathrm{t}_1, \mathrm{t}_2, \mathrm{t}_3$	s	auxiliary time variables with $\mathrm{t}_3 \stackrel{\text{def}}{=} t - \mathrm{t}_1 + \mathrm{t}_2$
τ_1, τ_2	s	time difference variable with $\tau_1 \stackrel{\text{def}}{=} t - \mathrm{t}_1$, $\tau_2 \stackrel{\text{def}}{=} \mathrm{t}_1 - \mathrm{t}_2$
T		time duration for channel symbols
T_{b}		time duration for information bit/binary source symbol
$u(z,t) \in \mathbb{C}$		optical receive signal
$U(z,\omega)$	s	frequency-domain optical signal
$\boldsymbol{u}(z,t)$		column vector of optical signal u_{x} and u_{y} in Jones space
$\boldsymbol{u}_{\mathrm{LIN}}(z,t)$		linearly propagating optical signal
$\vec{\boldsymbol{u}}(z,t)$		column vector of optical signal in Stokes space
$\mathbf{v}_{\mathrm{g}}(\omega_0)$	m/s	group velocity at ω_0
υ_1, υ_2	Hz	frequency difference with $\upsilon_1 \stackrel{\text{def}}{=} \omega_1 - \omega$, $\upsilon_2 \stackrel{\text{def}}{=} \omega_2 - \omega_1$
ξ	Hz^2	frequency difference product $\xi \stackrel{\text{def}}{=} \upsilon_1 \upsilon_2$
$y[k]$		discrete-time receive symbol
$\boldsymbol{y}[k]$		column vector of receive symbols
$z \in \mathbb{R}$	m	spatial position in direction of propagation

Bibliography

[AB98] M. J. Ablowitz, G. Biondini. Multiscale Pulse Dynamics in Communication Systems with Strong Dispersion Management. *Opt. Lett.*, vol. 23, no. 21, pp. 1668–1670, 1998.

[AFAK+18] A. Ali, F. Ferreira, M. Al-Khateeb, D. Charlton, C. Laperle, A. D. Ellis. The Impact of Dispersion Slope on Fiber Nonlinearity in Ultra-Wideband Optical Communication System. In *44th Eur. Conf. Opt. Commun.*, paper Tu4G.5, 2018.

[Agr02] G. P. Agrawal. *Fiber-Optic Communication Systems.* John Wiley & Sons, Inc., 3rd edition, 2002.

[Agr06] G. P. Agrawal. *Nonlinear Fiber Optics.* Optics and Photonics. Academic Press, 4th edition, 2006.

[Agr10] G. P. Agrawal. *Fiber-Optic Communication Systems.* John Wiley & Sons, Inc., 4th edition, 2010.

[Agr17] E. Agrell. Capacity Bounds in Optical Communications. In *43rd Eur. Conf. Opt. Commun.*, paper M.2.D.1, 2017.

[AH00] M. J. Ablowitz, T. Hirooka. Resonant Nonlinear Intrachannel Interactions in Strongly Dispersion-managed Transmission Systems. *Opt. Lett.*, vol. 25, no. 24, pp. 1750–1752, 2000.

[AH02a] M. J. Ablowitz, T. Hirooka. Managing Nonlinearity in Strongly Dispersion-managed Optical Pulse Transmission. *J. Opt. Soc. Am. B Opt. Phys.*, vol. 19, no. 3, pp. 425–439, 2002.

[AH02b] M. J. Ablowitz, T. Hirooka. Resonant Intrachannel Pulse Interactions in Dispersion-Managed Transmission Systems. *IEEE J. Sel. Top. Quantum Electron.*, vol. 8, no. 3, pp. 603–615, 2002.

[AHB01] M. J. Ablowitz, T. Hirooka, G. Biondini. Quasi-Linear Optical Pulses in Strongly Dispersion-Managed Transmission Systems. *Opt. Lett.*, vol. 26, no. 7, pp. 459–461, 2001.

[AW05] G. B. Arfken, H. J. Weber. *Mathematical Methods for Physicists.* Elsevier Academic Press, 6th edition, 2005.

[BAL+20] A. Bononi, J.-C. Antona, M. Lonardi, A. Carbo-Méseguer, P. Serena. The Generalized Droop Formula for Low Signal to Noise Ratio Optical Links. *J. Lightwave Technol.*, vol. 38, no. 8, pp. 2201–2213, 2020.

[BC85] S. Boyd, L. Chua. Fading Memory and the Problem of Approximating Nonlinear Operators with Volterra Series. *IEEE Trans. Circ. Syst.*, vol. 32, pp. 1150–1161, 1985.

[BCRC16] A. Bakhshali, W. Chan, A. Rezania, J. C. Cartledge. Sequential MAP Detection for High Baud-Rate Systems with Pattern-Dependent Distortions. In *42th Eur. Conf. Opt. Commun.*, 2016.

[BDS+20] A. Bononi, R. Dar, M. Secondini, P. Serena, P. Poggiolini. Fiber Nonlinearity and Optical System Performance. In *Springer Handbook of Optical Networks*, pp. 287–351. Springer International Publishing, 2020.

[Boy85] S. P. Boyd. *Volterra Series: Engineering Fundamentals.* PhD thesis, University of California, Berkeley, 1985.

[BSB08] A. Bononi, P. Serena, M. Bertolini. Unified Analysis of Weakly-Nonlinear Dispersion-Managed Optical Transmission Systems Using a Perturbative Approach. *C. R. Phys.*, vol. 9, no. 9-10, pp. 947–962, 2008.

[BSO08] A. Bononi, P. Serena, A. Orlandini. A Unified Design Framework for Terrestrial Systems. *J. Lightwave Technol.*, vol. 26, no. 22, pp. 3617–3631, 2008.

[CBCJ14] A. Carena, G. Bosco, V. Curri, Y. Jiang. EGN Model of Non-Linear Fiber Propagation. *Opt. Express*, vol. 22, no. 13, pp. 16335–16362, 2014.

[CGK+17] J. C. Cartledge, F. P. Guiomar, F. R. Kschischang, G. Liga, M. P. Yankov. Digital Signal Processing for Fiber Nonlinearities [Invited]. *Opt. Express*, vol. 25, no. 3, pp. 1916–1936, jan 2017.

[CS99] J. H. Conway, N. J. A. Sloane. *Sphere Packings, Lattices and Groups.* Springer, 3rd edition, 1999.

[Dar16] R. Dar. Analytical and Semi-Analytical Models for Nonlinear Transmission. In *42th Eur. Conf. Opt. Commun.*, paper W.1.D.5, 2016.

[DFM+15] R. Dar, M. Feder, A. Mecozzi, M. Shtaif, P. J. Winzer. Nonlinear Phase and Polarization Rotation Noise in Fully Loaded WDM Systems. In *41th Eur. Conf. Opt. Commun.*, 2015.

[DFMS13] R. Dar, M. Feder, A. Mecozzi, M. Shtaif. Properties of Nonlinear Noise in Long, Dispersion-Uncompensated Fiber Links. *Opt. Express*, vol. 21, no. 22, pp. 25685–25699, 2013.

[DFMS14] R. Dar, M. Feder, A. Mecozzi, M. Shtaif. Accumulation of Nonlinear Interference Noise in Fiber-Optic Systems. *Opt. Express*, vol. 22, no. 12, pp. 14199–14211, 2014.

[DFMS15] R. Dar, M. Feder, A. Mecozzi, M. Shtaif. Inter-Channel Nonlinear Interference Noise in WDM Systems: Modeling and Mitigation. *J. Lightwave Technol.*, vol. 33, no. 5, pp. 1044–1053, 2015.

[DFMS16] R. Dar, M. Feder, A. Mecozzi, M. Shtaif. Pulse Collision Picture of Inter-Channel Nonlinear Interference in Fiber-Optic Communications. *J. Lightwave Technol.*, vol. 34, no. 2, pp. 593–607, 2016.

[DW17] R. Dar, P. Winzer. Nonlinear Interference Mitigation: Methods and Potential Gain. *J. Lightwave Technol.*, vol. 35, no. 4, pp. 903–930, 2017.

[Eng15] R. Engelbrecht. *Nichtlineare Faseroptik: Grundlagen und Anwendungsbeispiele.* Springer-Verlag, 2015.

[FABH02] Y. Frignac, J.-C. Antona, S. Bigo, J.-P. Hamaide. Numerical Optimization of Pre-and In-line Dispersion Compensation in Dispersion-managed Systems at 40Gbit/s. In *Opt. Fiber Commun. Conf.*, paper Th.FF.5, 2002.

[Fan54] U. Fano. A Stokes-Parameter Technique for the Treatment of Polarization in Quantum Mechanics. *Phys. Rev.*, vol. 93, no. 1, p. 121, 1954.

[FBP09] J. K. Fischer, C.-A. Bunge, K. Petermann. Equivalent Single-Span Model for Dispersion-Managed Fiber-Optic Transmission Systems. *J. Lightwave Technol.*, vol. 27, no. 16, pp. 3425–3432, 2009.

[FDT$^+$12] Y. Fan, L. Dou, Z. Tao, L. Lei, S. Oda, T. Hoshida, J. C. Rasmussen. Modulation Format Dependent Phase Noise Caused by Intra-channel Nonlinearity. In *38th Eur. Conf. Opt. Commun.*, paper We.2.C.3, 2012.

[FES$^+$18] F. Frey, R. Emmerich, C. Schubert, J. K. Fischer, R. F. H. Fischer. Improved Perturbation-based Fiber Nonlinearity Compensation. In *44th Eur. Conf. Opt. Commun.*, paper Tu1F.6, 2018.

[Fey51] R. P. Feynman. An Operator Calculus Having Applications in Quantum Electrodynamics. *Physical Review*, vol. 84, no. 1, p. 108, 1951.

[FFF20] F. Frey, J. K. Fischer, R. F. H. Fischer. On Discrete-Time/Frequency-Periodic End-to-End Fiber-Optical Channel Models. *arXiv Prepr.*, vol. 1907.06087, no. v3, 2020.

[Fis02] R. F. H. Fischer. *Precoding and Signal Shaping for Digital Transmission.* John Wiley & Sons, 2002.

[Fis09] J. K. Fischer. *Scaling Laws of Optical Fibre Nonlinearity in Dispersion-Managed Transmission Systems.* PhD thesis, Technische Universität Berlin, 2009.

[FME⁺17] F. Frey, L. Molle, R. Emmerich, C. Schubert, J. K. Fischer, R. F. H. Fischer. Single-step Perturbation-based Nonlinearity Compensation of Intra- and Inter-Subcarrier Nonlinear Interference. In *43rd Eur. Conf. Opt. Commun.*, 2017.

[Fra96] W. A. Frank. Sampling Requirements for Volterra System Identification. *IEEE Sig. Proc. Lett.*, vol. 3, pp. 266–268, 1996.

[FS05] E. Forestieri, M. Secondini. Solving the Nonlinear Schrödinger Equation. In *Opt. Commun. Theor. Tech.*, pp. 3–11. Springer, 2005.

[GAMP15] F. P. Guiomar, S. B. Amado, C. S. Martins, A. N. Pinto. Time-Domain Volterra-Based Digital Backpropagation for Coherent Optical Systems. *J. Lightwave Technol.*, vol. 33, no. 15, pp. 3170–3181, 2015.

[GCK⁺14] Y. Gao, J. C. Cartledge, A. S. Karar, S. S.-H. Yam, M. O'Sullivan, C. Laperle, A. Borowiec, K. Roberts. Reducing the Complexity of Perturbation Based Nonlinearity Pre-compensation Using Symmetric EDC And Pulse Shaping. *Opt. Express*, vol. 22, no. 2, pp. 1209–1219, 2014.

[GCKY13] Y. Gao, J. C. Cartledge, A. Karar, S.-H. Yam. Reducing the Complexity of Nonlinearity Pre-compensation Using Symmetric EDC and Pulse Shaping. In *39th Eur. Conf. Opt. Commun.*, paper PD.3.E5, 2013.

[GD91] C. R. Giles, E. Desurvire. Propagation of Signal and Noise in Concatenated Erbium-Doped Fiber Optical Amplifiers. *J. Lightwave Technol.*, vol. 9, no. 2, pp. 147–154, 1991.

[Gha17] A. Ghazisaeidi. A Theory of Nonlinear Interactions Between Signal and Amplified Spontaneous Emission Noise in Coherent Wavelength Division Multiplexed Systems. *J. Lightwave Technol.*, vol. 35, no. 23, pp. 5150–5175, 2017.

[GK00] J. P. Gordon, H. Kogelnik. PMD Fundamentals: Polarization Mode Dispersion in Optical Fibers. *Proc. Natl. Acad. Sci. U.S.A.*, vol. 97, no. 9, pp. 4541–50, 2000.

[GP13] F. P. Guiomar, A. N. Pinto. Simplified Volterra Series Nonlinear Equalizer for Polarization-multiplexed Coherent Optical Systems. *J. Lightwave Technol.*, vol. 31, no. 23, pp. 3879–3891, 2013.

[GRTP11] F. P. Guiomar, J. D. Reis, A. L. Teixeira, A. N. Pinto. Digital Post-Compensation Using Volterra Series Transfer Function. *IEEE Photonics Technol. Lett.*, vol. 23, no. 19, pp. 1412–1414, 2011.

[GT96] I. R. Gabitov, S. K. Turitsyn. Averaged Pulse Dynamics in a Cascaded Transmission System with Passive Dispersion Compensation. *Opt. Lett.*, vol. 21, no. 5, pp. 327–329, 1996.

[Hal15] B. Hall. *Lie Groups, Lie Algebras, And Representations: An Elementary Introduction*, volume 222. Springer, 2015.

[Hub92] J. Huber. *Trelliscodierung*. Springer Verlag, Berlin, 1992.

[IK07] E. Ip, J. M. Kahn. Digital Equalization of Chromatic Dispersion and Polarization Mode Dispersion. *J. Lightwave Technol.*, vol. 25, no. 8, pp. 2033–2043, 2007.

[JA14] P. Johannisson, E. Agrell. Modeling of Nonlinear Signal Distortion in Fiber-Optic Networks. *J. Lightwave Technol.*, vol. 32, no. 23, pp. 3942–3950, 2014.

[JK13] P. Johannisson, M. Karlsson. Perturbation Analysis of Nonlinear Propagation in a Strongly Dispersive Optical Communication System. *J. Lightwave Technol.*, vol. 31, no. 8, pp. 1273–1282, 2013.

[Joh12] P. Johannisson. Analytical Modeling of Nonlinear Propagation in a Strongly Dispersive Optical Communication System. *arXiv Prepr.*, vol. 1205.2193, no. v2, 2012.

[Jon41] R. C. Jones. A New Calculus for the Treatment of Optical Systems—I: Description and Discussion of the Calculus. *J. Opt. Soc. Am.*, vol. 31, no. 7, pp. 488–493, 1941.

[Kam08] K. D. Kammeyer. *Nachrichtenübertragung*. Vieweg + Teubner, 2008.

[Kam13] I. Kaminow. *Optical Fiber Telecommunications Volume VIB, Systems and Networks*. Elsevier Science Publishing, 6th edition, 2013.

[Kar14] M. Karlsson. Four-Dimensional Rotations in Coherent Optical Communications. *J. Lightwave Technol.*, vol. 32, no. 6, pp. 1246–1257, 2014.

[KD14] S. Kumar, M. J. Deen. *Fiber Optic Communications: Fundamentals and Applications*. John Wiley & Sons, 2014.

[KL02] I. Kaminow, T. Li. *Optical Fiber Telecommunications IV B*. Elsevier Science Publishing, 2002.

[Kor02] N. Korneev. Application of Magnus Series for Polarization Evolution in Fibers. *Revista Mexicana De Física*, vol. 48, no. 3, pp. 250–254, 2002.

[KP04] M. Karlsson, M. Petersson. Quaternion Approach to PMD and PDL Phenomena in Optical Fiber Systems. *J. Lightwave Technol.*, vol. 22, no. 4, pp. 1137–1146, 2004.

[LHP+05] H. Louchet, A. Hodzic, K. Petermann, A. Robinson, R. Epworth. Simple Criterion for the Characterization of Nonlinear Impairments in Dispersion-Managed Optical Transmission Systems. *IEEE Photonics Technol. Lett.*, vol. 17, no. 10, pp. 2089–2091, 2005.

[LLH+12] L. Liu, L. Li, Y. Huang, K. Cui, Q. Xiong, F. N. Hauske, C. Xie, Y. Cai. Intrachannel Nonlinearity Compensation by Inverse Volterra Series Transfer Function. *J. Lightwave Technol.*, vol. 30, no. 3, pp. 310–316, 2012.

[Lou06] H. Louchet. *Top-Down Analysis of High-Capacity Fiber-OpticTransmission.* PhD thesis, Technische Universität Berlin, 2006.

[MCS00a] A. Mecozzi, C. B. Clausen, M. Shtaif. Analysis of Intrachannel Nonlinear Effects in Highly Dispersed Optical Pulse Transmission. *IEEE Photonics Technol. Lett.*, vol. 12, no. 4, pp. 392–394, 2000.

[MCS00b] A. Mecozzi, C. B. Clausen, M. Shtaif. System Impact of Intra-channel Nonlinear Effects in Highly Dispersed Optical Pulse Transmission. *IEEE Photonics Technol. Lett.*, vol. 12, no. 12, pp. 1633–1635, 2000.

[MCS+01] A. Mecozzi, C. B. Clausen, M. Shtaif, S. G. Park, A. H. Gnauck. Cancellation of Timing and Amplitude Jitter in Symmetric Links Using Highly Dispersed Pulses. *IEEE Photonics Technol. Lett.*, vol. 13, no. 5, pp. 445–447, 2001.

[ME12] A. Mecozzi, R. J. Essiambre. Nonlinear Shannon Limit in Pseudolinear Coherent Systems. *J. Lightwave Technol.*, vol. 30, no. 12, pp. 2011–2024, 2012.

[Mec11] A. Mecozzi. A Unified Theory of Intrachannel Nonlinearity in Pseudolinear Transmission. In *Impact of Nonlinearities on Fiber Optic Communications*, pp. 253–291. Springer Science & Business Media, 2011.

[Men89] C. R. Menyuk. Pulse Propagation in an Elliptically Birefringent Kerr Medium. *IEEE J. Quantum Electron.*, vol. 25, no. 12, pp. 2674–2682, 1989.

[Men99] C. R. Menyuk. Application of Multiple-Length-Scale Methods to the Study of Optical Fiber Transmission. *J. Eng. Math.*, vol. 875, pp. 113–136, 1999.

[MM06] C. R. Menyuk, B. S. Marks. Interaction of Polarization Mode Dispersion and Nonlinearity in Optical Fiber Transmission Systems. *J. Lightwave Technol.*, vol. 24, no. 7, pp. 2806–2826, 2006.

[MM12] A. Mecozzi, F. Matera. Polarization Scattering by Intra-channel Collisions. *Opt. Express*, vol. 20, no. 2, pp. 1213–1218, 2012.

[ONO+14] T. Oyama, H. Nakashima, S. Oda, T. Yamauchi, Z. Tao, T. Hoshida, J. Rasmussen. Robust and Efficient Receiver-side Compensation Method for Intra-channel Nonlinear Effects. In *Opt. Fiber Commun. Conf.*, paper Tu3A.3, 2014.

[OW83] A. V. Oppenheim, A. S. Willsky. *Signals and Systems.* Prentice Hall, 1983.

[PBC+12] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, F. Forghieri. A Detailed Analytical Derivation of the GN Model of Non-Linear Interference in Coherent Optical Transmission Systems. *arXiv Prepr.*, vol. v13, no. 1209.0394, 2012.

[PBP97] K. V. Peddanarappagari, M. Brandt-Pearce. Volterra Series Transfer Function of Single-mode Fibers. *J. Lightwave Technol.*, vol. 15, no. 12, pp. 2232–2241, 1997.

[PCC+11] P. Poggiolini, A. Carena, V. Curri, G. Bosco, F. Forghieri. Analytical Modeling of Nonlinear Propagation in Uncompensated Optical Transmission Links. *IEEE Photonics Technol. Lett.*, vol. 23, no. 11, pp. 742–744, 2011.

[PJ17] P. Poggiolini, Y. Jiang. Recent Advances in the Modeling of the Impact of Nonlinear Fiber Propagation Effects on Uncompensated Coherent Transmission Systems. *J. Lightwave Technol.*, vol. 35, no. 3, pp. 458–480, 2017.

[Pog12] P. Poggiolini. The GN Model of Non-linear Propagation in Uncompensated Coherent Optical Systems. *J. Lightwave Technol.*, vol. 30, no. 24, pp. 3857–3879, 2012.

[PS08] J. G. Proakis, M. Salehi. *Digital Communications.* McGraw-Hill Higher Education, 5th edition, 2008.

[RE10] A. A. Robert, C. Essex. *Calculus: A Complete Course.* Prentice Hall, 7th edition, 2010.

[SA15] L. Szczecinski, A. Alvarado. *Bit-interleaved Coded Modulation: Fundamentals, Analysis and Design.* John Wiley & Sons, 2015.

[Sav10] S. J. Savory. Digital Coherent Optical Receivers: Algorithms and Subsystems. *IEEE J. Sel. Top. Quantum Electron.*, vol. 16, no. 5, pp. 1164–1179, 2010.

[SB13] P. Serena, A. Bononi. An Alternative Approach to the Gaussian Noise Model and its System Implications. *J. Lightwave Technol.*, vol. 31, no. 22, pp. 3489–3499, 2013.

[SB15] P. Serena, A. Bononi. On the Nonlinear Reference Phase in Regular Perturbation Models. In *2015 Tyrrhenian Int. Work. Digit. Commun.*, 2015.

[SBO07] P. Serena, A. Bononi, A. Orlandini. Fundamental Laws of Parametric Gain in Periodic Dispersion-managed Optical Links. *J. Opt. Soc. Am. B*, vol. 24, no. 4, pp. 773–787, 2007.

[Sch80] M. Schetzen. *The Volterra and Wiener Theories of Nonlinear Systems.* John Wiley and Sons, Inc., 1980.

[Sei09] M. Seimetz. *High-order Modulation for Optical Fiber Transmission.* Springer Verlag, 2009.

[SF12] M. Secondini, E. Forestieri. Analytical Fiber-Optic Channel Model in the Presence of Cross-Phase Modulation. *IEEE Photonics Technol. Lett.*, vol. 24, no. 22, pp. 2016–2019, 2012.

[SFM09] M. Secondini, E. Forestieri, C. R. Menyuk. A Combined Regular-Logarithmic Perturbation Method for Signal-Noise Interaction in Amplified Optical Systems. *J. Lightwave Technol.*, vol. 27, no. 16, pp. 3358–3369, 2009.

[SFP13] M. Secondini, E. Forestieri, G. Prati. Achievable Information Rate in Nonlinear WDM Fiber-Optic Systems With Arbitrary Modulation Formats and Dispersion Maps. *J. Lightwave Technol.*, vol. 31, no. 23, pp. 3839–3852, 2013.

[Sha48] C. E. Shannon. A Mathematical Theory of Communication. *Bell Labs Tech. J.*, vol. 14, no. 4, pp. 306–17, 1948.

[Shy92] J. J. Shynk. Frequency-Domain and Multirate Adaptive Filtering. *IEEE Signal Process. Mag.*, vol. 9, no. 1, pp. 14–37, 1992.

[SHZM03] O. V. Sinkin, R. Holzlohner, J. Zweck, C. R. Menyuk. Optimization of the Split-step Fourier Method in Modeling Optical-fiber Communications Systems. *J. Lightwave Technol.*, vol. 21, no. 1, pp. 61–68, 2003.

[SKP93] A. Splett, C. Kurtzke, K. Petermann. Ultimate Transmission Capacity of Amplified Optical Fiber Communication Systems taking into Account Fiber Nonlinearities. In *19th Eur. Conf. Opt. Commun.*, paper MoC2.4, 1993.

[Spi10] B. Spinnler. Equalizer Design and Complexity for Digital Coherent Receivers. *IEEE J. Sel. Top. Quantum Electron.*, vol. 16, no. 5, pp. 1180–1192, 2010.

[Sto51] G. G. Stokes. On the Composition and Resolution of Streams of Polarized Light from Different Sources. *Transactions of the Cambridge Philosophical Society*, vol. 9, p. 399, 1851.

[TDY+11] Z. Tao, L. Dou, W. Yan, L. Li, T. Hoshida, J. Rasmussen. Multiplier-Free Intrachannel Nonlinearity Compensating Algorithm Operating at Symbol Rate. *J. Lightwave Technol.*, vol. 29, no. 17, pp. 2570–2576, 2011.

[TFS+00] S. K. Turitsyn, M. P. Fedoruk, E. G. Shapiro, V. K. Mezenrsev, E. G. Turitsyna. Novel Approaches to Numerical Modeling of Periodic Dispersion-managed Fiber Communication Systems. *IEEE J. Sel. Top. Quantum Electron.*, vol. 6, no. 2, pp. 263–275, 2000.

[TTMF00] S. K. Turitsyn, E. G. Turitsyna, S. B. Medvedev, M. P. Fedoruk. Averaged Model and Integrable Limits in Nonlinear Double-periodic Hamiltonian Systems. *Physical Review E*, vol. 61, no. 3, pp. 3127–3132, 2000.

[TZF+14] Z. Tao, Y. Zhao, Y. Fan, L. Dou, T. Hoshida, J. Rasmussen. Analytical Intrachannel Nonlinear Models to Predict the Nonlinear Noise Waveform. *J. Lightwave Technol.*, vol. 33, no. 10, pp. 2111–2119, 2014.

[VB02] A. Vannucci, A. Bononi. Statistical Characterization of the Jones Matrix of Long Fibers Affected by Polarization Mode Dispersion (PMD). *J. Lightwave Technol.*, vol. 20, no. 5, pp. 811–821, 2002.

[VSB02] A. Vannucci, P. Serena, A. Bononi. The RP Method: A New Tool for The Iterative Solution of the Nonlinear Schrödinger Equation. *J. Lightwave Technol.*, vol. 20, no. 7, pp. 1102–1112, 2002.

[WBSP09] M. Winter, C.-A. Bunge, D. Setti, K. Petermann. A Statistical Treatment of Cross-Polarization Modulation in DWDM Systems. *J. Lightwave Technol.*, vol. 27, no. 17, pp. 3739–3751, 2009.

[Wei06] X. Wei. Power-Weighted Dispersion Distribution Function Long-haul Optical Transmission Links. *Opt. Lett.*, vol. 31, no. 17, pp. 2544–2546, 2006.

[Win09] M. Winter. *A Statistical Treatment of Cross-Polarization Modulation in DWDM Systems & its Application.* PhD thesis, Technischen Universität Berlin, 2009.

[XBP01] B. Xu, M. Brandt-Pearce. Modified Volterra Series Transfer Function Method and Applications to Fiber-Optic Communications. In *35th Annu. Asilomar Conf. Signals, Syst. Comput.*, pp. 23–27, 2001.

[XBP02] B. Xu, M. Brandt-Pearce. Modified Volterra Series Transfer Function Method. *IEEE Photonics Technol. Lett.*, vol. 14, no. 1, pp. 47–49, 2002.

[ZNK97] J. L. Zyskind, J. A. Nagel, H. D. Kidorf. Erbium-doped Fiber Amplifiers For Optical Communications. *Optical Fiber Telecommunications III–B*, vol. 1, no. 1, pp. 13–68, 1997.

[ZRB+14] Q. Zhuge, M. Reimer, A. Borowiec, M. O'Sullivan, D. V. Plant. Aggressive Quantization on Perturbation Coefficients for Nonlinear Pre-distortion. In *Opt. Fiber Commun. Conf.*, paper Th4D.7, 2014.

[Zwi98] D. Zwillinger. *Handbook of Differential Equations (eprint).* Academic Press, 3rd edition, 1998.

Publications, Preprints, and Presentations of the Author

Journals and Preprints

[J1] F. Frey, J. K. Fischer, and R. F. H. Fischer, "On Discrete-Time/Frequency-Periodic End-to-End Fiber-Optical Channel Models," no. *arXiv preprint 1907.06087v3*, 2020.

[J2] F. Frey, S. Stern, J. K. Fischer, and R. F. H. Fischer, "Two-Stage Coded Modulation for Hurwitz Constellations in Fiber-Optical Communications," *J. Lightwave Technol.*, vol. 38, no. 12, pp. 3135–3146, 2020, (*Invited*).

[J3] A. Aimone, F. Frey, R. Elschner, I. G. Lopez, G. Fiol, P. Rito, M. Gruner, A. C. Ulusoy, D. Kissinger, J. K. Fischer, C. Schubert, and M. Schell, "DAC-Less 32-GBd PDM-256-QAM Using Low-Power InP IQ Segmented MZM," *IEEE Photonics Technol. Lett.*, vol. 29, no. 2, pp. 221–223, 2017.

[J4] Z. Zheng, F. Frey, P. Wilke Berenguer, and J. K. Fischer, "Low-Complexity Equalization Scheme for Multicarrier Offset-QAM Systems," *IEEE Photonics Technol. Lett.*, vol. 29, no. 23, pp. 2075–2078, 2017.

[J5] J. K. Fischer, C. Schmidt-Langhorst, S. Alreesh, R. Elschner, F. Frey, P. Wilke Berenguer, L. Molle, M. Nölle, and C. Schubert, "Generation, Transmission, and Detection of 4D Set-Partitioning QAM Signals," *J. Lightwave Technol.*, vol. 33, no. 7, pp. 1445–1451, 2015.

[J6] S. Alreesh, C. Schmidt-Langhorst, F. Frey, P. Wilke Berenguer, C. Schubert, and J. K. Fischer, "Transmission Performance of 4D 128SP-QAM with Forward Error Correction Coding," *IEEE Photonics Technol. Lett.*, vol. 27, no. 7, pp. 744–747, 2015.

[J7] C. Kottke, J. K. Fischer, R. Elschner, F. Frey, J. Hilt, C. Schubert, D. Schmidt, Z. Wu, and B. Lankl, "Coherent UDWDM PON with Joint Subcarrier Reception at OLT," *Opt. Express*, vol. 22, no. 14, pp. 16 876–88, 2014.

[J8] J. K. Fischer, S. Alreesh, R. Elschner, F. Frey, M. Nölle, C. Schmidt-Langhorst, and C. Schubert, "Bandwidth-Variable Transceivers based on Four-Dimensional Modulation Formats," *J. Lightwave Technol.*, vol. 32, no. 16, pp. 2886–2895, 2014.

[J9] T. Richter, M. Nölle, F. Frey, and C. Schubert, "Generation and Coherent Reception of 107-GBd Optical Nyquist BPSK, QPSK, and 16QAM," *IEEE Photonics Technol. Lett.*, vol. 26, no. 9, pp. 877–880, 2014.

[J10] J. K. Fischer, S. Alreesh, R. Elschner, F. Frey, C. Meuer, L. Molle, C. Schmidt-Langhorst, T. Tanimura, and C. Schubert, "Experimental Investigation of 126-Gb/s 6PolSK-QPSK signals," *Opt. Express*, vol. 20, no. 26, pp. B232–B237, 2012.

[J11] R. Elschner, F. Frey, C. Meuer, J. K. Fischer, S. Alreesh, L. Molle, T. Tanimura, and C. Schubert, "Experimental Demonstration of a Format-flexible Single-Carrier Coherent Receiver using Data-Aided Digital Signal Processing," *Opt. Express*, vol. 20, no. 27, pp. 28 786–28 791, 2012.

[J12] D. Hillerkuss, R. Schmogrow, T. Schellinger, M. Jordan, M. Winter, G. Huber, T. Vallaitis, R. Bonk, P. Kleinow, F. Frey, M. Roeger, S. Koenig, A. Ludwig, A. Marculescu, J. Li, M. Hoh, M. Dreschmann, J. Meyer, S. B. Ezra, N. Narkiss, B. Nebendahl, F. Parmigiani, P. Petropoulos, B. Resan, A. Oehler, K. Weingarten, T. Ellermeyer, J. Lutz, M. Moeller, M. Huebner, J. Becker, C. Koos, W. Freude, and J. Leuthold, "26 Tbit s^{-1} Line-Rate Super-Channel Transmission Utilizing All-Optical Fast Fourier Transform Processing," *Nat. Photonics*, vol. 5, no. 6, pp. 364–371, 2011.

Conference Contributions

[C1] S. Stern, M. Barakatain, F. Frey, J. K. Fischer, and R. F. H. Fischer, "Concatenated Non-Binary Coding with 4D Constellation Shaping for High-Throughput Fiber-Optic Communication," in *Sig. Proc. Photonic Commun.*, paper SpTh2F.3, virtual, 2021.

[C2] S. Stern, M. Barakatain, F. Frey, J. Pfeiffer, J. K. Fischer, and R. F. H. Fischer, "Coded Modulation for Four-Dimensional Signal Constellations with Concatenated Non-Binary Forward Error Correction," in *46th Eur. Conf. Opt. Commun.*, paper W.1.F.4, Brussels, Belgium, 2020.

[C3] J. Pfeiffer, C. Schmidt-Langhorst, R. Elschner, F. Frey, R. Emmerich, C. Schubert, and R. F. H. Fischer, "Security Gap Investigation of Multilevel Coding in Coherent Fiber-Optical Systems," in *21st ITG-Symp. Proc. Photonic Networks*, Leipzig, Germany, 2020.

[C4] F. Frey, S. Stern, R. Emmerich, C. Schubert, J. K. Fischer, and R. F. H. Fischer, "Coded Modulation using a 512-ary Hurwitz-Integer Constellation," in *45th Eur. Conf. Opt. Commun.*, paper W.2.D.2, Dublin, Ireland, 2019, (*Highly Scored*).

[C5] C. Schmidt-Langhorst, J. Pfeiffer, R. Elschner, F. Frey, R. Emmerich, R. F. H. Fischer, and C. Schubert, "Experimental Investigation of Security Gaps when using Punctured LDPC Coding in Coherent Systems up to 64-GBd DP-64QAM," in *20th ITG-Symp. Proc. Photonic Networks*, Leipzig, Germany, 2019.

[C6] S. Stern, F. Frey, J. K. Fischer, and R. F. H. Fischer, “Two-Stage Dimension-Wise Coded Modulation for Four-Dimensional Hurwitz-Integer Constellations,” in *12th Int. ITG Conf. Syst. Commun. Coding (SCC)*, Hamburg, Germany, 2019.

[C7] R. Elschner, R. Emmerich, C. Schmidt-Langhorst, F. Frey, P. W. Berenguer, J. K. Fischer, H. Grießer, D. Rafique, J.-P. Elbers, and C. Schubert, “Improving Achievable Information Rates of 64-GBd PDM-64QAM by Nonlinear Transmitter Predistortion,” in *Opt. Fiber Commun. Conf.*, paper M1C.2, San Diego, CA, USA, 2018.

[C8] F. Frey, R. Emmerich, C. Schubert, J. K. Fischer, and R. F. H. Fischer, “Improved Perturbation-based Fiber Nonlinearity Compensation,” in *44th Eur. Conf. Opt. Commun.*, paper Tu1F.6, Rome, Italy, 2018.

[C9] M. Baier, F. M. Soares, Z. Zheng, C. Schmidt-Langhorst, R. Elschner, F. Frey, M. Gruner, J. K. Fischer, C. Schubert, M. Moehrle, N. Grote, and M. Schell, “112 Gb/s PDM-PAM4 Generation and 80 Km Transmission Using A Novel Monolithically Integrated Dual-Polarization Electro-Absorption Modulator InP PIC,” in *43rd Eur. Conf. Opt. Commun.*, paper Th.1.C.3, Gothenburg, Sweden, 2017.

[C10] P. Wilke Berenguer, F. Frey, C. Schubert, and J. K. Fischer, “Nonlinear Digital Predistortion of Transmitter Components,” in *Proc. IEEE Photonics Conf. (IPC)*, Orlando, FL, USA, 2017, pp. 431–432, (*Invited*).

[C11] F. Frey, L. Molle, R. Emmerich, C. Schubert, J. K. Fischer, and R. F. H. Fischer, “Single-step Perturbation-based Nonlinearity Compensation of Intra- and Inter-Subcarrier Nonlinear Interference,” in *43rd Eur. Conf. Opt. Commun.*, paper P1.SC3.53, Gothenburg, Sweden, 2017.

[C12] F. Frey, F. Wäckerle, J. K. Fischer, and R. F. H. Fischer, “Optimized Detection and Decoding of 4D Signals for the Nonlinear Fiber-optical Channel,” in *Sig. Proc. Photonic Commun.*, paper SpTu2F.4, New Orleans, LA, USA, 2017, (*Invited*).

[C13] F. Frey, R. Elschner, C. Schmidt-Langhorst, R. Emmerich, C. Schubert, J. K. Fischer, and R. F. H. Fischer, “Investigation of Nonlinear Inter-Channel Interference in Heterogeneous Flexible Optical Networks,” in *11th Int. ITG Conf. Syst. Commun. Coding (SCC)*, Hamburg, Germany, 2017.

[C14] F. Frey, R. Elschner, and J. K. Fischer, “Estimation of Trends for Coherent DSP ASIC Power Dissipation for Different Bitrates and Transmission Reaches,” in *18. ITG-Symp. Proc. Photonic Networks*, Leipzig, Germany, 2017.

[C15] M. Nölle, F. Frey, R. Elschner, C. Schmidt-Langhorst, J. K. Fischer, and C. Schubert, “Investigation of CAZAC Sequences for Data-Aided Channel Estimation Considering Nonlinear Optical Transmission,” in *Opt. Fiber Commun. Conf.*, paper Th3G.2, Los Angeles, CA, USA, 2015.

[C16] C. Schmidt-Langhorst, R. Elschner, F. Frey, R. Emmerich, and C. Schubert, "Experimental Analysis of Nonlinear Interference Noise in Heterogeneous Flex-Grid WDM Transmission," in *41th Eur. Conf. Opt. Commun.*, paper Tu.1.4.3, Valencia, Spain, 2015.

[C17] S. Alreesh, C. Schmidt-Langhorst, R. Elschner, F. Frey, P. Wilke Berenguer, L. Molle, M. Noelle, C. Schubert, and J. K. Fischer, "Transmission of 2048SP-QAM Nyquist-WDM Signals," in *16. ITG-Fachtagung Photonische Netze*, Leipzig, Germany, 2015.

[C18] R. Elschner, F. Frey, C. Schmidt-langhorst, J. K. Fischer, and C. Schubert, "FPGA Implementation of a Data-Aided Single-Carrier Frequency-Domain Equalizer for Format-Flexible Receivers," in *16. ITG-Fachtagung Photonische Netze*, Leipzig, Germany, 2015.

[C19] T. A. Eriksson, S. Alreesh, C. Schmidt-langhorst, F. Frey, P. Wilke Berenguer, C. Schubert, J. K. Fischer, P. A. Andrekson, M. Karlsson, and E. Agrell, "Experimental Investigation of a Four-Dimensional 256-ary Lattice-based Modulation Format," in *Opt. Fiber Commun. Conf.*, paper W4K.3, Los Angeles, CA, USA, 2015.

[C20] K. Habel, L. Fernandez, S. Weide, J. Hilt, V. Jungnickel, R. Elschner, C. Schubert, F. Frey, J. K. Fischer, and R. Freund, "5 Gbit/s Real-Time Processing Using $\pi/4$ -shift DQPSK for Bidirectional Radio-Over-Fibre System," in *17th Int. Conf. Transparent Opt. Networks*, paper Tu.D2.3, Budapest, Hungary, 2015.

[C21] R. Elschner, F. Frey, C. Schmidt-Langhorst, J. K. Fischer, and C. Schubert, "Data-Aided DSP for Flexible Transceivers," in *Sig. Proc. Photonic Commun.*, paper SW1C.1, San Diego, CA, USA, 2014, (*Invited*).

[C22] J. K. Fischer, C. Schmidt-Langhorst, S. Alreesh, R. Elschner, F. Frey, P. Wilke Berenguer, L. Molle, M. Nölle, and C. Schubert, "Transmission of 512SP-QAM Nyquist-WDM Signals," in *40th Eur. Conf. Opt. Commun.*, paper Tu.3.3.2, Cannes, France, 2014.

[C23] J. K. Fischer, R. Elschner, F. Frey, J. Hilt, C. Kottke, and C. Schubert, "Digital Signal Processing for Coherent UDWDM Passive Optical Networks," in *15. ITG-Fachtagung Photonische Netze*, Leipzig, Germany, 2014.

[C24] F. Frey, R. Elschner, C. Kottke, C. Schubert, and J. K. Fischer, "Efficient Real-time Implementation of a Channelizer Filter with a Weighted Overlap-Add Approach," in *40th Eur. Conf. Opt. Commun.*, paper Tu.3.1.3, Cannes, France, 2014.

[C25] M. Nölle, F. Frey, R. Elschner, C. Schmidt-Langhorst, A. Napoli, and C. Schubert, "Performance Comparison of Different 8QAM Constellations for the Use in Flexible Optical Networks," in *Opt. Fiber Commun. Conf.*, paper W3B.2, San Francisco, CA, USA, 2014.

[C26] R. Elschner, F. Frey, C. Schmidt-Langhorst, J. K. Fischer, M. Nölle, and C. Schubert, "Software-defined Transponders for Future Flexible Grid Networks," in *Photonic Networks and Devices (NETWORKS)*, paper NT2C.4, Rio Grande, PR, USA, 2013.

[C27] J. K. Fischer, S. Alreesh, R. Elschner, F. Frey, M. Nölle, and C. Schubert, "Bandwidth-Variable Transceivers Based on 4D Modulation Formats for Future Flexible Networks," in *39th Eur. Conf. Opt. Commun.*, paper Tu.3.C.1, London, UK, 2013.

[C28] C. Schmidt-Langhorst, F. Frey, M. Nölle, R. Elschner, C. Meuer, P. Wilke Berenguer, and C. Schubert, "Optimization of Subcarrier Spacing of 400-Gb/s Dual-Carrier Nyquist PDM-16QAM in a Flexgrid Scenario," in *39th Eur. Conf. Opt. Commun.*, paper P.5.1, London, UK, 2013.

[C29] T. Tanimura, S. Alreesh, J. K. Fischer, C. Schmidt-Langhorst, F. Frey, C. Meuer, R. Elschner, L. Molle, and C. Schubert, "Nonlinear Transmission of 6PolSK-QPSK Signals using Coded Modulation and Digital Back Propagation," in *Opt. Fiber Commun. Conf.*, paper OTu3B.3, Anaheim, CA, USA, 2013.

[C30] R. Elschner, F. Frey, C. Meuer, J. K. Fischer, S. Alreesh, C. Schmidt-Langhorst, L. Molle, T. Tanimura, and C. Schubert, "Experimental Demonstration of a Format-Flexible Single Carrier Coherent Receiver using Data-Aided Digital Signal Processing," in *38th Eur. Conf. Opt. Commun.*, paper We.1.A.4, Amsterdam, Netherlands, 2012.

[C31] J. K. Fischer, S. Alreesh, R. Elschner, F. Frey, C. Meuer, L. Molle, T. Tanimura, and C. Schubert, "Experimental Investigation of 126-Gb/s 6PolSK-QPSK Signals," in *38th Eur. Conf. Opt. Commun.*, paper We.1.C.4, Amsterdam, Netherlands, 2012.

Workshop Presentations

[W1] F. Frey, J. K. Fischer, and R. F. H. Fischer, "Aliasing of the Frequency-Domain Kernel in Perturbation-Based End-to-End Channel Models," in *21. ITG-Workshop KT 3.1: Modellierung photonischer Komponenten und Systeme*, Karlsruhe, Germany, 2020.

[W2] F. Frey, S. Stern, J. K. Fischer, and R. F. H. Fischer, "Four-Dimensional Hurwitz-Integer Constellations: Implementation Aspects and Practical Results," in *33. ITG–Workshop KT 1: Angewandte Informationstheorie*, Ulm, Germany, 2019.

[W3] F. Frey, J. K. Fischer, and R. F. H. Fischer, "Probability Distributions of Nonlinear Distortions Obtained from the Sampled Nonlinear Transfer Function," in *19. ITG-Workshop KT 3.1: Modellierung photonischer Komponenten und Systeme*, Hamburg, Germany, 2018.

[W4] F. Frey, J. K. Fischer, and R. F. H. Fischer, "On the Frequency Dependence of the FWM Efficiency and its Application to Analytical Channel Models," in *18. ITG-Workshop KT 3.1: Modellierung photonischer Komponenten und Systeme*, Munich, Germany, 2017.

[W5] F. Frey, R. Elschner, C. Schmidt-Langhorst, R. Emmerich, C. Schubert, J. K. Fischer, and R. Fischer, "Analysis of Temporal Correlations of Inter-Channel Interference in Heterogeneous Flexible Optical Networks," in *Munich Workshop on Information Theory of Optical Fiber (MIO)*, Munich Germany, 2016.

[W6] F. Frey, R. Elschner, C. Schmidt-Langhorst, M. Nölle, J. K. Fischer, and C. Schubert, "Data-Aided Signal Processing for Format Flexible Transceivers," in *1. Tobias Eriksson Workshop (TEW)*, Gothenburg, Sweden, 2015.

[W7] F. Frey, R. Elschner, J. Fischer, and C. Schubert, "Efficient-Real Time Implementation of DSP Algorithms for Coherent Receivers," in *14. ITG-Workshop KT 3.1: Modellierung photonischer Komponenten und Systeme*, Stuttgart, Germany, 2013.

Curriculum Vitæ

For data protection reasons, the curriculum vitæ has been removed from the online version.

ISBN: 4978-3-948303-12-9