APPROVAL SHEET

Title of Dissertation: A Covariance Matrix Method for the Computation of Bit Errors in Optical Transmission Systems

Name of Candidate: Ronald Holzlöhner
Doctor of Philosophy, 2003

Dissertation and Abstract Approved: __________________________
Professor Curtis R. Menyuk
Computer Science and Electrical Engineering,
University of Maryland Baltimore County

Date Approved: ________________________

Curriculum Vitæ

Name: Ronald Holzlöhner

Degree and date to be conferred: Ph.D., May 2003.

Collegiate institutions attended: 1998–2003, University of Maryland
Baltimore County, Ph.D., May 2003

1990–1998, Technische Universität Berlin (TUB),
Germany, M.S. in physics, August 1998

1995–1996, University of California,
Santa Barbara (UCSB)

Major: Photonics (Electrical Engineering).

Professional position: Research Assistant
Department of Computer Science and Electrical Engineering
University of Maryland, Baltimore County
TRC 201 B,
1000 Hilltop Circle,
Baltimore, MD 21250
USA

Publications in Refereed Archival Journals:

1. R. Holzlöhner, C. R. Menyuk, W. L. Kath, and V. S. Grigoryan, “A covariance matrix method to compute accurate bit error rates in a highly nonlinear dispersion-managed soliton system,” *IEEE Photon. Technol. Lett.*, to appear in May 2003.
2. O. V. Sinkin, R. Holzlöhner, J. Zweck, and C. R. Menyuk, “Optimization of the split-step Fourier method in modeling optical fiber communications systems.” *J. Lightwave Technol.*, vol. 21, pp. 61–68, 2003.
3. R. Holzlöhner, C. R. Menyuk, W. L. Kath, and V. S. Grigoryan, “Efficient and accurate calculation of eye diagrams and bit-error rates in a single-channel CRZ system.” *IEEE Photon. Technol. Lett.*, vol. 14, pp. 1079–1081, 2002.
4. R. Holzlöhner, H. N. Ereifej, G. M. Carter, and C. R. Menyuk, “Experimental and theoretical characterization of a 40 Gb/s long-haul single-channel transmission system.” *J. Lightwave Technol.*, vol. 20, pp. 1124–1131, 2002.
5. R. Holzlöhner, V. S. Grigoryan, C. R. Menyuk, and W. L. Kath, “Accurate calculation of eye diagrams and bit error rates in long-haul transmission systems using linearization,” *J. Lightwave Technol.*, vol. 20, pp. 389–400, 2002.
6. H. N. Ereifej, R. Holzlöhner, and G. M. Carter, “Inter-symbol interference and timing jitter measurements in a 40 Gb/s long-haul dispersion-managed soliton system,” *IEEE Photon. Technol. Lett.*, vol. 14, pp. 343–345, 2002.
7. D. Marcuse, C. R. Curtis, and R. Holzlöhner, “Time shift of pulses due to dispersion slope and nonlinearity,” *IEEE Photon. Technol. Lett.*, vol. 11, pp. 1611–1613, 1999.
8. R. Holzlöhner and M. Schoen, “Attractive forces between anisotropic inclusions in the membrane of a vesicle,” *Eur. Phys. J. B*, vol. 12, pp. 413–419, 1999.
9. R. Holzlöhner, “Schweissen mit Strom,” *ELRAD*, vol. 1, pp. 50–53, 1989.

Invited Papers:

1. C. R. Menyuk, R. Holzlöhner, I. T. Lima, Jr., B. S. Marks, and J. Zweck “Advances in modeling high data rate optical fiber communication systems” *SIAM Conference on Computational Science and Engineering (CSE03)*, (San Diego, CA), 2003.

2. J. Zweck, I. T. Lima Jr., R. Holzlöhner, and C. R. Menyuk, “New advances in modeling optical fiber communication systems,” in *Proc. OSA Topical Meeting*, (Vancouver, Canada), 2002.

3. C. R. Menyuk, R. Holzlöhner, and I. T. Lima Jr., “New approaches for modeling high data rate optical communication systems,” in *Proc. OSA Annual Meeting*, ThFF1, (Long Beach, CA), 2001.

4. C. R. Menyuk, R. Holzlöhner, and I. T. Lima Jr., “Advances in modeling optical fiber transmission systems,” in *IEEE LEOS Newsletter*, pp. 21–23, 2001.

5. C. R. Menyuk, R. Holzlöhner, and I. T. Lima Jr., “Advances in modeling of optical fiber transmission systems,” in *Proc. LEOS Summer Topical Meeting*, MD 1.2, (Copper Mountain, CO), pp. 5–6, 2001.

6. C. R. Menyuk, D. Wang, R. Holzlöhner, I. T. Lima Jr., and E. Ibragimov, “Polarization mode dispersion in optical transmission systems,” in *Proc. OFC’00*, Tutorial, (Baltimore, MD), pp. 78–97, 2000.

Contributed Papers at Conferences:

1. R. Holzlöhner and C. R. Menyuk, “Accurate bit error rates from multicanonical Monte Carlo simulations,” *Proc. CLEO’03*, talk CThJ3 (Baltimore, MD), 2003.

2. R. Holzlöhner, C. R. Menyuk, W. L. Kath, V. S. Grigoryan, "A Covariance Matrix Method for Accurate Bit Error Ratios in a DWDM CRZ System," *Proc. OFC’03*, talk ThW3 (Atlanta, GA), 2003.

3. O. V. Sinkin, R. Holzlöhner, J. Zweck, C. R. Menyuk "Optimization of the split-step Fourier method in modeling optical fiber communications systems," *Proc. OSA Annual Meeting 2002*, (Orlando, FL), 2002.

4. R. Holzlöhner, C. R. Menyuk, V. S. Grigoryan, and W. L. Kath, “Direct calculation of the noise evolution in a highly nonlinear transmission system using the covariance matrix,” in *Proc. CLEO’02*, talk CThG5, (Long Beach, CA), 2002.

5. R. Holzlöhner, C. R. Menyuk, V. S. Grigoryan, and W. L. Kath, “Efficient and accurate computation of eye diagrams and bit error rates in a single-channel CRZ system,” in *Proc. CLEO’02*, poster CThO44, (Long Beach, CA), 2002.

6. R. Holzlöhner, C. R. Menyuk, V. S. Grigoryan, and W. L. Kath, “Accurate calculation of eye diagrams and error rates in long-haul transmission systems using linearization,” in *Proc. UMBC Student Conference*, (Baltimore, MD), 2002.
7. R. Holzlöhner, H. N. Ereifej, G. M. Carter, and C. R. Menyuk, “Timing jitter in a 40 Gb/s dispersion-managed soliton system,” in *Proc. OFC’02*, talk ThQ2, (Anaheim, CA), 2002.
8. R. Holzlöhner, C. R. Menyuk, V. S. Grigoryan, and W. L. Kath, “Accurate eye diagrams and error rates using linearization,” in *Proc. ECOC’01*, poster P41, (Amsterdam, Netherlands), 2001.
9. R. Holzlöhner, C. R. Menyuk, V. S. Grigoryan, and W. L. Kath, “Accurate calculation of eye diagrams and error rates in long-haul transmission systems,” in *Proc. OFC’01*, paper MF3, (Anaheim, CA), 2001.
10. R. Holzlöhner and S. Hess, “Fluidmembranen mit anisotropen Einschlüssen,” in *Proc. DPG Spring Meeting*, (Regensburg, Germany), 1998.
11. R. Holzlöhner and J. B. Fournier, “Elastomechanik von Lipidmembranen mit anisotropen Einschlüssen,” in *Proc. DPG Spring Meeting*, (Muenster, Germany), 1997.

Unpublished Material:

1. R. Holzlöhner, “Mesoscopic and statistical properties of lipid membranes,” Master’s thesis, Technische Universität Berlin, Berlin, Germany, 1998.

Abstract

Title of Dissertation: A Covariance Matrix Method for the Computation
of Bit Errors in Optical Transmission Systems

Ronald Holzlöhner, Doctor of Philosophy, 2003

Dissertation directed by: Professor Curtis R. Menyuk
Computer Science and Electrical Engineering,
University of Maryland Baltimore County

Bit error rates (BERs) are traditionally estimated by running Monte Carlo simulations and extrapolating the results under the assumption that the electrical voltage at the receiver after narrow-band filtering is Gaussian distributed in the marks (Ones) and spaces (Zeros). This method is computationally expensive and not always reliable. Thus, it is often replaced in practice by approximations—the most common of which is to simply neglect noise altogether during propagation and add Gaussian white noise at the receiver. This and similar approaches are usually not well-validated.

The covariance matrix method that I describe in my dissertation is fully deterministic and relies on the assumption that the interaction of the optical noise in the fiber with itself is negligible. I calculate the linearized evolution of the noise around the noise-free signal. The optical noise at the receiver is multivariate Gaussian distributed after the phase jitter is separated, and therefore completely described by a covariance matrix.

I successfully apply the covariance matrix method to compute the BER in a highly nonlinear dispersion-managed soliton system over 24,000 km and to a 10 Gb/s 5-channel chirped return-to-zero (CRZ) submarine system over 6,100 km. The results are far more accurate than what those from extrapolating Monte Carlo simulations, while requiring a fraction of the computational cost.

A Covariance Matrix Method for the Computation of Bit Errors in Optical Transmission Systems

by

Ronald Holzlöhner

Dissertation submitted to the Faculty of the Graduate School
of the University of Maryland in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2003

Grau, teurer Freund, ist alle Theorie,
Und grün des Lebens goldner Baum.

Johann Wolfgang von Goethe

Acknowledgments

My years at UMBC were very rewarding from a scientific point of view, and I had the chance to interact with and learn from a number of outstanding scientists. In particular, I owe my advisor Curtis R. Menyuk many thanks for his support and for selecting challenging and interesting problems. I am impressed with his scientific insight, his vision, and his tenacity. I did not always agree with his view of organizational issues, but I feel that his unrelenting battle for scientific excellence indeed paid off. On a personal level, I always appreciated the support and concern that he has for his students and colleagues.

Only shortly after arriving at UMBC in the Fall of 1998, I started my main Ph.D. project, which eventually led to the development of the Covariance Matrix Method. I felt very fortunate to work with Vladimir Grigoryan, who helped me start work on this problem, and who made major contributions until he left our group in February 2000. During his frequent visits to UMBC, Bill Kath helped me gain a deeper understanding of the mathematical aspects, as well as the numerical solution of this method, which I am thankful for. I consider both Vladimir and Bill outstanding scientists, and it was and is a pleasure to work with them.

While theory is intriguing, comparing it with experiment is even better. Getting both to agree can be a real challenge, and I would like to thank Gary Carter and Heider Ereifej for their patience and help when I simulated their 40 Gb/s soliton system. There are few places in the academic world where theory and experiment come together in fiber optics, and UMBC is one of them.

Much, if not most, of the strength in a research team lies between the advisor and the student levels, and I owe many thanks to John Zweck and Brian Marks, as well as Gino Biondini, who I could always approach with scientific questions. All three of them have more or less recently been promoted to professors, and I wish them the best for their careers.

I enjoyed working with my fellow students. Most closely, at least spatially, I interacted with my office mate Oleg Sinkin, but I equally thank Ivan and Aurenice Lima, Jonathan

Hu, Ruomei Mu, Tao Yu, Ding Wang, Gevorg Grigoryan, Yu Sun, Hai Xu, Hua Jiao, and Jiping Wen. Furthermore, I wish to thank Tulay Adalı, Li Yan, and Joel Morris, as well as Moshe Horowitz, Gastón Tudury, Anthony Lenihan, Amit Mahadevan, Jochen Dörring, and Wen Ze Xi.

No acknowledgment would be complete without mentioning Lyn Randers, the heart and soul of the TRC building, who always had an open ear for our concerns.

Science can be fun, but it is not everything. Without my parents Ulrich and Ines Holzlöhner and their support, I certainly would not be sitting here writing these lines, and I thank them as well as my girlfriend Fadime Verhoeven.

Contents

List of Figures

Chapter 1

Introduction

Computing accurate bit error rates (BERs) is one of most important tasks when designing optical fiber transmission systems. The covariance matrix method that I derive in this dissertation yields a speedup in simulation time of orders of magnitude over the commonly used Monte Carlo simulation technique. Unlike some *ad-hoc* approaches, the covariance matrix method is based on a sound mathematical foundation that makes use of the result that the optical noise in a transmission system is multivariate-Gaussian distributed after phase and timing jitter are separated. This result is derived from the assumption that the signal evolution is linearizable, which means that the interactions of the noise with itself in the fiber are negligible once phase and timing jitter are separated. I show that this assumption is valid for currently used systems. Unlike Monte Carlo methods, the covariance matrix method is completely deterministic; it does not require random number generators. In the remainder of this introduction, I will outline the problem of computing bit error rates and provide a guide to the following chapters.

Optical fiber transmission systems offer the highest data rates among all communications systems, and their bandwidth-distance product significantly exceeds radio, microwave, satellite, and free-space optical transmission in the distance range of a few tens to thousands of kilometers. Recent experiments have demonstrated the transmission of 5 Tb/s over a distance of 1,200 km [1] and 1.5 Tb/s over a distance of 6,500 km [2] at channel data rates of 40 Gb/s. This excellent transmission quality is due to low attenuation and distortion over a very large bandwidth in modern optical fiber waveguides. In addition, the light in the fiber can be optically amplified, either using doped lumped fiber amplifiers such as erbium-doped fiber amplifiers (EDFAs) [3], or, exploiting stimulated Raman scattering in the transmission fiber itself, as in Raman amplifiers [4].

Despite these virtues, the combination of confinement of the light beam to a small effective fiber area and long propagation distance leads to a significant Kerr nonlinearity in the transmission. Optical fibers also have chromatic dispersion that interacts with the fiber nonlinearity in a complicated way. Dispersion can be used to mitigate some of the detrimental effects of nonlinearity, but in general a quote by Neal Bergano holds true: "Nonlinearity in optical fibers is always bad." [5].

The large available bandwidth of about 5 THz in optical fibers is a consequence of the high carrier frequency of 200 THz, which is more than six orders of magnitude larger than the UHF radio frequency and four orders of magnitude beyond typical satellite frequencies. However, because the photon energy is proportional to frequency, the photon flow at the typical low optical signal powers of fractions of a milliwatt to a few milliwatts per channel is small and hence the field quantization becomes relevant. As a consequence, quantization noise is significant and dominates over thermal noise in optical amplifiers [3], where it is known as amplified spontaneous emission (ASE) noise.

Modeling optical fiber communications is a young discipline. All-optical systems, in which the light propagates from the transmitter all the way to the receiver without ever being converted to electrical signals, have only been in existence since the advent of low-loss fibers and EDFAs in the mid-1990s. As the systems have become faster and more complex, physical modeling has become more important as well. In particular, fast modeling tools are required in order to design systems, tune parameters, and verify error-free operation. Given the rapid progress in commercial systems, it is not surprising that modeling algorithms are lagging somewhat behind.

The task of computing BERs essentially consists of computing the evolution of the ASE noise that is caused by optical amplifiers, taking into account its interaction with the noise-free signal due to the fiber nonlinearity. Typical BERs are in the range of 10^{-15}–10^{-9} and below, although many systems now employ forward-error correction (FEC) and can tolerate raw BERs of up to 10^{-3} [1], [6]. A commonly used method to simulate these rare events is still to run Monte Carlo simulations of the noise and extrapolate the results [7]. This approach relies on the speed of the signal propagation equation solver, which usually employs a split-step fast Fourier transform method [8] that requires a large number of discrete Fourier transforms. Running a Monte Carlo simulation is like repeatedly rolling dice or flipping a coin, and it is easy to implement on a computer and very robust [9]. However, when evaluating the probability of extremely rare events, such as that of the coin standing on its rim after being flipped, standard Monte Carlo simulations are too inefficient. One can increase the Monte Carlo simulation efficiency by biasing the perturbations toward regions of the phase space where bit errors are believed to be more likely to occur [10], [11], but the success of this method relies on a detailed knowledge of the system, and the biasing must be optimized in order to achieve a sizable performance gain over the unbiased simulation. On the other hand, this detailed system knowledge is sometimes the result of a simulation, and hence it is not known at the time the simulation is started.

Monte Carlo methods are very expensive computationally. For that reason, they are often replaced in practice with simplified models. The most common of these models is to run a noise-free simulation and then to simply add white Gaussian-distributed noise at the receiver that corresponds to the noise that would arrive at the receiver in the absence of any noise-signal interaction due to fiber nonlinearity. It is known, however, that parametric pumping of noise by the signal is a significant issue in some systems. For this reason, Hui, *et al.* [12], [13] introduced a model, later extended by Pilipetskii, *et al.* [14] that takes into account parametric pumping by treating the signal as a cw wave. Due to difficulties in carrying out extensive Monte Carlo simulations, this simplified approach is not well-validated in general.

Systems that suffer from noise are well known in many areas of science. In all communications systems there is a certain amount of noise that distorts signals during transmission or storage of information, and therefore limits the capacity [15], [16]. Unless a given physical system is completely dominated by nonlinearities, perturbation theory based on the assumption that the noise interaction with itself is negligible is always a powerful technique for computing error rates. The linearization assumption has been successfully applied to optical fiber communications and the known results can be divided into several groups:

- Soliton perturbation theory was applied to transmission with gain and loss, ASE noise, soliton-soliton interactions, filtering, and synchronous modulation [17]–[20]. However, the results are only strictly valid for classical solitons.

- Standard perturbation theory was applied to an unmodulated (cw) signal [12], [13], [21], [22].

- Computation of the evolution of noise moments such as timing and amplitude jitter has been carried out for an arbitrary signal [23].

However, none of the approaches that have been developed to date are capable of computing the full probability density function (pdf) of the optical noise in the presence of an arbitrary signal modulation and the pdf of the electrical receiver current. The goal of this dissertation is to develop a method that

1. linearizes the optical noise propagation in a wide range of transmission systems,
2. yields the accurate pdf of the optical noise over the entire transmission distance,
3. yields the pdf of the electrical receiver current after passing a narrow-band filter,
4. can be easily implemented and is computationally efficient.

I derive such an approach in this dissertation and call it the covariance matrix method. The approach is capable of computing accurate probability density functions (pdfs) of the receiver current after narrow-band filtering. With this information, the BER and contour eye diagrams can then be obtained. I develop two different methods to compute the key statistical quantity, namely the covariance matrix of the optical noise, and I show that both lead to a significant efficiency enhancement over traditional Monte Carlo simulations. The most complex system that I treat in this dissertation is a 10 Gb/s 5-channel WDM chirped return to zero (CRZ) system, in which I simulate the transmission of 32 bits.

In Chapter 2, I cover some basics of optical fiber communications systems and optical amplifier noise. I outline different previous linearization approaches in Chapter 3 and formulate the theory of the covariance matrix method in Chapter 4. I show the results of calculating the covariance matrix using Monte Carlo simulation techniques in Chapter 5 and using the fully deterministic covariance matrix method in Chapter 6. The work in Chapter 5 validates the linearization assumption. Chapter 7 contains the conclusions. The Appendix contains some mathematical details and a description of the numerical algorithms.

Chapter 2

Basics of Optical Fiber Transmission Systems

This dissertation is concerned with the nonlinear interaction of optical noise with a signal and this chapter lays the foundation for the subsequent derivation of the covariance matrix method. In Section 2.1, I introduce the most important components of a typical optical fiber transmission system. Section 2.2 contains a brief description of optical noise sources. I introduce the modified nonlinear Schrödinger equation (NLS) that describes the evolution of light in nonlinear optical fibers in Section 2.3. Finally, I discuss the effective nonlinearity in modern transmission systems in Section 2.4.

2.1 Design

Modern optical transmission systems contain a large number of individual optical components such as lasers, modulators and demodulators, multiplexers and demultiplexers, filters, optical fibers, and amplifiers. In addition, there is much electrical equipment at the receiver, such as photodiodes, electrical filters, amplifiers, and decision circuits. When simulating these components, the level of detail of the model must be appropriate for the system under study. For example, optical amplifiers are often modeled by simply multiplying the optical field by a factor; I use this method in the simulations for this dissertation. More realistic models can include ASE noise, gain saturation, the gain profile, polarization hole burning, and transients. The first step of any simulation is to simplify the transmission system and to restrict the model of the optical propagation to the essential effects. The nature of the most important effects strongly depends on the type of the optical system as I will show in Section 2.4, and the effects can be expected to vary when the data rate is increased. In particular, the amount of nonlinearity has an important impact on the evolution of the signal and the noise.

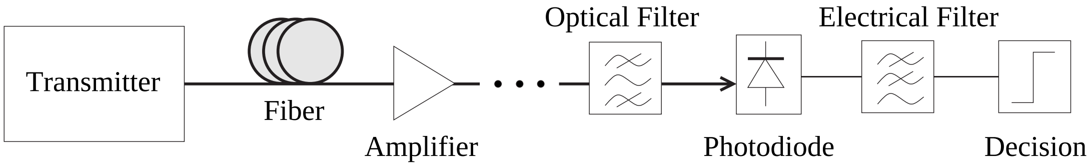

Figure 2.1: Simple optical transmission system.

Fig. 2.1 shows the schematics of a simple optical transmission system. The optical signal is generated by a transmitter, and inserted into the fiber. It then passes through a transmission line that primarily consists of fiber spans and optical amplifiers. At the end of the transmission line, the signal is optically filtered and enters a receiver, where it is converted to electrical current by a high-speed photodiode. This current is low-pass filtered and enters a decision circuit.

In a typical digital optical fiber transmission system, the pulses are either directly created by an optical laser, or the output of a constant wave (cw) laser is modulated by an external modulator. I will only consider the case of an intensity-modulated return-to-zero signal in which a light pulse or “mark” represents a logical “1” and the absence of light or “space” represents a logical “0”. The pulse stream might then be multiplexed, either by interleaving pulses in the time domain, called time division multiplexing (TDM), and/or by combining channels at different frequencies, called wavelength-division multiplexing (WDM). All recently-deployed systems employ dispersion management [8], in which the dispersion varies periodically. Each period consists of a concatenation of several fiber spans with different local fiber dispersions, and the variation of dispersion in one period is referred to as the dispersion map.

The systems that I model in this dissertation are depicted in Fig. 5.1 and in Fig. 6.4 and will be discussed at a later point.

2.2 Noise Sources in Optical Systems

All active components in a communications system produce noise that degrades the signal. These active components include the pulse source, the receiver, and, most important, the amplifiers. One can distinguish between optical and electrical noise, and this dissertation will be mainly concerned with the evolution of optical noise. The photodiode and the electrical amplifiers produce electrical noise, and I will show how it can be taken into account in principle. In this section, I will first outline a theoretical proof that all optical amplifiers must produce noise. This discussion is followed by a brief description of noise creation in the EDFAs.

I define noise as a random fluctuation of a signal that an active component *adds to the signal*. Other sources of apparent randomness, such as polarization mode dispersion (PMD) or a pseudorandom data sequence, are not considered noise. When dealing with electromagnetic signals, there are two principal sources of noise: (i) shot or quantization noise and (ii) thermal noise. Shot noise is directly connected to the quanta of the electrical field, the photons. Thermal noise is caused by the thermal fluctuations in a medium that for instance manifests itself as Brownian motion. The relative strength of these two noise sources in an optical amplifier is given by the ratio $h\nu/k_BT$, where h is Plank's constant, $\nu = c/\lambda$ is the frequency of light, c is the speed of light, k_B is Boltzmann's constant, and T is temperature [3]. At $\lambda \approx 1550$ nm and room temperature, this ratio is about 30, showing that shot noise dominates, and I will therefore neglect thermal noise in the optical domain.

We may mathematically characterize the light evolution in the fiber as a point in the phase space that consists of all Fourier amplitudes of the electric field. In principle, this space is infinite-dimensional, but in practice we restrict it to a finite number of dimensions by considering a finite time domain and a finite frequency bandwidth. Without noise, the field evolves along a deterministic path in this phase space. Once we take into account noise, however, this point becomes a cloud whose volume normalized to the signal increases as the light propagates along the fiber. Quantum mechanics dictates the minimum noise that an amplifier can add to the signal on average, and this value determines the minimum increase in the phase space volume.

A single photon at a single frequency occupies a two-dimensional subspace of this phase space. The minimum area that it can occupy, after an appropriate normalization [24], has units of action, which is energy times time. This action can be written as the product of two conjugate variables such as momentum p and position x, and Heisenberg's uncertainty principle states that this area is quantized in multiples of $\hbar/2$, where $\hbar = h/2\pi$. The uncertainty principle is usually expressed in terms of the uncertainties Δp and Δx as $\Delta p \Delta x \geq \hbar/2$. An equivalent way of writing this relation for an ensemble of photons that propagate in the same direction at the same frequency is $\Delta n \Delta\varphi \geq 1/2$, where n is the number of photons that arrive at a detector per time, $\varphi = 2\pi\nu t$ is their phase, and t is time. I now consider an amplifier with power gain G. Under the assumptions that the input signal is shot noise limited with $\Delta n \Delta\varphi = 1/2$, that the amplifier contributes added white Gaussian noise (AWGN) to the signal that increases only Δn but not $\Delta\varphi$, and that an optimized detector is used, Heffner [25] showed that the minimum added noise power P_N may be written,

$$P_N = \tfrac{1}{2} h\nu B(G-1), \tag{2.2.1}$$

where B represents the full width half maximum (FWHM) amplifier bandwidth. One way to interpret (2.2.1) is to view it as a quantization error of $h\nu/2$ in the measurement of the field intensity at the amplifier input at each frequency during an observation period of $1/B$. Detailed studies show that the noise in EDFAs is not Gaussian distributed but rather Bose-Einstein distributed [26]; however, unless the number of photons in the system is small, the distinction with the usual Gaussian approximation is negligible. That is always the case in today's communications systems.

Realistic optical amplifiers produce more noise than the quantum limit (2.2.1), although modern devices do not exceed it by much. In laser amplifiers, the ratio of the true noise power to the quantum limit is given by the *spontaneous emission factor* n_{sp}. Describing the physics of EDFAs in detail would exceed the scope of this dissertation, and I therefore just mention that quantum noise in laser amplifiers is caused by *spontaneous emission*, where electrons randomly lower their energy state by emitting noise photons. The EDFA amplifies these photons, giving rise to amplified spontaneous emission (ASE) noise. In a greatly simplified, two-state model of the erbium, the spontaneous emission factor is related to the population densities N_1 and N_2 of the ground state and excited state respectively in the erbium-doped fiber by [3]

$$n_{\text{sp}} = \frac{N_2}{N_2 - N_1}. \tag{2.2.2}$$

Typical values of n_{sp} are 1.2–2.5. Considering that current optical fibers have almost reached the physical minimum of the light attenuation due to Rayleigh scattering, the amplifier gains and, hence, the accumulated ASE noise cannot be reduced very much.

2.3 The Nonlinear Schrödinger Equation

In order to study the light evolution in optical fibers, we begin by writing the electric field in the fiber as [27]

$$\mathbf{E}(\mathbf{r},\tau) = \sqrt{\frac{\omega_0}{2\epsilon_0 c^2 \beta(\omega_0, z_p)}}\, e(z_p,\tau)\mathbf{R}(x,y) \exp\left[i\int_0^{z_p} \beta(\omega_0, z_p')dz_p' - i\omega_0\tau \right], \tag{2.3.1}$$

where $\mathbf{E}$ is the electric field vector in MKS units, while $\mathbf{r} = (x, y, z_p)$ and τ are position in the fiber and physical time, $\omega_0 \approx 10^{15}$ Hz is the carrier frequency, $\epsilon_0 = 8.85 \times 10^{-12}$ F/m is the permittivity of the vacuum, $c = 2.99 \times 10^8$ m/s is the speed of light in a vacuum, and $\beta(\omega_0, z_p)$ is the wavenumber at the carrier frequency ω_0, which may vary slowly relative to the wavelength along the fiber. The coordinate z_p measures distance along the fiber. The transverse modal field is normalized so that

$$\int dx \int dy \left|\mathbf{R}_{\perp(x,y)}\right|^2 = 1, \tag{2.3.2}$$

where $\mathbf{R}_{\perp(x,y)}$ is a vector obtained by projecting $\mathbf{R}$ onto the plane perpendicular to the propagation direction. With this normalization, $|e(z_p,t)|^2$ equals the local power. We note that we are assuming that only one polarization is propagating. From equation (2.3.1) for the electric field, it is possible to derive the modified nonlinear Schrödinger equation in the following form [8], [27],

$$i\frac{\partial u}{\partial z} + \frac{1}{2}D(z)\frac{\partial^2 u}{\partial t^2} + |u|^2 u = ig(z)u + \hat{F}(z,t), \tag{2.3.3}$$

where $u(z,t) = e(z,t)\sqrt{\gamma L_D}$, while $\gamma = n_2\omega_0/(A_{\text{eff}}c)$ is the nonlinear coefficient and L_D is a characteristic length. The distance z is normalized as $z = z_p/L_D$. We have transformed from physical time τ to retarded time $t = \tau - \int_0^z \beta'(z')dz'$. The quantity $n_2 = 2.6 \times 10^{-16}$ cm^2/W is the Kerr coefficient, and $A_{\text{eff}} = 30$–$200\,\mu\text{m}^2$ is the effective fiber core area. The quantity $D(z) = -\beta''(z)/\beta_0''$ is the normalized dispersion parameter with the local dispersion $\beta''(z)$ and a scaling dispersion β_0''. The dispersion β'' is measured in units of ps^2/nm. The characteristic length L_D equals $T_0^2/|\beta_0''|$, where T_0 is a characteristic time scale. When modeling soliton systems, it is customary to set T_0 equal to the FWHM soliton duration and β_0'' equal to the path average dispersion, in which case L_D is the *dispersive scale length* [8]. The normalized field gain coefficient $g(z)$ is

$$g(z) = \begin{cases} g_m(z), & z_m < z < z_m + L_{\text{amp}}/L_D\,, \\ -\gamma_l & \text{elsewhere}, \end{cases} \tag{2.3.4}$$

where g_m represents the gain coefficient inside the m-th amplifier normalized by L_D, which I assume to begin at $z = z_m$ and to be of length L_{amp}, and γ_l is the normalized fiber loss coefficient. Since the typical length of an EDFA is only a few meters, the dispersion and nonlinearity in the EDFA can be neglected and hence amplification and ASE noise input can be considered as lumped. For later use, I define the *nonlinear scale length* [8]

$$L_{\text{nl}} = \frac{1}{\gamma P_{\text{peak}}}, \tag{2.3.5}$$

which is the distance over which the nonlinear phase rotation is approximately 2π.

Summarizing the meaning of the terms that appear in (2.3.3), the second term on the left-hand side describes chromatic dispersion, the third term describes the Kerr nonlinearity, and the first term on the right-hand side describes fiber attenuation and gain within the amplifiers. The last term is a *Langevin* noise term and the quantity $\hat{F}$ represents the ASE white noise contribution with zero mean

$$\langle \hat{F}(z,t) \rangle = 0, \tag{2.3.6}$$

and the autocorrelation

$$\left\langle \hat{F}(z,t)\, \hat{F}^{*}(z',t') \right\rangle = 2\eta\, \delta(z-z')\, \delta(t-t'), \tag{2.3.7}$$

where

$$\eta = n_{\text{sp}} g_m \hbar \omega_0 L_D \gamma / T_0. \tag{2.3.8}$$

In the test systems that I consider in this dissertation, $n_{\text{sp}} = 1.4$–2.0 inside the fiber amplifiers and $n_{\text{sp}} = 0$ in the transmission fiber. The angular brackets $\langle \cdot \rangle$ denote the noise ensemble average, and the asterisk denotes complex conjugation. Equation (2.3.3) can be extended to include effects such as higher-order dispersion, saturated absorption, and Raman and Brillouin scattering [27]. However, saturable absorbers are not used in the test systems that I study, while Raman and Brillouin scattering effects are small and can be neglected. Mu, *et al.* [28], [29] showed that third-order dispersion does not have a strong impact on the pulse propagation in the DMS system and the CRZ system that I simulate in this dissertation. I solve (2.3.3) using a standard split-step approach [8].

2.4 From Solitons to Quasilinear Systems

When damping, noise, and the z-variation of the dispersion relation can all be neglected, the nonlinear Schrödinger equation (2.3.3) is integrable, and it possesses a set of closed-form solutions called *solitons* [30]. Solitons were originally defined as solitary waves that do not change shape in a collision, while their phases and temporal positions undergo an offset. We note that solitary waves are wave packets that propagate without changing shape except possibly in collisions with other entities. So, the original definition of solitons was highly restrictive. Over time, this definition has evolved in the optical communications community so that now virtually any return-to-zero pulse is called a soliton [29], [31]. In this section, we will use the original, restrictive definition.

The solution for a single soliton of peak power P_{peak} and duration T_0 in a lossless fiber with a constant anomalous dispersion ($\beta'' < 0$) of

$$\beta'' = -P_{\text{peak}}\gamma A^{-2} \tag{2.4.1}$$

is [20]

$$u_s(A,\tau,\Phi,\Omega;t,z) = A\,\text{sech}\big\{A\,[t-\tau(z)]\big\}\,\exp\big\{i\,[\Phi(z)-\Omega t]\big\}, \tag{2.4.2}$$

with the four parameters

A: soliton amplitude and inverse duration,
Ω: central (angular) frequency offset relative to ω_0,
Φ : soliton phase,
τ : soliton center in time.

The quantities A and Ω are arbitrary constants, while τ and Φ depend on z and satisfy the relations

$$\tau(z) = \tau_0 - \Omega z, \tag{2.4.3a}$$

$$\Phi(z) = \Phi_0 + \tfrac{1}{2}(A^2 - \Omega^2)z, \tag{2.4.3b}$$

where $\tau_0 = \tau(0)$ and $\Phi_0 = \Phi(0)$. The soliton energy is proportional to

$$E_s = \int_{-\infty}^{\infty} |u_s|^2\,dt = 2A. \tag{2.4.4}$$

Higher-order solutions called “breathers” exist whose shape oscillates in time.

Because the first-order soliton (2.4.2) is localized in time and space, it is analogous to a particle that is described by the four parameters (A,τ,Φ,Ω). The mechanism that keeps the soliton stable is the balance between dispersion and nonlinearity: The dispersion tends to spread the pulse in time, while the nonlinearity tends to compress it due to self-phase modulation (SPM). These two forces neutralize each other if the pulse shape is given by (2.4.2), and the shape remains constant while propagating in a lossless waveguide. As noted earlier, solitons pass through each other without changing their shape. They merely shift their central times and phases.

Mollenauer *et al.* [32] demonstrated in 1980 that it is possible to transmit solitons in optical fibers over large distances. The loss in optical fibers can be compensated by optical amplifiers such as EDFAs or Raman fiber amplifiers. However, as mentioned in Section 2.2, amplifiers also add noise to the signal and distort the pulse shapes. Both the soliton shape and its four parameters are affected by noise, and I will outline a soliton perturbation theory in the next chapter that describes these distortions in the case of small noise. Of the four soliton parameters, the change in the central time τ is the most severe at the receiver and leads to *timing jitter*, giving rise to bit errors. After passing through an EDFA, the central time of a soliton will have shifted slightly. Moreover, the central frequency Ω will have changed by a small amount. The latter effect, by means of (2.4.3a), in turn induces a shift in the central time in the subsequent propagation. This effect was described by Gordon and Haus in 1986 [17], [19]. The Gordon-Haus timing jitter is a severe impairment in soliton systems. I note that a analogous discussion applies to *phase jitter*, which is enhanced by fluctuations in the soliton energy A. While phase jitter is irrelevant in direct-detection receivers that I consider in this thesis, it is relevant in the context of the covariance matrix method, and I will elaborate on this issue later.

Since ASE noise is additive rather than multiplicative, one can in principle try to increase the soliton energy, but (2.4.2) shows that an increase in A also leads to shorter pulses. These pulses are harder to make, are more susceptible to modulation effects like the Raman self-frequency shift, and require more bandwidth per wavelength channel. Fortunately, a technique called dispersion management, in which the fiber with constant anomalous dispersion is replaced by fiber spans that alternate between normal $(\beta'' > 0)$ and anomalous $(\beta'' < 0)$ dispersion, can effectively reduce the nonlinearity of the soliton and thereby reduce the impact of noise. This method permits one to increase the soliton peak power by an amount known as the *energy enhancement factor* [33], [34]. In this case, the pulse shape deviates from the soliton shape (2.4.2) and is closer to a Gaussian shape. Moreover, the pulse shape is not constant, but it varies periodically with the period of the dispersion map (periodically stationary) or even with the period of the entire propagation distance. An additional advantage of dispersion management is that in a multi-channel or wavelength-division multiplexed (WDM) system, pulses in neighboring channels pass through each other rapidly, although repeatedly, which reduces and average the impairment due to nonlinear channel crosstalk.

Using dispersion management, it is possible to reduce the path average dispersion. If one reduces the peak power at the same time, one can approach the linear transmission regime, where the fiber acts as a linear waveguide. In the limit where nonlinearity is negligible, the accumulated dispersion at the receiver is zero, and noise is absent, one finds that any pulse shape is preserved. Traditionally, the non-return to zero (NRZ) modulation format has been popular in the communications community because the pulses are relatively easy to generate, and because it is spectrally efficient, which means that at a given data rate it occupies less bandwidth that other easily produced modulation formats [31], [35]. However, in 1996, Bergano *et al.* [36] showed that the transmission quality can be improved by subjecting the NRZ signal to a bit-synchronous phase modulation or *chirp*. Applying chirp reduces the spectral efficiency, but it was found that the error-free transmission distance could be increased in some cases. Another consequence is that the optical power at the receiver appears in the form of separated return-to-zero (RZ) pulses. This development led to the implementation of chirped RZ (CRZ) modulation, one of the modulation formats that I will study in the simulations in this dissertation.

All commercial and almost all academic optical transmission systems to date employ pulse formats that fall somewhere in the middle between the classical soliton and the linear regime. In fact, it is possible to go continuously from one limiting case to the other, and it is hard to distinguish wavelength-division-multiplexed (WDM) dispersion-managed soliton (DMS) systems, where the pulse shapes do not repeat periodically, from CRZ systems [31]. Both operate at per-channel peak powers of a few mW and are called "quasilinear." Periodically stationary DMS systems [37] form a class by themselves as they operate at higher peak powers of a few mW and above and are significantly nonlinear in the sense that the solitons would spread dispersively in the absence of fiber nonlinearity. As I will show in subsequent chapters, this significant nonlinearity is a challenge and stringent test for the covariance matrix method that is described in this dissertation.

Chapter 3

Transmission Linearization

3.1 Introduction

The covariance matrix method that I describe in this dissertation can be divided into two steps:

1. the computation of the covariance matrix that describes the accumulation of ASE noise from the optical amplifiers and its evolution due to optical fiber nonlinearity, and
2. the computation of the pdf of the electrical current in the receiver after a photodiode and an electrical filter.

As I will show in the next chapters, the first step, in which we treat transmission through an optical fiber with a series of erbium-doped fiber amplifiers along the path, is amenable to a linearization approach that assumes that the nonlinear interaction of the optical noise with itself in the fiber is negligible once phase jitter and, in some cases, timing jitter are separated. By contrast, the noise interaction with itself in the receiver cannot be ignored. The aim of this chapter is to describe some of the previous work that was done on transmission linearization and to outline the limitations of the linearization approach.

As I noted in the Introduction, most of the previous work on optical transmission linearization falls into one of three categories: One is soliton perturbation theory, described in Section 3.2, and the second may be called *cw linearization*, described in Section 3.3. The two categories differ in the zero-order signal about which the noise fluctuates; in the first case it is the standard soliton, and in the latter it is cw (continuous wave) radiation, which is unmodulated light. In both cases, the simple form of the zeroth-order signal makes it possible to use analytical approaches, which cannot be used with the modulation formats that are currently employed in communications systems. Nevertheless, both cases yield important insights. The third category was recently pioneered by Grigoryan *et al.* [23], who have shown that important scalar noise moments such as amplitude and timing jitter can be derived by linearizing around an arbitrary signal, and I briefly review the results in Section 3.4. Finally, I discuss the limits of linearization in Section 3.5.

The starting point of the linearization approach is the nonlinear Schrödinger equation (2.3.3). One can decompose the noisy optical signal $u = u_0 + \delta u$ as a sum of a noise-free signal, $u_0 = \langle u \rangle$, and accumulated transmitted noise, δu. The accumulated noise is the sum of all the individual contributions and experiences nonlinear interaction with the signal, as well as dispersion, attenuation, and amplification. The angular brackets $\langle\cdot\rangle$ denote the noise ensemble average. The difference of (2.3.3) and the statistical average of (2.3.3) then yields the evolution equation for δu

$$i\frac{\partial\,\delta u}{\partial z} + \frac{D}{2}\frac{\partial^2\,\delta u}{\partial t^2} + 2\,|u_0|^2\,\delta u + u_0^2(\delta u)^* = ig\,\delta u + \hat{F}, \tag{3.1.1}$$

where the terms that are quadratic and third order in δu, describing the optical noise-noise interactions, are omitted. The third and fourth terms on the left-hand side stem from the nonlinear term $|u|^2 u$ in (2.3.3) and describe the interactions of the noise-free signal u_0 with the noise δu. The term $2\,|u_0|^2\,\delta u$ leads to a phase rotation that can be regarded as a cross-phase modulation (XPM) between u_0 and δu. The term $u_0^2(\delta u)^*$ describes an energy exchange between u_0 and δu. Since $|u_0^2| \gg |\delta u|^2$, energy mainly flows from the signal to the noise and leads to *parametric growth* of the noise, an effect that I will describe in Chapters 5 and 6. Parametric gain can be viewed as a four-wave mixing (FWM) in which the annihilation of two signal photons at ω_1 create two noise photons at ω_3 and ω_4, so that $\omega_4 = 2\omega_1 - \omega_3$ [8]. The energy conversion process is most efficient if the frequency mismatch $|\omega_1 - \omega_3| = |\omega_1 - \omega_4|$ is small. We note that the linearized NLS (3.1.1) depends on both δu as well as on its complex conjugate $(\delta u)^*$. As a consequence, (3.1.1) is not self-adjoint, which substantially complicates the solution, as I will show in the next section. By contrast, the original nonlinear equation (2.3.3) is self-adjoint.

3.2 Soliton Perturbation Theory

Soliton perturbation theory describes the stability of a soliton in the presence of small perturbations and was derived and frequently used between the 1970s and the 1990s. The mathematical theory is beautiful, but somewhat complex, and even a summary of the results tends to be lengthy. However, soliton perturbation theory is the key to understanding the jitter in optical systems, such as Gordon-Haus timing jitter [17], [19], and is a prerequisite to subsequent chapters.

I quote results from the papers by Georges [20] and Kaup [18], but mainly follow Iannone [38]. The perturbed NLS is

$$\frac{\partial u}{\partial z} = \frac{i}{2}\frac{\partial^2 u}{\partial t^2} + i|u|^2 u + P, \tag{3.2.1}$$

where P is a small perturbation. The exact solution of the unperturbed NLS ($P = 0$) is known. I now set $u = u_0 + \delta u \exp(iA^2 z/2)$, where $u_0 = u_s = u_s(A, \tau, \Phi, \Omega; t, z)$ is the soliton solution (2.4.2). The first-order perturbation expansion is then

$$\frac{\partial \delta u}{\partial z} = \mathcal{L}[\delta u] + P, \tag{3.2.2}$$

where $\mathcal{L}$ is a linear operator given by

$$\mathcal{L}[\delta u] = i\left(\frac{1}{2}\frac{\partial^2}{\partial t^2} + 2\left|u_0\right|^2 - \frac{1}{2}A^2\right)\delta u + iu_0^2(\delta u)^*. \tag{3.2.3}$$

Any perturbation P can now be expanded in the spectrum of $\mathcal{L}$, that is, decomposed into a sum of multiples of the eigenfunctions of $\mathcal{L}$. The expansion coefficients then yield the perturbations of the soliton parameters $\{A, \tau, \Phi, \Omega\}$ to first order. First, I define the scalar product

$$(g, h) = \text{Re}\int_{-\infty}^{\infty} g(t)h^*(t)\,dt = \text{Re}\int_{-\infty}^{\infty} h(t)g^*(t)\,dt, \tag{3.2.4}$$

and the adjoint operator $\overline{\mathcal{L}}$

$$(\mathcal{L}[g], f) = (g, \overline{\mathcal{L}}[f]), \tag{3.2.5}$$

note that $\overline{\mathcal{L}}[if] = -i\mathcal{L}[f]$. Similar to the spectrum of the stationary eigensystem of a hydrogen atom [24], the spectrum of $\mathcal{L}$ consists of discrete (bound) modes and a continuous part. The discrete part has four modes, one each for the four soliton parameters A, Φ, Ω, and τ, in contrast to the hydrogen operator that has a countably infinite number of bound states. However, there are two complications: First, $\mathcal{L}$ is not a unitary operator so that $\mathcal{L}\overline{\mathcal{L}} \neq \mathcal{I}$, where $\mathcal{I}$ is the identity operator, implying that the eigenvectors of $\mathcal{L}$ do not form a single orthonormal basis, but rather there are "left" and "right" eigenvectors.

Second, $\mathcal{L}$ has a zero eigenvalue. I denote the four discrete generalized eigenfunctions of $\mathcal{L}$ by f_A, f_Φ, f_Ω, and f_τ, and analogously the generalized discrete eigenfunctions of $\overline{\mathcal{L}}$ by $\overline{f}_A$, $\overline{f}_\Phi$, $\overline{f}_\Omega$, and $\overline{f}_\tau$. These functions satisfy

$$\begin{aligned}
\mathcal{L}[f_A] &= -iA\overline{f}_A = Af_\Phi, && \text{(3.2.6a)}\\
\mathcal{L}[f_\Phi] &= 0, && \text{(3.2.6b)}\\
\mathcal{L}[f_\Omega] &= -iA\overline{f}_\Omega = f_\tau, && \text{(3.2.6c)}\\
\mathcal{L}[f_\tau] &= 0. && \text{(3.2.6d)}
\end{aligned}$$

The identities $\mathcal{L}[f_\Phi] = 0$ and $\mathcal{L}[f_\tau] = 0$ are manifestations of the phase invariance and translational invariance of the NLS respectively.[1] the eigenfunctions satisfy the orthonormality relations

$$\begin{aligned}
(f_i, \overline{f}_j) &= \delta_{i,j}, && \text{(3.2.7a)}\\
(f_i, f_j) &= (\overline{f}_i, \overline{f}_j) = 0, \quad \text{if} \quad i \neq j, && \text{(3.2.7b)}
\end{aligned}$$

where $i, j = A, \Phi, \Omega, \tau$. The remaining products are $(f_A, f_A) = (\overline{f}_\Phi, \overline{f}_\Phi) = (2/3A)(1 + \pi^2/12)$, $(f_\Phi, f_\Phi) = (\overline{f}_A, \overline{f}_A) = 2A$, $(f_\Omega, f_\Omega) = A^2(\overline{f}_\tau, \overline{f}_\tau) = \pi^2/(6A)$, and $(f_\tau, f_\tau) = A^2(\overline{f}_\Omega, \overline{f}_\Omega) = 2A^3/3$. The explicit form of the eigenfunctions of $\mathcal{L}$ is

$$\begin{aligned}
f_A &= -i\overline{f}_\Phi = \frac{1}{A}\{1 - A(t-\tau)\tanh[A(t-\tau)]\}\,u_s, && \text{(3.2.8a)}\\
f_\Phi &= i\overline{f}_A = iu_s, && \text{(3.2.8b)}\\
f_\Omega &= -iA\overline{f}_\tau = -i(t-\tau)u_s, && \text{(3.2.8c)}\\
f_\tau &= iA\overline{f}_\Omega = A\tanh[A(t-\tau)]\,u_s, && \text{(3.2.8d)}
\end{aligned}$$

and the generalized eigenfunctions of $\overline{\mathcal{L}}$ are

$$\begin{aligned}
\overline{f}_A &= -if_\Phi = u_s, && \text{(3.2.9a)}\\
\overline{f}_\Phi &= if_A = \frac{i}{A}\{1 - A(t-\tau)\tanh[A(t-\tau)]\}\,u_s, && \text{(3.2.9b)}\\
\overline{f}_\Omega &= -\frac{i}{A}f_\tau = -i\tanh[A(t-\tau)]\,u_s, && \text{(3.2.9c)}\\
\overline{f}_\tau &= \frac{i}{A}f_\Omega = \frac{t-\tau}{A}u_s. && \text{(3.2.9d)}
\end{aligned}$$

Any accumulated noise δu can be expanded as

$$\delta u = f_A\delta A + f_\Phi\delta\Phi + (f_\Omega - \tau f_\Phi)\delta\Omega + f_\tau\delta\tau + \delta u_c, \qquad \text{(3.2.10)}$$

---

[1] If $u(z,t)$ is a solution of (3.2.1), then $\exp(i\varphi)u(z,t)$ and $u(z+z', t+t')$ are solutions as well.

where δA, $\delta\Phi$, $\delta\Omega$,and $\delta\tau$ denote deviations in the four soliton parameters, and δu_c is the residual or continuous part of the noise that is orthogonal to the $\overline{f}_i$ so that $(\overline{f}_i, \delta u_c) = 0$. The residual noise disperses out as the soliton propagates and does not interact with the soliton to first order. With the help of the projections of the perturbation P on the $\overline{f}_i$, one can derive the following differential equations for the four parameters

$$
\begin{aligned}
\frac{dA}{dz} &= (P, \overline{f}_A), && \text{(3.2.11a)}\\
\frac{d\Phi}{dz} &= (P, \overline{f}_\Phi) + \frac{A^2 - \Omega^2}{2} + \tau \frac{d\Omega}{dz}, && \text{(3.2.11b)}\\
\frac{d\Omega}{dz} &= (P, \overline{f}_\Omega), && \text{(3.2.11c)}\\
\frac{d\tau}{dz} &= (P, \overline{f}_\tau) - \Omega. && \text{(3.2.11d)}
\end{aligned}
$$

Let us now assume that the perturbation P is due to ASE noise, $P = \hat{F}$ and consider the projection

$$\hat{F}_i \equiv (\overline{f}_i, \hat{F}). \tag{3.2.12}$$

Since the noise components satisfy (2.3.7)

$$
\begin{aligned}
\left\langle \hat{F}(z,t) \right\rangle &= 0, && \text{(3.2.13a)}\\
\left\langle \hat{F}(z,t)\, \hat{F}^*(z',t') \right\rangle &= 2\eta\, \delta(z - z')\, \delta(t - t'), && \text{(3.2.13b)}
\end{aligned}
$$

where $\eta = n_{\text{sp}} g_m \hbar\omega_0 L_D \gamma / T_0$ as defined in (2.3.8), one can infer that

$$
\begin{aligned}
\left\langle \hat{F}_i \right\rangle &= (\overline{f}_i, \langle \hat{F} \rangle) = 0, && \text{(3.2.14a)}\\
\left\langle \hat{F}_i \hat{F}_j \right\rangle &= \left\langle (\overline{f}_i, \hat{F})(\overline{f}_j, \hat{F}) \right\rangle = \eta(\overline{f}_i, \overline{f}_j). && \text{(3.2.14b)}
\end{aligned}
$$

Inspection of (3.2.7b) shows that the only non-zero correlations are the $\left\langle (\hat{F}_i)^2 \right\rangle$, and hence the perturbations in the four soliton parameters are uncorrelated if they are driven by white noise. The four variances of the projections are

$$
\begin{aligned}
\left\langle (\hat{F}_A)^2 \right\rangle &= 2\eta A, && \text{(3.2.15a)}\\
\left\langle (\hat{F}_\Omega)^2 \right\rangle &= \frac{2\eta A}{3}, && \text{(3.2.15b)}\\
\left\langle (\hat{F}_\tau)^2 \right\rangle &= \frac{\pi^2 \eta}{6A^3}, && \text{(3.2.15c)}\\
\left\langle (\hat{F}_\Phi)^2 \right\rangle &= \frac{2\eta}{3A}\left(1 + \frac{\pi^2}{12}\right). && \text{(3.2.15d)}
\end{aligned}
$$

With these expectation values, (3.2.11a)–(3.2.11d) can be solved to yield

$$\begin{aligned}
\left\langle \delta A(z)^2 \right\rangle &= \left\langle (\hat{F}_A)^2 \right\rangle z, && (3.2.16a)\\
\left\langle \delta \Omega(z)^2 \right\rangle &= \left\langle (\hat{F}_\Omega)^2 \right\rangle z, && (3.2.16b)\\
\left\langle \delta \Phi(z)^2 \right\rangle &= \left\langle (\hat{F}_\Phi)^2 \right\rangle z + \left\langle (\hat{F}_A)^2 \right\rangle A^2 \frac{z^3}{3}, && (3.2.16c)\\
\left\langle \delta \tau(z)^2 \right\rangle &= \left\langle (\hat{F}_\tau)^2 \right\rangle z + \left\langle (\hat{F}_\Omega)^2 \right\rangle \frac{z^3}{3}. && (3.2.16d)
\end{aligned}$$

All four parameter variances contain a part that grows linearly over distance with a magnitude that is determined by the projection of the noise on the adjoint eigenfunction. The linear growth of the variances is the power law of a one-dimensional random walk [39]. However, phase jitter is also driven by amplitude jitter, and timing jitter is also driven by frequency jitter. These indirect effects become stronger than the direct perturbations of the phase and the central time as z becomes large. Equation (3.2.16d) states that the timing jitter of an unfiltered classical soliton, $\left\langle \delta\tau(z)^2 \right\rangle^{1/2}$, grows proportionally to $z^{3/2}$ for large z, the well known result first derived by Gordon and Haus [17]. Besides the timing jitter that is induced by frequency shifts $\delta\Omega$, there is a direct shift in the central time whose standard deviation grows as $z^{1/2}$. The ratio of the two contributions to the variance is [38]

$$\frac{\left\langle (\hat{F}_\Omega)^2 \right\rangle z^3}{3\left\langle (\hat{F}_\tau)^2 \right\rangle z} = \frac{4}{3\pi} z^2 A^4. \qquad (3.2.17)$$

If one requires that this ratio be larger than unity, one arrives at the condition $Z \gg 2.72\, L_{\text{nl}}$, where Z is transmission distance in meters and $L_{\text{nl}} = 1/(\gamma P_{\text{peak}})$ is a nonlinear scale length of the soliton (2.3.5). With typical values of $\gamma = 2.1\,(\text{W}\cdot\text{km})^{-1}$ and $P_{\text{peak}} = 8$ mW, one obtains $Z \gg 160$ km, which is small compared to the typical transmission distance in a soliton system. This comparison shows that it is sensible to assume that timing jitter grows like $z^{3/2}$ in *unfiltered* optical communications using solitons. This large timing jitter poses a serious problem in soliton transmission systems. The phase jitter $\left\langle \delta\Phi(z)^2 \right\rangle^{1/2}$ grows in an analogous way, except that the jitter is proportional to A, not $A^{1/2}$ as in the case of the timing jitter. These large growth rates are another manifestation of the phase and timing invariances of the NLS. The distance at which the cubic and the linear contributions to the phase jitter in (3.2.16c) equal each other is $Z = 1.35\, L_{\text{nl}} \approx 80$ km.

I now revisit equations (3.2.11a)–(3.2.11b) and try to interpret their meaning in the context of noise. They state that the deviation of the four soliton parameter contains the projection of the noise onto the eigenfunctions $\left\langle \hat{F}_i \right\rangle$. All of the adjoint eigenfunctions

$\overline{f}_i$ are proportional to $u_s(t,z)$ multiplied by a bounded or a linear function. They are all localized in time since sech(t) decays rapidly toward infinity. According to Parseval's theorem, the projection or scalar product between two functions $g(t)$ and $h(t)$ can be written as

$$(g,h) = \int_{-\infty}^{\infty} g(t)h^*(t)\,dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} \widetilde{g}(\omega)\tilde{h}^*(\omega)\,d\omega, \tag{3.2.18}$$

where the asterisk denotes the complex conjugate and the tilde denotes the Fourier transform. Applying (3.2.18) to (3.2.11a)–(3.2.11d) shows that only the frequency components of the noise that overlap with the spectrum of the eigenmodes are relevant. This result is well known and can be applied to all signals $u_0(t,z)$. As a consequence, it is sufficient to focus on the range of the noise spectrum where the signal spectrum is nonzero. Other noise components do not interact with the signal in the fiber to first order. Solitons with short duration, corresponding to large values of A, have a wider spectrum and hence incorporate more noise. (I note that this rule does not hold at the receiver, which includes a square law detector.)

This reasoning leads to the natural question of how many modes are relevant in soliton perturbation theory. I note that the following discussion is approximate and qualitative. Suppose one samples the function $u_s(t)$ so that it is represented by N samples in the range $\pm T_s/2$, where $T_s = \rho T_{\text{FWHM}}$ is a multiple of the full width soliton duration at half the peak power. If one requires that a discrete Fourier transform yield N samples that lie in the range $\pm B_s/2$, with $B_s = \rho B_{\text{FWHM}}$ with the FWHM soliton bandwidth B_{FWHM}, one obtains the relation $N = T_s B_s = \rho^2 T_{\text{FWHM}} B_{\text{FWHM}} = 0.63\,\rho^2$. If one chooses T_s and B_s so that the soliton power does not exceed 1% of its peak value outside the time interval of $-T_s/2 < t < T_s/2$ and outside the frequency interval of $-B_s/2 < f < B_s/2$, one obtains $\rho \approx 3.4$ and $N \approx 7.3$, where I take advantage of the fact that the Fourier transform of the sech-function is again a sech-function. Thus, a small number of modes, only 7–8, suffices to simulate a soliton in the presence of noise. The above Fourier decomposition with a period of T_s corresponds to an infinite train of solitons and if this signal passes an EDFA, it follows from (2.2.1) that the noise power $\frac{1}{2}h\nu B_s(G-1)n_{\text{sp}}$ will be added to each soliton. If, on the other hand, one neglects the soliton overlap in the tails and applies soliton perturbation theory to the soliton at $t = 0$, then the four soliton eigenmodes f_A, f_Φ, f_τ, and f_Ω will each receive an average energy of $\frac{1}{2}h\nu B_s(G-1)n_{\text{sp}}$ as well. The rest of the noise energy that overlaps with the central soliton in both the time and the frequency domains has approximately $N-4$ times as much power as any of the discrete modes receives. Hence, one can decompose the ASE noise into three parts: a first part that perturbs the soliton parameters, a second part consisting of the portion of the noise continuum that overlaps with the soliton in both frequency and time, and a third part that neither perturbs the soliton parameters nor overlaps with the soliton. The first two parts receive about equal energy.

To summarize the results of soliton perturbation analysis, one finds that a perturbation to the soliton, such as ASE noise, causes the four soliton parameters amplitude, frequency, central time and phase to fluctuate. The rest of the noise produces a radiation continuum and disperses. Noise can be expanded into the four eigenfunctions and the continuum, and the parameter perturbations are given by the projections of the noise on the adjoint eigenfunctions $\overline{f}_i$. Although this dissertation does not deal with standard solitons, these qualitative observations will be of importance in the remainder of this chapter and in the next chapter.

3.3 CW Linearization

In addition to linearizing the NLS around a zeroth-order soliton solution, the NLS can be linearized around a constant power solution. In this case, the theory is quite simple. Starting with the lossless NLS,

$$i\frac{\partial u}{\partial z} - \frac{\beta''}{2}\frac{\partial^2 u}{\partial t^2} + \gamma |u|^2 u = 0, \tag{3.3.1}$$

and linearizing around an unmodulated signal, or cw radiation of power P_0 [8], [40], we write,

$$u(t,z) = \left[\sqrt{P_0} + a(t,z)\right] \exp\left(i\phi_{NL}\right), \tag{3.3.2}$$

where $a(t,z)$ is the time-dependent perturbation. Equation (3.3.2) solves (3.3.1) with the nonlinear phase rotation $\phi_{NL} = P_0 z$. Without going into the details of this calculation, I will just state that one obtains a linearized NLS for $a(t,z)$ similar to (3.1.1). Upon making a plane-wave *ansatz* of the form

$$a(t,z) = a_0 \exp\left[i(kz - \omega t)\right], \tag{3.3.3}$$

where a_0 is a complex number, one obtains the dispersion relation

$$k = \pm\tfrac{1}{2}|\beta''|\omega\sqrt{\omega^2 + \mathrm{sgn}(\beta'')\omega_c^2}, \tag{3.3.4}$$

where $\omega_c = 4/(|\beta''|L_{NL})$ with $L_{NL} = 1/(\gamma P_0)$. Equation (3.3.4) shows that the stability of the perturbed system depends on the sign of β''. If the dispersion is normal ($\beta'' > 0$), then the wave number k will be real for all ω, resulting in a noise radiation (3.3.3) that is stable in the presence of small perturbations. However, if the dispersion is anomalous ($\beta'' < 0$), then k becomes imaginary for $\omega < \omega_c$, and the perturbation grows exponentially with z. This instability is called the *modulational instability* and implies that cw radiation in an optical fiber with anomalous dispersion will eventually break up and form a periodic pulse train, along with background radiation. The gain spectrum of the modulational instability induced by a cw peak consists of two symmetric sidelobes with their maxima offset by $\omega_{\max} = \pm\left(2\gamma P_0/|\beta''|\right)^{1/2}$ and the value $g_{\max} = 2\gamma P_0$ [8].

The discussion above provides neither for amplification and attenuation, nor for variable dispersion. Hui *et al.* [12], [13] removed this limitation by dividing the simulated transmission path into pieces and applying the cw linearization separately. The result for constant dispersion confirms the simple calculation in (3.3.1)–(3.3.4) by showing that sidelobes are only produced in the presence of anomalous dispersion. The presence of a strong signal leads to noise growth in the wavelength vicinity, and the additional gain that the noise experiences is called *parametric gain* [8]. The computational method starts with the linearized NLS (3.1.1) and its complex conjugate. This system of equations is then solved for each fiber span by multiplication with a transfer matrix, where the matrix depends on the locally constant fiber parameters and signal power. In this work, the noise is represented by a vector containing the values of the discrete noise spectrum. Moreover, Hui *et al.* [12], [13] considered dispersion-managed links and the optimization of dispersion maps to reduce the total noise in a single-channel system.

Hui's analysis was extended by Bosco *et al.* [21] and Carena *et al.* [22] to WDM systems, who also computed the pdfs of the receiver current and the BER. They split the noise into its in-phase and quadrature components, $a_p(t,z)$ and $a_q(t,z)$

$$u(t,z) = \left[\sqrt{P_0} + a_p(t,z) + i a_q(t,z)\right] \exp\left(i\phi_{NL}\right). \tag{3.3.5}$$

Bosco and Carena implicitly assume that a_p and a_q are Gaussian distributed and consider the autocorrelations $\langle a_p a_p \rangle$ and $\langle a_q a_q \rangle$ and the crosscorrelations $\langle a_p a_q \rangle$. They show that the growth ratio of a_p and a_q depends on the fiber dispersion. In normal dispersive fiber, the quadrature component grows much faster than the in-phase component [41]. The in-phase noise component distorts the signal power, while the quadrature noise component leads to phase noise that is irrelevant in systems with square-law detectors that are used in the vast majority of all experimental and commercial systems. The reduction of the in-phase noise component at the expense of the quadrature noise component is called *noise squeezing* [42] and can indeed improve the BER, but is very difficult to achieve in long-haul WDM systems.

In conclusion, the cw linearization approach shows that the stability of noise growth depends critically on the fiber dispersion. However, linearizing around cw radiation implies that the signal-noise interactions in the presence of a modulated signal cannot be calculated. Pilipetskii *et al.* [14] report on a method in which they apply the noise statistics computed in a cw linearization to the noise-free modulated signal, but this method obviously does not properly take into account the nonlinear signal-noise interaction during the transmission, and it is not well-validated.

3.4 Noise Moments

Considering a single pulse, the most important moments are the pulse energy U, the central time t_p, and the central frequency $\mathbf{\Omega}_p$, defined as

$$U = \int_0^{T_{\text{bit}}} |u|^2 \, dt, \tag{3.4.1a}$$

$$t_p = \frac{1}{U} \int_0^{T_{\text{bit}}} t|u|^2 \, dt, \tag{3.4.1b}$$

$$\mathbf{\Omega}_p = \frac{1}{U} \int_{-\infty}^{\infty} \omega |\tilde{u}(\omega)|^2 \, d\omega = \frac{1}{U} \text{Im} \int_0^{T_{\text{bit}}} u_t^* u \, dt, \tag{3.4.1c}$$

where T_{bit} is the inverse of the data rate, $\tilde{u}(\omega)$ is the Fourier transform of $u = u(t)$, and $u_t = \partial u(t)/\partial t$. In the case of the soliton, $u = u_s(A, \alpha, \mathbf{\Phi}, \omega; t, z)$, $U = 2A$, $t_p = \tau$, and $\mathbf{\Omega}_p = \mathbf{\Omega}$.

As I discussed in the last two chapters, one would like to linearize around a computationally determined solution u_0 of the signal. Grigoryan *et al.* [23] generalized the Gordon-Haus result in 1999 to arbitrary signal shapes and extended the analysis to different kinds of jitter. They found that the deviation of t_p and $\mathbf{\Omega}_p$ with distance is given by projections of the noise on the field

$$\frac{\partial t_p}{\partial z} = D\mathbf{\Omega}_p + \frac{i}{U} \int_0^{T_{\text{bit}}} (t - t_p) \left(u\hat{F}^* - u^*\hat{F} \right) dt, \tag{3.4.2}$$

$$\frac{d\mathbf{\Omega}_p}{dz} = -\frac{i\mathbf{\Omega}_p}{U} \int_0^{T_{\text{bit}}} \left(u\hat{F}^* - u^*\hat{F} \right) dt - \frac{1}{U} \int_0^{T_{\text{bit}}} \left(u_t\hat{F}^* - u_t^*\hat{F} \right) dt, \tag{3.4.3}$$

where D is the local dispersion. Linearizing the propagation equation around an arbitrary field u, one can now numerically compute the evolution of the timing jitter

$$\sigma_p = \left(\left\langle t_p^2\right\rangle - \left\langle t_p\right\rangle^2\right)^{1/2} \tag{3.4.4}$$

and the frequency jitter $\sigma_\Omega = \left\langle (\delta\Omega)^2\right\rangle^{1/2}$. Grigoryan *et al.* [23] successfully applied this method to DMS, RZ, and NRZ systems.

3.5 Limits of Linearization

What lies beyond linearization? Is it possible and is it even necessary to derive results for transmission systems in which noise-noise interactions are relevant? Mecozzi [43] has taken an interesting approach by neglecting the dispersion terms in the NLS and, using Itô's formalism [44], derives exact expressions for arbitrary field averages of the form $\langle u(t,z)^n\,(u^*(t,z))^m\rangle$. I note that neglecting the dispersion is a very strong simplification, and hence the resulting equation describes a system which is significantly different from a realistic optical communications system. Mecozzi is able to define characteristic distances that define two regimes of propagation, a linear regime where the noise is additive, and a nonlinear one where significant signal-noise interactions occur. Moreover, he shows that the signal spectrum broadens due to SPM and the signal-noise interactions add a noise background to the signal spectrum. However, the results are strongly limited by the constraint that the dispersion equals zero.

A traditional method for obtaining the full pdf of the optical noise at the receiver is to run Monte Carlo simulations that pick random noise realizations and hence sample the noise ensemble. From this noise distribution, one can then calculate the distribution of the receiver current. This method is robust and works in principle in all settings, including situations with strong amplifier gain saturation, frequency-dependent dispersion, and strong nonlinearity. This method can handle arbitrarily strong pulse distortion, including complete pulse breakdown. In addition, it is easy to program. If a system is to be investigated in a significantly nonlinear regime, then Monte Carlo simulations might be one of the few possible options. However, I will show in the following chapters that the linearization assumption is valid in a soliton system, whose error-free (BER < 10^{-9}) transmission distance is hundreds of times larger than the nonlinear scale length, once phase and timing jitter are properly treated. Consequently, I claim that linearization is capable of describing the noise evolution in a wide range of transmission systems—probably any practical system.

Chapter 4

Statistics of the Noise Evolution

4.1 Theoretical Foundation

In this chapter, I describe a covariance matrix method that yields a complete and accurate statistical description of the optical noise at the end of the transmission line. In Section 4.6, I develop exact equations to compute the pdf of the received current from this information. The only limitation of the procedure is that the noise-noise interactions *in the fiber* are neglected in an appropriate basis set in which phase and in some cases timing jitter are separated. The receiver contains a square law detector and hence it is substantially nonlinear. Hence, we retain noise-noise terms in the receiver.

The utility of the linearization assumption stems from two key mathematical results. The first is the Karhunen-Loève theorem [45], which states that a combination of signal and noise over any finite time can be expanded in an orthonormal basis whose coefficients are independent random variables. When the noise is white, any orthonormal basis will satisfy the Karhunen-Loève theorem. In optical fiber communications systems, the ASE noise is effectively white when it is contributed by the amplifiers, but it only remains white for short distances over which the nonlinear interactions between the signal and the noise can be neglected. Over longer distances, the noise becomes correlated, and the Karhunen-Loève basis becomes unique and distance-dependent. The second mathematical result is Doob's theorem [46], which states that when the system is linearizable and driven by Gaussian-distributed noise, each of these independent random variables is Gaussian. Thus, it suffices in principle to determine the Karhunen-Loève modes, as well as the mean and variance of its coefficients, to calculate the effective noise pdf! I emphasize that this powerful result allows the signal to interact nonlinearly with itself and with the noise; it only requires that the noise not interact with itself. In practice, one must use an approximate static basis from which to compute the Karhunen-Loève modes. The choice of this basis set is important. The linearization assumption may hold in one basis set and not in another. As we will see, it is necessary to use a basis set in which the phase jitter, and in some systems also the timing jitter, are explicitly separated. Otherwise, the linearization assumption only holds for short distances.

The study of the dispersion-managed soliton (DMS) system that I introduce in the next chapter and in [47] shows that it is possible to use the linearization assumption to calculate the effects of accumulated noise. However, this assumption breaks down after a short propagation distance, unless the phase and timing jitter are handled separately from the other noise components, as I show in Section 4.3. This separation is necessary because the nonlinear equation that governs the fiber transmission (2.3.3) implies that small amounts of amplitude and frequency noise can lead to large amounts of phase and timing jitter respectively, which invalidates the linearization assumption. I have discussed the relationship between the parameter perturbations in Section 3.2 for a standard soliton, but the qualitative behavior remains the same for arbitrary pulse shapes. As stated previously, the nonlinear propagation equation is also phase and time invariant, which implies that phase and timing jitter can be separated from the standard Fourier basis without affecting the subsequent evolution. We will find that it is necessary to separate the phase and timing jitter separately for each pulse, in which case the phase invariance no longer strictly holds. However, it holds sufficiently well for the linearization to remain valid over the distances of interest. Once phase and in some cases timing jitter are separated, the coefficients of the modified Fourier basis, along with the phase and timing jitter, remain multivariate-Gaussian distributed far longer than the original Fourier coefficients [47]. I note that the phase is only Gaussian distributed if one tracks the phase change on the infinite line. If one only tracks it in the range $[0, 2\pi]$, then it is Jacobi-Θ function distributed. This function is the periodic analogue of a Gaussian distribution. Therefore, soliton perturbation theory is a special case where the expansion of optical noise into discrete modes and a continuum is appropriate, rather than an expansion in the usual Fourier basis [18]–[20], as shown in Section 3.2. I validate this assumption using extensive Monte Carlo simulations.

I note again that the approach assumes that noise-noise interactions in an appropriate basis set are negligible during the transmission, but I take them into account in the receiver, which I also assume has a realistic, narrow-band electrical Bessel filter. Thus, the work presented in this dissertation is a generalization of [48], in that the noise that enters the receiver is not white, but is determined by the actual transmission, and I apply realistic electrical filtering.

The optical noise in a wide range of optical transmission systems is multivariate-Gaussian distributed with zero mean, after the phase and timing jitter are separated. This pdf only depends on the covariance matrix of the noise. I present two ways to compute the covariance matrix:

1. First, it is possible to run Monte Carlo simulations and average the results to approximate the covariance matrix. In essence, one fits a multivariate-Gaussian distribution to the optical noise immediately prior to the receiver using a Monte Carlo simulation, after separating the phase and timing jitter. This procedure is self-validating and relatively simple to program, but requires on the order of 2000 noise realizations, as I will show in Chapter 5. However, this approach makes better use of the simulation data and yields more accurate results than traditional Monte Carlo simulations with the same number of noise realizations.

2. In Section 4.4, I derive an ordinary differential equation (ODE) that allows me to propagate the covariance matrix deterministically from amplifier to amplifier, a much more computationally efficient method than running Monte Carlo simulations. I refer to this approach as the covariance matrix method, and I use it in Chapter 6.

The entire discussion in this dissertation applies to one optical polarization only, which is appropriate for the example transmission systems that I am using [28]. Moreover, this choice somewhat simplifies the theoretical development. There is no reason to doubt that the covariance matrix method can be extended to take polarization effects into account.

The remainder of this chapter is organized as follows: I derive the theory of the linearization approach and define the covariance matrix in Section 4.2. The separation of phase and timing jitter and the computation of the covariance matrix from Monte Carlo simulations are described in Section 4.3. A deterministic computation based on the ODE is described in Section 4.4. Section 4.5 is devoted to saturated amplifiers in the context of noise linearization. In Section 4.6, I describe how the pdfs of the receiver current and the accurate BERs are computed. Finally, in Section 4.7, I consider the computation of the BER in systems with multiple bits.

4.2 Noise Covariance Matrix

The starting point is the system of equations for u_0 and δu as shown in the previous chapter

$$\begin{aligned} i\frac{\partial u_0}{\partial z}+\frac{D(z)}{2}\frac{\partial^2 u_0}{\partial t^2}+|u_0|^2 u_0 &= ig(z)u_0, & \text{(4.2.1a)} \\ i\frac{\partial \delta u}{\partial z}+\frac{D(z)}{2}\frac{\partial^2 \delta u}{\partial t^2}+2\left|u_0\right|^2\delta u+u_0^2(\delta u)^* &= ig(z)\,\delta u+\hat{F}, & \text{(4.2.1b)} \end{aligned}$$

where (4.2.1a) describes the evolution of the noise-free solution u_0 and (4.2.1b) describes the accumulated noise if $\delta u \ll u_0$. The signal u_0 must be known when solving (4.2.1b) and in practice it is most convenient to solve both equations in parallel. Note that I neglect any influence of δu on u_0. Considering all the terms in the Kerr nonlinearity of the NLS with $u = u_0 + \delta u$,

$$\left|u\right|^2 u = \left|u_0\right|^2 u_0^* + 2\left|u_0\right|^2 \delta u + u_0^2\,\delta u^* + u_0^*\,(\delta u)^2 + 2u_0\left|\delta u\right|^2 + \left|\delta u\right|^2 \delta u, \qquad (4.2.2)$$

one finds the second and third terms on the right-hand side of (4.2.2) that appear in (4.2.1b). The fourth and fifth terms on the right-hand side are linear in u_0 and represent the next order beyond noise linearization. The term $\left|\delta u\right|^2 \delta u$ is cubic in δu and can be expected to contribute least. The term $u_0^*\,(\delta u)^2$ describes the depletion of u_0 due to the four-wave mixing with the noise, and the term $2u_0\left|\delta u\right|^2$ describes a cross-phase modulation of the signal due to the noise. These two terms can in principle be included into (4.2.1a), which is the statistical average of (2.3.3) for $u = u_0 - \delta u$, yielding

$$i\frac{\partial u_0}{\partial z}+\frac{D(z)}{2}\frac{\partial^2 u_0}{\partial t^2}+|u_0|^2 u_0 = ig(z)u_0 - 2u_0\left\langle\left|\delta u\right|^2\right\rangle - u_0^*\left\langle(\delta u)^2\right\rangle. \qquad (4.2.3)$$

The averages $\langle|\delta u|^2\rangle$ and $\left\langle(\delta u)^2\right\rangle$ could be obtained by solving 4.2.1b, in which case the equations (4.2.3) and (4.2.1b) would form a mutually dependent equation system. However, if second-order terms become relevant, the linearization assumption is invalid and optical noise will not be Gaussian distributed anymore. I neglect the influence of δu on u_0 by solving the system (4.2.1a), (4.2.1b), and I show that the linearization assumption holds in an appropriate basis set for highly nonlinear optical fiber transmission systems.

One can expand $u_0(t)$ and $\delta u(t)$ as Fourier series,[1]

$$u_0(t) = \sum_{n=-N_{\mathrm{FFT}}/2}^{N_{\mathrm{FFT}}/2-1} A_n(z)\,\exp(i\omega_n t), \tag{4.2.4a}$$

$$\delta u(t) = \sum_{n=-N_{\mathrm{FFT}}/2}^{N_{\mathrm{FFT}}/2-1} a_n(z)\,\exp(i\omega_n t), \tag{4.2.4b}$$

where T is the period, N_{FFT} is the number of sample points in the Fourier transform, and $\omega_n \equiv 2\pi n T_0/T$. Typical values of N_{FFT} are powers of 2 such as 1024, 2048, and 4096. The values of T and N_{FFT} are chosen so that the simulation result does not change when either one is increased. After substituting (4.2.4a) and (4.2.4b) into (4.2.1b), one finds that

$$\begin{aligned}\frac{da_k}{dz} &= \Big(g - i\frac{D}{2}\omega_k^2\Big)a_k \\ &\quad + i\sum_{n,l,m=-N_{\mathrm{FFT}}/2}^{N_{\mathrm{FFT}}/2-1}\Big[2\,A_n\,A_l^*\,a_m\,\delta_{n-l,k-m} + A_n\,A_l\,a_m^*\,\delta_{n+l,k+m}\Big] - i\Gamma_k(z),\end{aligned} \tag{4.2.5}$$

where the Γ_k are the Fourier coefficients of the white noise input $\hat{F}$, and δ is Kronecker's delta. Using (2.3.7), the correlation of the Γ_k is $\langle\Gamma_k(z)\,\Gamma_m^*(z')\rangle = (2\eta T_0/T)\,\delta(z-z')\,\delta_{k,m}$, where $\eta = n_{\mathrm{sp}} g_m \hbar\omega_0 L_D \gamma/T_0$ is again zero outside of the amplifiers.

I now introduce a frequency cutoff and will only consider $N = 2Tf_{\max}$ frequencies with $|f| < f_{\max} = (N/2)\Delta f_{\mathrm{FFT}} = N/(2T)$ in the following. I choose the frequency range $[-f_{\max}, f_{\max}]$ so that the signal power outside of this range is smaller than 1% of its total value. Hence, I neglect frequencies from the linearization at which the signal power is small, as justified by the discussion in Section 3.2. Typical values of N lie in the range 50–150. The large ratio of $N_{\mathrm{FFT}}/N \gtrsim 10$ demonstrates the advantage of working in the frequency domain: The relevant modes are all concentrated at low frequencies, and high frequency modes can be neglected. The covariance matrix can be expressed in any basis, for example in the time domain, but then reducing the number of modes might not be as simple.

I define the complex column vectors $\boldsymbol{\alpha} = (a_{-N/2}, \dots, a_{N/2-1})^T$ and $\boldsymbol{\alpha}^* = (a_{-N/2}^*, \dots, a_{N/2-1}^*)^T$, as well as $\Gamma = (\Gamma_{-N/2}, \dots, \Gamma_{N/2-1})^T$, where the superscript T

---

[1] Note that by choosing this sign convention in the exponential functions, the optical frequency that corresponds to ω_n lies at $\omega_{\mathrm{total}} = \omega_0 - \omega_n$, where ω_0 is the carrier frequency.

indicates the transpose operation. Then I can rewrite (4.2.5) in matrix form as

$$\frac{d\boldsymbol{\alpha}}{dz} \;=\; \mathsf{B}\boldsymbol{\alpha}+\mathsf{E}\boldsymbol{\alpha}^{*}-i\Gamma, \tag{4.2.6}$$

where the complex matrices B and E are defined as[2]

$$\mathsf{B}_{km} \;=\; \Big(g-i\frac{D}{2}\omega_k^2\Big)\delta_{km}+2i\sum_{n,l}A_n\,A_l^{*}\,\delta_{n-l,k-m}, \tag{4.2.7a}$$

$$\mathsf{E}_{km} \;=\; i\sum_{n,l}A_n\,A_l\,\delta_{n+l,k+m}. \tag{4.2.7b}$$

The matrix E is symmetric ($\mathsf{E}_{km}=\mathsf{E}_{mk}$), and, if g is zero, the matrix B is anti-self-adjoint ($\mathsf{B}_{km}^{*}=-\mathsf{B}_{mk}$). The matrix B is circulant and thus corresponds to a convolution in the time domain, while E can be termed anti-circulant. (A matrix M is circulant if there is a vector $\boldsymbol{x}$ with $M_{km}=x_{k-m}$). The number of operations required to evaluate B and E grows like N^3. Equation (4.2.5) depends on both a_k and a_k^{*} since the linearized noise propagation is not self-adjoint [12], [13].

The probability space of the optical noise in the frequency domain is spanned by the $2N$ real variables $a_{k,R}$ and $a_{k,I}$. It is therefore convenient to split (4.2.5) into its real and imaginary parts and consider the resulting system of equations. I define the real partitioned vector $\boldsymbol{a}=(\boldsymbol{\alpha}_R,\boldsymbol{\alpha}_I)=(a_{-N/2,R},a_{-N/2+1,R}\cdots,a_{N/2-1,R},$ $a_{-N/2,I},a_{-N/2+1,I}\cdots,a_{N/2-1,I})^T$ of length $2N$, consisting of the real and imaginary parts of $\boldsymbol{\alpha}$ at the N lowest frequencies $-N/2,-N/2+1,\ldots N/2-1$. Similarly, I define $\boldsymbol{w}=(\Gamma_I,-\Gamma_R)^T$. I can express (4.2.6) as

$$\frac{d\boldsymbol{a}}{dz} \;=\; \mathcal{R}(z)\boldsymbol{a}+\boldsymbol{w}(z), \tag{4.2.8a}$$

$$\mathcal{R} \;=\; \begin{bmatrix} \mathsf{B}_R+\mathsf{E}_R & -\mathsf{B}_I+\mathsf{E}_I \\ \mathsf{B}_I+\mathsf{E}_I & \mathsf{B}_R-\mathsf{E}_R \end{bmatrix}, \tag{4.2.8b}$$

where $\mathcal{R}$ is a real $2N\times 2N$ block matrix, and I have used the notation $\mathsf{B}=\mathsf{B}_R+i\mathsf{B}_I$ and $\mathsf{E}=\mathsf{E}_R+i\mathsf{E}_I$. One may formally write the solution to (4.2.8b) as [49]

$$\boldsymbol{a}(z) \;=\; \boldsymbol{\Psi}(z,z_0)\boldsymbol{a}(z_0)+\int_{z_0}^{z}\boldsymbol{\Psi}(z,z')\boldsymbol{w}(z')\,dz', \tag{4.2.9}$$

---

[2] In the rest of this dissertation, I will use a sans serif font to denote complex $N\times N$ matrices like B, a script font for real $2N\times 2N$ matrices like $\mathcal{R}$, and a bold face font for real $2N$-vectors such as $\boldsymbol{a}$. The only exception will be $\boldsymbol{\alpha}$, which is a complex N-vector.

where $\boldsymbol{\Psi}(z,\zeta)$ is a propagator matrix that obeys the equation

$$\frac{d}{dz}\boldsymbol{\Psi}(z,\zeta)=\mathcal{R}(z)\boldsymbol{\Psi}(z,\zeta)\,,\qquad \boldsymbol{\Psi}(\zeta,\zeta)=\mathcal{I}\quad\forall\, z,\zeta\,, \tag{4.2.10}$$

and $\mathcal{I}$ is the identity matrix. Equation (4.2.8b) describes the spatial evolution of the noise Fourier modes. I now assume that the $\boldsymbol{a}(z)$ satisfy a multivariate-Gaussian distribution, which is completely described by its first two moments. Making this assumption is equivalent to assuming that the linearization assumption holds in a Fourier basis set. For the system that I consider in this dissertation, this assumption only remains valid for short distances. However, it is a useful starting point for our subsequent discussion. From (2.3.6), it follows that the mean of $\boldsymbol{a}(z)$ is zero, while the second moments are given by the covariance matrix $\mathcal{K}(z)$. The pdf $p_{\boldsymbol{a}}$ of $\boldsymbol{a}$ may be written as [50]

$$p_{\boldsymbol{a}}(\boldsymbol{a},z)=(2\pi)^{-N}\sqrt{\det\mathcal{K}^{-1}(z)}\,\exp\Big[-\tfrac{1}{2}\,\boldsymbol{a}^T\mathcal{K}^{-1}(z)\boldsymbol{a}\Big], \tag{4.2.11}$$

where the real symmetric $2N\times 2N$ covariance matrix $\mathcal{K}$ is defined by,

$$\begin{aligned}\mathcal{K} &= \Big\langle\big(\boldsymbol{a}-\langle\boldsymbol{a}\rangle\big)\big(\boldsymbol{a}^T-\langle\boldsymbol{a}\rangle^T\big)\Big\rangle\\ &= \left\langle\begin{bmatrix}\boldsymbol{\alpha}_R\boldsymbol{\alpha}_R^T & \boldsymbol{\alpha}_R\boldsymbol{\alpha}_I^T\\ \boldsymbol{\alpha}_I\boldsymbol{\alpha}_R^T & \boldsymbol{\alpha}_I\boldsymbol{\alpha}_I^T\end{bmatrix}\right\rangle\equiv\begin{bmatrix}\text{RR} & \text{RI}\\ \text{IR} & \text{II}\end{bmatrix},\end{aligned} \tag{4.2.12}$$

where RR, RI, IR and II are four real $N\times N$ block matrices. In (4.2.12) all vector products above are outer products. This definition of $\mathcal{K}$ embodies the full covariance information in $4N^2$ real numbers of which $N(2N+1)$ are independent. The alternative complex $N\times N$ matrix $\langle a_k a_l^*\rangle=\langle\boldsymbol{\alpha}\boldsymbol{\alpha}^\dagger\rangle_{kl}$, where the † denotes the conjugate transpose, contains only $2N^2$ real numbers of which N^2 are independent, and thus $\langle a_k a_l^*\rangle$ does not contain complete information. From (4.2.8b) and (4.2.9), one now finds that $\mathcal{K}$ evolves over distance according to

$$\frac{d}{dz}\mathcal{K}=\mathcal{R}\mathcal{K}+\mathcal{K}\mathcal{R}^T+\frac{\eta T_0}{T}\mathcal{I}, \tag{4.2.13}$$

where η is defined after (2.3.7). Equation (4.2.13) is a Lyapunov equation [49] and is the linear evolution equation for the covariance matrix. The right-hand side of (4.2.13) is symmetric since $(\mathcal{R}\mathcal{K})^T=\mathcal{K}^T\mathcal{R}^T=\mathcal{K}\mathcal{R}^T$, so that $\mathcal{K}$ remains symmetric as it evolves over z. Initially, $\mathcal{K}$ is zero since the launched signal is assumed to be noise-free. The matrix $\mathcal{R}$ is distance-dependent and includes amplification/attenuation as well as the nonlinear interactions of the signal with the noise. The last term describes the white noise input and is only nonzero inside the optical amplifiers. Newly-added noise only contributes to the diagonal elements $\mathcal{K}_{kk}$. In addition to being symmetric, $\mathcal{K}$ is also positive definite, so that its determinant is positive.

I note that the direct derivation of (4.2.13) from (4.2.8b) is only one possible way of obtaining (4.2.13). One can show that the pdf in (4.2.11), where the covariance matrix $\mathcal{K}(z)$ is described by (4.2.13), represents the exact solution of the Fokker-Planck equation [51] corresponding to the Langevin equation (4.2.8b). Yet another approach to derive (4.2.13) uses Itô's method [44]. All of these methods are, of course, equivalent.

4.3 Separation of Phase and Timing Jitter

Soliton perturbation theory is based on the fact that standard solitons remain constant as a function of the propagation distance, except for a constant phase rotation. Moreover, solitons are uniquely characterized by the four parameters A, τ, Φ, and Ω as defined in (2.4.2). Kodama [52] proved that a realistic fiber system with gain, loss, and variable dispersion does not generally support constant or strictly periodic pulse shapes; instead, pulses decay over extremely long distances. One can nevertheless generalize some results of soliton perturbation theory.

As shown by equations (3.2.8b) and (3.2.8d), the part of the noise that is responsible for a phase perturbation of a soliton is to first order proportional to $f_\Phi = iu_0$, while the part of the noise that shifts its central time is proportional to $f_\tau = A\tanh[A(t-\tau)]\,u_s = (du_s/dt)|_{\Omega=0}$. Although these results were derived for a classical soliton, they also hold for any other pulse shape: The expansion $u_0\exp(i\varphi) = u_0[1+i\varphi+O(\varphi^2)]$ shows that the phase of u_0 is rotated by φ when adding $i\varphi u_0$. Analogously, consider $\partial u_0(t)/\partial t \approx [u_0(t+\tau)-u_0(t)]/\tau$ for small τ, implying $u_0(t)+\tau u_0'(t) = u_0(t+\tau)$. It is possible to integrate arbitrarily large perturbations: Consider the initial value problem

$$\frac{\partial u}{\partial z} = i\varphi u + q\frac{\partial u}{\partial t}, \tag{4.3.1a}$$

$$u(t,z=0) = u_0(t), \tag{4.3.1b}$$

which is exactly solved by $u(t,z) = u_0(t+zq)\exp(i\varphi z)$. Equation (4.3.1a) describes the optical field in an idealized transmission line without nonlinearity, dispersion, or loss, driven by noise that produces phase and timing jitter. Moreover, the functions

$$\begin{aligned} f_\varphi(t) &\equiv iu_0(t), && \text{(4.3.2a)} \\ f_\tau(t) &\equiv \frac{\partial u_0(t)}{\partial t}. && \text{(4.3.2b)} \end{aligned}$$

are solutions of the linearized NLS (4.2.1b), showing that both perturbations are stable to first order when propagating with an arbitrary signal u_0.

Based on these observations, I define the *pulse energy* U, the *average phase* φ, the *central time* τ, and the *central frequency* Ω of a pulse $u(t,z)$ that is confined to the bit slot $[t_0, t_0+T_{\text{bit}}]$ as

$$\begin{aligned} U &= \int_{t_0}^{t_0+T_{\text{bit}}} |u|^2 dt && \text{(4.3.3a)} \\ \tau &= \frac{1}{U}\int_{t_0}^{t_0+T_{\text{bit}}} t\,|u|^2 dt, && \text{(4.3.3b)} \\ \varphi &= \frac{1}{U}\int_{t_0}^{t_0+T_{\text{bit}}} \arg(u)\,|u|^2 dt, && \text{(4.3.3c)} \\ \Omega &= \frac{1}{U}\text{Im}\int_{t_0}^{t_0+T_{\text{bit}}} u_t^* u\,dt, && \text{(4.3.3d)} \end{aligned}$$

respectively, where $\arg(x) = \arctan[\text{Im}(x)/\text{Re}(x)]$. In the case of a soliton, the parameter set $U/2$, τ, φ, Ω agrees with the soliton parameters set A, τ, Φ, Ω. Moreover, I argue that the set $U/2$, τ, φ, Ω is the appropriate perturbation expansion parameter set for arbitrary pulse shapes because perturbations in $U/2$, τ, φ, Ω are independent of each other to first order at the point in the fiber where the perturbation occurs. The pulse energy can be varied without influencing the other parameters by scaling $u(t)$, the central time can be shifted in isolation, and so forth.

In (3.2.16c) and (3.2.16d), I showed that phase and timing jitter grow proportionally to $z^{3/2}$. In order for the linearization assumption to remain valid over large propagation distances, one must separate the phase and timing jitter from the covariance matrix, while keeping track of their magnitude. If these contributions are not separated, they distort the distribution functions so that the noise $\boldsymbol{a}$ is no longer Gaussian distributed. Physically, (3.2.16c) and (3.2.16d) imply that small amounts of amplitude and frequency noise can lead to large amounts of phase and timing jitter respectively, which can invalidate the linearization. The nonlinear propagation equation is phase and time invariant, which implies that phase and timing jitter can be removed from the accumulated noise surrounding a

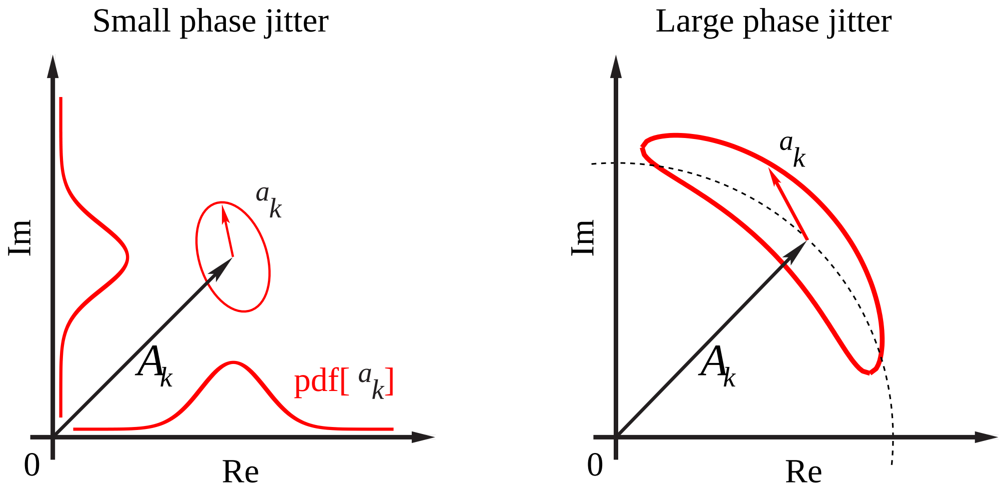

Figure 4.1: Distortion of the marginal noise distributions due to phase jitter. The vectors are shown in the complex plane, A_k: noise-free vector, a_k: accumulated noise vector at frequency component k.

pulse without affecting the pulse's subsequent evolution. While this statement is strictly valid only for isolated pulses, not for systems with multiple pulses and/or multiple channels, we have found it works well enough in practice for the linearization assumption to hold, once phase and timing jitter are properly separated from the individual pulses.

Fig. 4.1 shows the distortion of the marginal distributions due to phase jitter. The impact of timing jitter on a system is similar in concept but not as easy to depict. In the left diagram, I show the vectors of one Fourier coefficient A_k of the noise-free signal u_0 in the complex plane and one of the Fourier coefficients a_k of the accumulated noise δu. The contour of constant noise probability density around A_k in the presence of weak phase jitter is in general an ellipse and the projection of the pdf onto the axes yields the marginal pdfs. If the accumulated noise δu is multivariate-Gaussian distributed, then all the marginals must be one-dimensional Gaussian pdfs [45]. However, if the phase jitter becomes large and the linearization breaks down, the jitter tends to rotate the signal around the origin. The contour of constant probability density spreads out along a circle around the origin, and it is slightly skewed since noise realizations with large phase deviations are usually caused by large amplitude deviations, resulting in a tilted, banana-shaped contour as shown in the right diagram. The projections of this contour are not Gaussian. A similar discussion applies to timing jitter.

In the dispersion-managed soliton system that I consider in the next chapter, the timing jitter grows only linearly with distance, rather than proportionally to $z^{3/2}$, due to inline filtering in the loop [23]. Nevertheless, timing jitter can become on the order of the pulse duration once it reaches the receiver so that it is not a small perturbation anymore. Phase jitter is mainly driven by amplitude jitter, as shown in (3.2.16c), so that it is not reduced by the inline filter. I find that phase jitter has to be separated in the DMS system that I introduce in the next chapter, as well as in the much less nonlinear CRZ system that I will introduce in Section 6.2. Timing jitter only has to be separated in the DMS system.

In the following, I show how to separate large phase and timing jitter from the calculation of the covariance matrix by using Monte Carlo simulations. I will deal with the deterministic covariance matrix method and the incremental separation of small phase and timing jitter in the next section.

In Monte Carlo simulations, one focuses on the received optical signal $u(t) = u_0(t) + \delta u(t)$, where $u_0 = \langle u \rangle$ is the signal averaged over all noise realizations, and δu is one particular realization of the accumulated noise at the receiver. The Fourier expansion of u is $u(t) = \sum_{k=-N/2}^{N/2-1} B_k \exp(i\omega_k t)$, where $B_k = A_k + a_k$ and $\omega_k = 2\pi k T_0/T$, conforming to (4.2.4a) and (4.2.4b). The phase and timing offsets φ and τ cannot be considered small anymore at the receiver, corresponding to the situation depicted in Fig. 4.1(b). In fact, the jitter can become so strong that the marginal pdfs become Gaussians around zero and then the averages $A_k = \langle B_k \rangle$ tend to zero!

The approach that my colleagues and I use is to apply the nonlinear transformation from B_k to $\overline{B}_k$

$$\overline{B}_k \equiv B_k \exp[-i(\varphi + \omega_k \tau)] = \overline{A}_k + r_k, \tag{4.3.4}$$

where $\overline{A}_k \equiv \langle \overline{B}_k \rangle$ is the new signal average and $\langle r_k \rangle = 0$ is a residual noise. For single pulse transmission and for each noise realization, we determine φ and τ by fitting the linear function $\alpha + \beta\omega_k$ to the phase of the B_k using the least-squares criterion

$$H = \min_{\alpha,\beta} \sum_{k=-N/2}^{N/2-1} |B_k|^2 \left[\arg B_k - (\alpha + \beta\omega_k)\right]^2, \tag{4.3.5}$$

where $B_k = B_{k,R} + i\,B_{k,I}$, and then setting $\varphi = \alpha$ and $\tau = \beta$. We have found that the linear phase assumption of (4.3.5) in the DMS system is good as long as the receiver is placed at the chirp-free maximum pulse compression point of the dispersion map. This renormalization separates the large phase and timing fluctuations from the total signal and hence the new average power $|\overline{A}_k|^2 \gg |A_k|^2$ is larger than without the separation.

Conversely, the average power of the residual noise $\langle\sum_k |r_k|^2\rangle \ll \langle\sum_k |a_k|^2\rangle$ is reduced. I show in Appendix A.1 that (4.3.5) and (4.3.3c) are consistent for arbitrary pulse shapes.

One can compute the pdfs of φ and τ by averaging over all Monte Carlo noise realizations. Moreover, with the real partitioned $2N$-vector $\boldsymbol{r} = (r_{-N/2,R}, \cdots, r_{N/2-1,R}, r_{-N/2,I}, \ldots, r_{N/2-1,I})^T$ and, analogously, $\overline{\boldsymbol{B}} = (\overline{B}_{-N/2,R}, \ldots)^T$, one may define a reduced covariance matrix $\mathcal{K}^{(r)}$as

$$\mathcal{K}^{(r)} = \left\langle \overline{\boldsymbol{B}}\overline{\boldsymbol{B}}^T \right\rangle - \left\langle \overline{\boldsymbol{B}} \right\rangle \left\langle \overline{\boldsymbol{B}} \right\rangle^T = \left\langle \boldsymbol{r}\boldsymbol{r}^T \right\rangle. \tag{4.3.6}$$

I will use (4.3.6) in the next chapter. The quantity $\boldsymbol{r}$ obeys a multivariate-Gaussian distribution, and one may replace $\boldsymbol{a}$ by $\boldsymbol{r}$ and $\mathcal{K}$ by $\mathcal{K}^{(r)}$ in (4.2.11).

Fig. 4.2 shows simulation results of the dispersion-managed soliton (DMS) system that I will treat in the next chapter. I transmit a single soliton with a peak power of 7.9 mW and a FWHM duration of 9 ps. The dots show 750 Monte Carlo noise realizations of the 10 GHz Fourier mode B_8 for $n_{\text{sp}} = 1.2$ in the complex plane. Fig. 4.2(a)–(c) show B_8 without jitter separation, while Fig. 4.2(d)–(f) show B_8 after employing phase and timing jitter separation. The solid curve is a probability density contour under the assumption that the magnitude and the angles of the dots are independently Gaussian distributed. The radius of the dash-dotted circle equals $\langle|B_8|\rangle \approx |A_8|$, the average magnitude of the dots. Next, I consider the set of the angles $\{\arg(B_8)\}$. The quantity r is the ratio of the standard deviation of these angles over the standard deviation of the magnitude of the samples. When r is large, the contour assumes a banana-like or even doughnut-like shape and wraps around the unit circle. The upper row of figures shows that r grows quickly with the transmission distance, indicating strong phase jitter. However, r can be kept close to unity if we separate the phase and timing jitter. I note also that if n_{sp} or the degree of fiber nonlinearity is reduced, the value of r converges toward unity.

Fig. 4.2(g)–(i) show $Ue^{i\varphi}$ for the same 750 noise realizations, using (4.3.3a) and (4.3.3c). The quantity r here is the ratio of the standard deviation of φ to that of U. Again, the solid sickle-shaped contours are drawn at a constant probability density. The contours are visibly skewed, indicating a coupling between U and φ that is analogous to a similar coupling that occurs in soliton perturbation theory, expressed in (3.2.11b) and (3.2.16c). This skew is not visible in the upper six subfigures, since the B_k only represent a small fraction of the signal energy, and hence their individual fluctuations overwhelm the skew.

After the jitter separation, the coefficients of the Fourier basis remain multivariate-Gaussian distributed far longer than the original Fourier coefficients. In principle, one

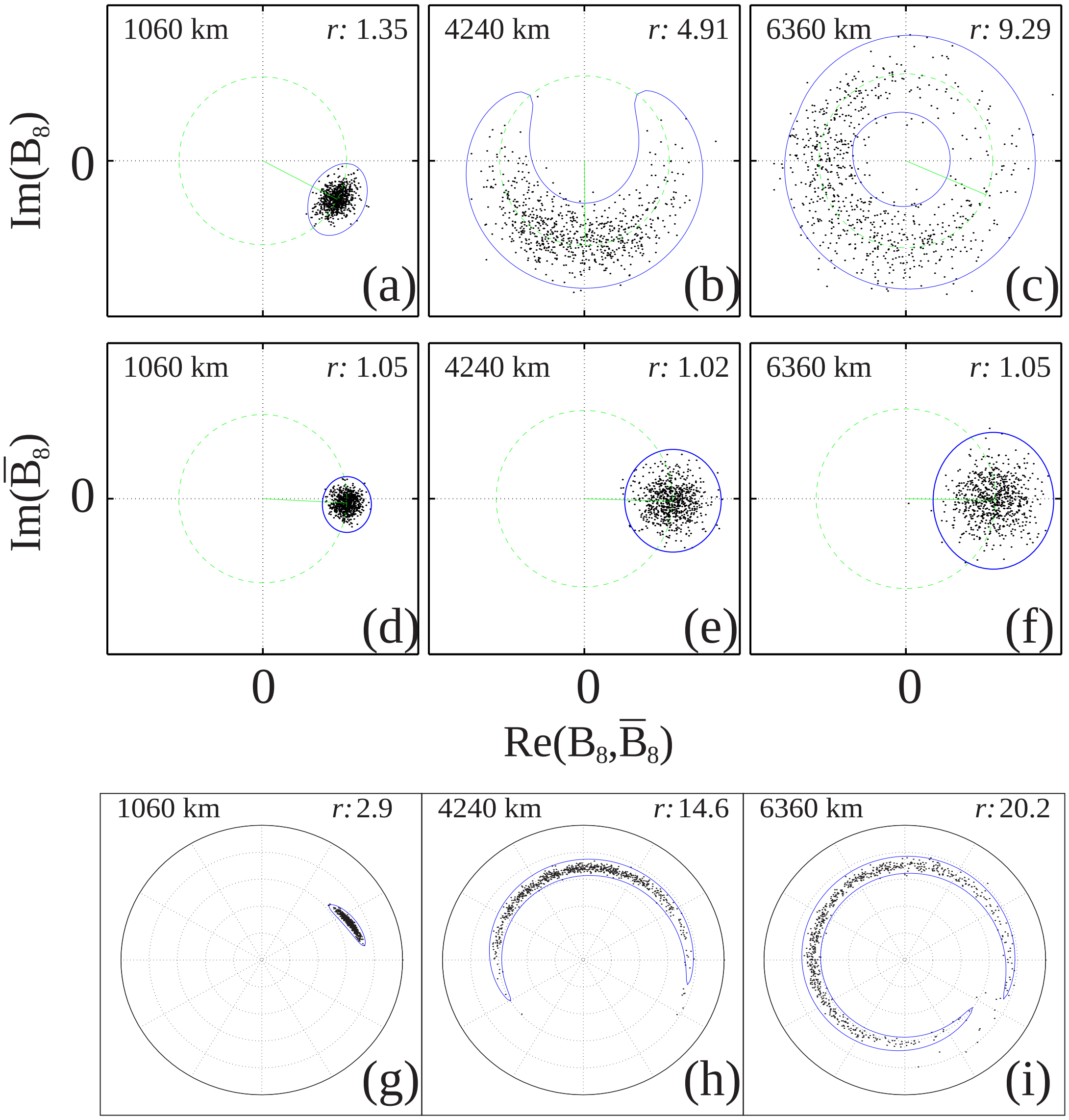

Figure 4.2: (a) Simulation results in the DMS system introduced in the next chapter with ASE noise. The dots show 750 Monte Carlo realizations of the 10 GHz noisy Fourier mode B_8 for $n_{\text{sp}} = 1.2$. (a)–(c): B_8 without jitter separation, (d)–(f): B_8 with jitter separation. (g)–(h): $Ue^{i\varphi}$ for the same 750 noise realizations, using (4.3.3a) and (4.3.3c).

should take into account the off-diagonal matrix elements of the covariance matrix that result from the interactions of the modified Fourier coefficients $\overline{B}_k$ and the phase and timing jitter. However, these elements can be ignored for reasons that I will explain later. If one assumes that soliton perturbation theory can be approximately generalized to arbitrary pulse shapes, then the noise in the Fourier basis after the phase and timing jitter separation is a representation of the energy jitter δU and central frequency jitter $\delta\boldsymbol{\Omega}$, plus the noise continuum. Energy and frequency jitter can be expected to grow much slower than phase and timing jitter due to (3.2.16a)–(3.2.16d), and the simulations presented here show that it is not necessary to separate them from the continuum.

4.4 Deterministic calculation of the covariance matrix

In this section, we obtain the covariance matrix by solving the linearized evolution equation directly, without the use of Monte Carlo simulations. The basic approach that I follow is to propagate the covariance matrix from amplifier to amplifier, while projecting out and separating the contribution to the phase and timing jitter, before the fluctuations have had the opportunity to accumulate significantly.

The propagation of the accumulated noise is governed by (4.2.1b), which is a linear equation and is homogeneous everywhere except at the amplifiers since $\hat{F} = 0$ in the fiber. Its Fourier transform must be linear and homogeneous as well, and I can write it in terms of $\boldsymbol{a}$ as in (4.2.8b),

$$\frac{d\boldsymbol{a}}{dz} = \mathcal{R}(z)\boldsymbol{a}, \tag{4.4.1}$$

where I set the ASE term $\boldsymbol{w}(z)$ to zero. I write the solution of (4.4.1) as $\boldsymbol{a}(z) = \boldsymbol{\Psi}(z)\boldsymbol{a}(0)$, consistent with (4.2.9). The evolution of the noise covariance matrix $\mathcal{K} = \left\langle \boldsymbol{a}\boldsymbol{a}^T \right\rangle$ over one fiber leg from $z = 0$ to $z = L$, in which no noise is added, followed by an EDFA with the power gain G, is given by

$$\mathcal{K}(L) = G\boldsymbol{\Psi}\mathcal{K}(0)\boldsymbol{\Psi}^T + \eta\frac{T_0}{T}\mathcal{I}, \tag{4.4.2}$$

where $\boldsymbol{\Psi}$ is a propagator matrix, $\mathcal{I}$ is the identity matrix, and η is defined in (2.3.8). By successive application of (4.4.2), one can propagate the covariance matrix from amplifier to amplifier.

We choose a perturbative method to compute $\boldsymbol{\Psi}$. Let $u_0(t,0)$ and $u_0(t,L)$ be the noise-free optical field at the beginning and end of the fiber span respectively. Then we perturb

$u_0(t,0)$ in a single frequency mode m by a small amount ϵ and launch the perturbed signal

$$u^{(m)}(t,0) = u_0(t,0) + \epsilon \exp(i\omega_m t). \tag{4.4.3}$$

At $z = L$, we obtain $u^{(m)}(t,L)$ by solving the nonlinear transmission equation (4.2.1a) and calculate the deviation $\delta u^{(m)}(t) = u^{(m)}(t,L) - u_0(t,L)$ and its Fourier space vector $\boldsymbol{a}^{(m)}$. The elements of $\boldsymbol{\Psi}$ are given by

$$\boldsymbol{\Psi}_{km} = \frac{a_k^{(m)}}{\epsilon}. \tag{4.4.4}$$

This perturbative approach corresponds to the Lyapunov method described by Bennetin *et al.* [53]. Note that ϵ must be a positive real number if $1 \leq k \leq N$ and a purely imaginary positive number if $N+1 \leq k \leq 2N$, since $\boldsymbol{\Psi}$ is defined in the $2N$-dimensional space of partitioned vectors. The real and imaginary perturbations must be computed separately since the operator $\mathcal{L}$ defined in (3.2.3) is not linear when δu is scaled by an imaginary number: $\mathcal{L}[i\delta u] \neq i\mathcal{L}[\delta u]$, due to the non-self-adjoint FWM term $u_0^2(\delta u)^*$.

When carrying out Monte Carlo simulations as described in the previous section, the phase and timing jitter are large by the time they reach the receiver. By contrast, when propagating $\mathcal{K}$, one can separate the jitter incrementally, on a scale that is small compared to the nonlinear scale length of the system. We do that at every amplifier. Following the discussion of the previous section, I consider the functions $f_\varphi(t) = iu_0(t)$ and $f_\tau(t) = \partial u_0(t)/\partial t$, defined in (4.3.2a) and (4.3.2b). Each pulse in the signal has a different average phase φ (4.3.3c) and central time τ (4.3.3b). In a system in which pulses do not overlap, such as a soliton system, φ and τ evolve independently for each pulse and must be removed separately from each other. To deal with non-overlapping pulses, we decompose the functions $f_\varphi(t)$, $f_\tau(t)$, and $\delta u^{(m)}(t)$ into sums of mutually orthogonal pulse functions, *i.e.* $h(t) = \sum_l h_l(t)$, where $h_l(t) = h(t)$ for $(l-1)T_{\text{bit}} \leq t < lT_{\text{bit}}$, and $h_l(t) = 0$ otherwise for an arbitrary function h. I later discuss how to generalize this decomposition to the case where pulses overlap during the transmission.

The next step is to orthogonalize the functional basis $[f_{\varphi,l}(t), f_{\tau,l}(t), \delta u_l^{(m)}(t)]$ analogously to [23]. The goal is to obtain a new set of basis vectors $[f_{\varphi,l}(t), \widetilde{f}_{\tau,l}(t), \widetilde{\delta u_l}^{(m)}(t)]$ so that $(f_{\varphi,l}, \widetilde{f}_{\tau,l}) = (\widetilde{f}_{\tau,l}, \widetilde{\delta u_l}^{(m)}) = (f_{\varphi,l}, \widetilde{\delta u_l}^{(m)}) = 0$ with the real scalar product (3.2.4).

We apply the Gram-Schmidt orthogonalization procedure [54]

$$\widetilde{f}_{\tau,l} = f_{\tau,l} - \frac{(f_{\tau,l}, f_{\varphi,l})}{(f_{\varphi,l}, f_{\varphi,l})} f_{\varphi,l}, \tag{4.4.5a}$$

$$\widetilde{\delta u}_l^{(m)} = \delta u_l^{(m)} - \frac{(\delta u_l^{(m)}, \widetilde{f}_{\tau,l})}{(\widetilde{f}_{\tau,l}, \widetilde{f}_{\tau,l})} \widetilde{f}_{\tau,l} - \frac{(\delta u_l^{(m)}, f_{\varphi,l})}{(f_{\varphi,l}, f_{\varphi,l})} f_{\varphi,l}. \tag{4.4.5b}$$

The Fourier decomposition vector $\boldsymbol{r}^{(m)}$ of the reduced noise $\widetilde{\delta u}^{(m)} = \sum_l \widetilde{\delta u}_l^{(m)}$ then replaces $\boldsymbol{a}^{(m)}$ in (4.4.4), and the covariance matrix $\mathcal{K}$ in (4.4.2) is reduced to $\mathcal{K}^{(r)}$ analogously to (4.3.6). Note that the functions $f_{\varphi,l}$ and $f_{\tau,l}$ are not orthogonal for asymmetric pulses; therefore, (4.4.5a) is not redundant in general. Even though $f_{\varphi,l}(t)$ and $\widetilde{\delta u}_l^{(m)}$ are orthogonal, the quantities φ and $\mathcal{K}^{(r)}$ are not statistically independent because φ depends on the pulse power due to the nonlinear phase rotation. A noise realization in which the noise increases the pulse power will tend to have both large $|\widetilde{\delta u}^{(m)}|$ and large $|\varphi|$, leading to a correlation. However, these correlations, like the phase jitter itself, have no effect on a receiver with a square law detector and can be neglected. There is a similar correlation between τ and $|\widetilde{\delta u}^{(m)}|$, and, in contrast to the phase jitter, the timing jitter cannot simply be ignored. However, I will show in Section 6.1 that the correlations between τ and the modified Fourier components has a negligible effect on the receiver current, so that these correlations do not influence the pdf of the receiver current.

I briefly return to the problem of overlapping pulses. In most modern transmission systems, optical pulses overlap during the transmission even though they are well-separated when they are launched and detected. In the CRZ system introduced in Chapter 6.2, the maximum FWHM pulse duration is 210 ps, leading to a significant overlap of adjacent pulses. In this case, the phase jitter can still be removed separately for each pulse after applying artificial dispersion compensation. One passes the functions $\{f_{\varphi,l}(t), f_{\tau,l}(t), \delta u_l^{(m)}(t)\}$ through an ideal linear and lossless fiber whose total dispersion is $-D(L)$, where $D(L)$ is the total accumulated dispersion at the point L in the transmission system. This procedure separates the pulses, Since it is linear, it is fully reversible [55]. Then, one applies the orthogonalization (4.4.5a) and (4.4.5b) to the separated pulses and sends the signal corresponding to $\widetilde{\delta u}_l^{(m)}$ back to the point L through an ideal fiber with total dispersion $+D(L)$. One might argue that the phases of the overlapped pulses will not evolve independently and hence might become correlated; however, we find by comparison to Monte Carlo simulations that the procedure just described leads to accurate BERs in the CRZ system, indicating that the coupling of the signal in one pulse to the phase jitter in an overlapping pulse is negligible.

I will summarize the algorithm for propagating $\mathcal{K}$ in Appendix A.4. Propagating the covariance matrix requires the propagation of u_0, plus the $2N$ perturbed fields. The total computational cost of this algorithm is thus roughly $2N+1$ times that of a single Monte Carlo noise realization, plus the time required for the matrix multiplications in (4.4.2). Since matrix multiplications scale with the cube of the dimension N^3, there is a practical limit to the size of N. I will demonstrate in Chapter 6.2 that the solution of a problem with $N = 140$ is computationally feasible on a Pentium P4 workstation.

I have also attempted a direct solution of (4.2.13) using the matrix ODE solver package CVODE [56], but the solution was numerically inefficient. I find that the perturbation method is numerically stable and its result is independent of the value of ϵ over several orders of magnitude.

4.5 Gain Saturation

In the dispersion-managed soliton system, I must include gain saturation in order to obtain good agreement with the experimentally-observed evolution [28]. The basic assumption of the linearization approach is that $u_0 = \langle u \rangle$, namely that the average of the received signal including the noise equals the noise-free transmission. However, one must be careful in the presence of saturable amplifiers, since the noise power that the amplifiers add to the signal increases the total power of $u(t) = u_0(t) + \delta u(t)$, according to

$$\left\langle \|u\|^2 \right\rangle = \|u_0\|^2 + \left\langle \|\delta u\|^2 \right\rangle, \tag{4.5.1}$$

where $\|f\|^2 = (1/T)\int_0^T |f(t)|^2\, dt$ for any function $f(t)$. The term $\left\langle \|\delta u\|^2 \right\rangle$ is always finite and positive. Saturable amplifiers tend to keep the power of signal plus noise $\|u\|^2$ constant, hence when more noise is added to the signal, the noise-free signal power decreases. If one attempts to compute u_0 by simply switching off the ASE noise input in the simulation of the saturated EDFAs, the gains and hence the magnitude of the resulting field will be too large. Fortunately, EDFA saturation is a slow process that happens on a time scale of 1 ms, corresponding to 10–40 million bit periods in modern systems. Consequently, the amplifier gains adapt to the constant average power $\left\langle \|u\|^2 \right\rangle$ and cannot follow variations in the noise. Hence, the effect of amplifier saturation is a mere gain renormalization. Setting the amplifiers to match these reduced gains in a static gain model (static gain corresponds to infinite saturation power), one can obtain the correct zero-order solution u_0. In practice, my colleagues and I run about 100 Monte Carlo simulations and record the effective amplifier gains. We use these gains in the propagation of the noise-free signal u_0, while setting the saturation power to infinity.

4.6 Derivation of the Eye Diagram

In this section, I derive the pdf of the filtered output current of a square law detector. Similar pdfs have already been derived by Marcuse [48], Lee *et al.* [57], Bosco *et al.* [21], and Forestieri [58]. One can use this pdf to compute an electrical eye diagram that displays a continuous probability density rather than, as is traditional in simulations, overlaying a finite number of traces of marks and spaces with different noise realizations. The following derivation is valid for both the Monte Carlo method of Section 4.3 and the ODE method of Section 4.4. The inputs are the pdf of the timing jitter p_τ, the Fourier modes $\overline{A}_k$ of the renormalized noise-free signal, and the reduced covariance matrix $\mathcal{K}^{(r)}$. Below, I will drop the superscript and just write $\mathcal{K}$ for convenience.

The photodetector in the receiver converts the optical input signal plus noise to an electrical current $I(t)$. One may assume that the photodetector is an ideal square law detector with $I = \kappa |u|^2$, where κ is the receiver responsivity. I apply the transformation (4.3.4) and hence start by describing $I(t)$ in the absence of timing jitter. The electrical current is

$$\begin{aligned} I(t) &= \kappa \left| u_0(t) + \delta u(t) \right|^2 \\ &= \kappa \sum_{k=-N/2}^{N/2-1} \sum_{l=-N/2}^{N/2-1} \left(\overline{A}_k + r_k\right)^* \left(\overline{A}_l + r_l\right) \exp\left[it(\omega_l - \omega_k)\right] \\ &= \kappa \sum_{n=-N+1}^{N-1} \exp\left(it\omega_n\right) \sum_{k=k_{!}(n)}^{k_2(n)} \left(\overline{A}_k + r_k\right)^* \left(\overline{A}_{n+k} + r_{n+k}\right), \end{aligned} \tag{4.6.1}$$

where the r_k, r_l are the residual noise coefficients, $n = l - k$, $k_1(n) = \max(-N/2, -N/2 - n)$, and $k_2(n) = \min(N/2 - 1, N/2 - 1 - n)$. The current $I(t)$ then passes through a low-pass electrical filter. The last line in (4.6.1) represents the Fourier decomposition of $I(t)$ in the *electrical* domain, whose $2N - 1$ coefficients are given by the sum over k. Optical and electrical filtering can be introduced by multiplying the respective Fourier components by filter functions H^{opt} and H^{el}, respectively. The combined filtering operation yields the filtered voltage $y(t)$

$$\begin{aligned} y(t) &= \kappa \sum_{k=-N/2}^{N/2-1} \sum_{l=-N/2}^{N/2-1} H_k^{\text{opt},*} \left(\overline{A}_k + r_k\right)^* H_l^{\text{opt}} \left(\overline{A}_l + r_l\right) H_{l-k}^{\text{el}} \exp\left[it(\omega_l - \omega_k)\right] \\ &= \kappa \sum_{n=-N+1}^{N-1} H_n^{\text{el}} \exp\left(it\omega_n\right) \sum_{k=k_1(n)}^{k_2(n)} H_k^{\text{opt},*} \left(\overline{A}_k + r_k\right)^* H_{n+k}^{\text{opt}} \left(\overline{A}_{n+k} + r_{n+k}\right). \end{aligned} \tag{4.6.2}$$

The frequency term and all three filter terms can be combined into a complex $N \times N$ matrix

$$\mathsf{W}_{kl}(t) = \kappa H_k^{\text{opt},*} H_l^{\text{opt}} H_{l-k}^{\text{el}} \exp\left(it\omega_{l-k}\right). \tag{4.6.3}$$

One can write the filtered current $y(t)$ more compactly as

$$y(t) = \sum_{k,l=-N/2}^{N/2-1} (\overline{A}_k + r_k)^* \mathsf{W}_{kl}(t) \, (\overline{A}_l + r_l). \tag{4.6.4}$$

Since y is a real quantity, W must be self-adjoint. I introduce the real partitioned vector $\overline{\boldsymbol{A}} = (\overline{A}_{-N/2,R}, \dots, \overline{A}_{N/2-1,R}, \overline{A}_{-N/2,I}, \dots, \overline{A}_{N/2-1,I})^T$ to rewrite $y(t)$ as

$$y(t) = \left(\overline{\boldsymbol{A}} + \boldsymbol{r}\right)^T \mathcal{W}(t) \left(\overline{\boldsymbol{A}} + \boldsymbol{r}\right), \tag{4.6.5a}$$

$$\mathcal{W} = \begin{bmatrix} \mathsf{W}_R & -\mathsf{W}_I \\ \mathsf{W}_I & \mathsf{W}_R \end{bmatrix} = \mathcal{W}^T. \tag{4.6.5b}$$

The minus sign in the matrix in (4.6.5b) appears because W_I is anti-symmetric. The right-hand side of (4.6.5a) is a symmetric bilinear form, but, due to filtering, it is not necessarily positive. The receiver current $y_{\text{nf}}(t)$ in the absence of noise at time t is

$$y_{\text{nf}}(t) \equiv \overline{\boldsymbol{A}}^T \mathcal{W}(t) \overline{\boldsymbol{A}}. \tag{4.6.6}$$

In order to obtain an eye diagram, I must derive $p_y(y,t)$, the pdf of the current y at time t. The derivation of p_y is a generalization of Marcuse's [48] in that I consider all noise correlations and allow for arbitrary optical and electrical filtering. By contrast, Marcuse assumes optical white noise at the receiver and an integrate-and-dump circuit. The procedure of computing p_y starts with computing the Karhunen-Loève modes, the basis in which (4.6.5a) can be decomposed into a sum of independent random variables. In a second step, I calculate the pdf $p_{y,\tau=0}(y,t)$, where the subscript $\tau = 0$ means the exclusion of the effect of the timing jitter, as the convolution of the individual pdfs of these random variables. I solve the convolution by multiplying characteristic functions. In a third step, I add the effect of the timing jitter by convolving $p_{y,\tau=0}$ with the pdf of the timing jitter p_τ.

One must first find the functional basis that diagonalizes both the square law detection followed by the filtering, $\mathcal{W}(t)$, and the inverse covariance matrix $\mathcal{K}^{-1}$. Since both $\mathcal{W}(t)$ and $\mathcal{K}^{-1}$ are symmetric matrices and in addition $\mathcal{K}^{-1}$ is positive definite, one can apply the theorem of simultaneous diagonalization [59] which states that there is a real square matrix $\mathcal{C}$ satisfying

$$\begin{aligned} \mathcal{K}^{-1} &= \mathcal{C}^T\mathcal{C}, \qquad \text{and} \qquad &(4.6.7\text{a})\\ \mathcal{W} &= \mathcal{C}^T\Lambda\mathcal{C}. &(4.6.7\text{b})\end{aligned}$$

One procedure to obtain $\Lambda(t)$ and $\mathcal{C}(t)$ is to solve the generalized eigenproblem $\mathcal{W}\mathcal{C}^{-1} = \mathcal{K}^{-1}\mathcal{C}^{-1}\Lambda$. The solution can be performed on the computer, using a generalized eigenvalue routine such as the routine `eig()` in Matlab, or the procedure outlined in Appendix A.3. The matrix $\Lambda(t)$ is diagonal, and I write it as $\Lambda = \text{diag}(\lambda_1, \ldots \lambda_{2N})$, where the $\lambda_k(t)$ are real [59]. Note that if the impulse response of the filter can become negative, as in the case of a Bessel filter, some of the λ_k are negative. The transformation $\mathcal{C}$ yields the Karhunen-Loève modes of y, which are the noise-free signal modes $Q_k(t) \equiv \mathcal{C}_{kl}(t)\overline{A}_l$ and the independent noise modes $q_k(t) \equiv \mathcal{C}_{kl}(t)r_l$. I simplify (4.6.5a) to

$$\begin{aligned} y(t) &= \left(\overline{\boldsymbol{A}} + \boldsymbol{r}\right)^T \mathcal{C}^T\Lambda\mathcal{C}\left(\overline{\boldsymbol{A}} + \boldsymbol{r}\right)\\ &= \left(\boldsymbol{Q} + \boldsymbol{q}\right)^T \Lambda\left(\boldsymbol{Q} + \boldsymbol{q}\right)\\ &= \sum_{k=1}^{2N} \lambda_k\left(Q_k^2 + 2Q_kq_k + q_k^2\right) = \sum_{k=1}^{2N} g_k, \qquad (4.6.8)\end{aligned}$$

where the $g_k(q_k, t) \equiv \lambda_k\left(Q_k^2 + 2Q_kq_k + q_k^2\right)$ represent a new set of independent random variables.

The noise pdf $p_{\boldsymbol{r}}$ (4.2.11), where I replaced $\boldsymbol{a}$ by $\boldsymbol{r}$, can be factored into independent

Gaussian pdfs, using (4.6.7a)

$$\begin{aligned} p_{\boldsymbol{r}}(\boldsymbol{r}) &= (2\pi)^{-N}\sqrt{\det(\mathcal{K}^{-1})}\,\exp\left[-\tfrac{1}{2}\,\boldsymbol{r}^T\mathcal{C}^T\mathcal{C}\boldsymbol{r}\right] \\ &= (2\pi)^{-N}\sqrt{\det(\mathcal{K}^{-1})}\,\exp\left[-\tfrac{1}{2}\textstyle\sum_{k=1}^{2N} q_k^2\right] \\ &= \sqrt{\det(\mathcal{K}^{-1})}\prod_{k=1}^{2N} f_{q_k}(q_k), \end{aligned} \tag{4.6.9}$$

where the marginal pdfs $p_{q_k}(q_k) \equiv \exp(-q_k^2/2)/\sqrt{2\pi}$ are Gaussians with zero mean and unit variance. The factor $\sqrt{\det(\mathcal{K}^{-1})} = \det\mathcal{C}$ does not appear in the definition of the p_{q_k} since $\det\mathcal{C}\,d\boldsymbol{r} = dq_k$. Equation (4.6.8) is a sum of $2N$ random variables. The pdf $p_{y,\tau=0}(y,t)$ therefore equals the $(2N-1)$-dimensional convolution [45]

$$\begin{aligned} p_{y,\tau=0}(y,t) &= \int_{-\infty}^{\infty} p_{g_1}(y-g_2-g_3-\ldots-g_{2N},t) \\ &\times\; p_{g_2}(g_2,t)p_{g_3}(g_3,t)\ldots p_{g_{2N}}(g_{2N},t)\,dg_2dg_3\ldots dg_{2N}, \end{aligned} \tag{4.6.10}$$

but this convolution can be transformed into simple multiplications using characteristic functions, taking advantage of the convolution theorem [60]. The characteristic function $\Phi_h(\zeta)$ of a random variable h is defined as the expectation $\Phi_h(\zeta) \equiv E[\exp(i\zeta h)] = \int_{-\infty}^{\infty}\exp(i\zeta h)p_h\,dh$ [45]. With the help of the derived distribution identity $p_{g_k}\,dg_k = p_{q_k}\,dq_k$ [45], one can write

$$\begin{aligned} \Phi_{g_k}(\zeta) &= \int_{-\infty}^{\infty}\exp\Big[i\zeta g_k(q_k)\Big]p_{q_k}\,dq_k \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\Big[-\frac{q_k^2}{2}+i\zeta\lambda_k\big(Q_k^2+2Q_kq_k+q_k^2\big)\Big]\,dq_k \\ &= \frac{1}{\sqrt{1-2i\lambda_k\zeta}}\exp\left(\frac{i\lambda_kQ_k^2\zeta}{1-2i\lambda_k\zeta}\right). \end{aligned} \tag{4.6.11}$$

Again, note that the integration variables are all real, and that the entire analysis so far neglects the timing jitter. The characteristic function $\Phi_{y,\tau=0}(\zeta,t)$ of $p_{t,\tau=0}(y,t)$ equals the product of the Φ_{g_k}

$$\begin{aligned} \Phi_{y,\tau=0} &= \prod_{k=1}^{2N}\Phi_{g_k} \\ &= \left(\prod_{k=1}^{2N}\frac{1}{\sqrt{1-2i\lambda_k\zeta}}\right)\exp\left(i\zeta\sum_{k=1}^{2N}\frac{\lambda_kQ_k^2}{1-2i\lambda_k\zeta}\right). \end{aligned} \tag{4.6.12}$$

Note the identity $y_{\text{nf}} = \sum_{k=1}^{2N} \lambda_k(t)\, Q_k^2(t)$. Additive noise sources such as electrical noise can in principle be accounted for by multiplying $\Phi_{y,\tau=0}$ with the appropriate characteristic functions [57]. From $\Phi_{y,\tau=0}$ one obtains $p_{y,\tau=0}$ [45]

$$p_{y,\tau=0}(y,t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi_{y,\tau=0}(\zeta,t)\, \exp(-iy\zeta)\, d\zeta. \tag{4.6.13}$$

Due to the complicated dependence of $\Phi_{y,\tau=0}(\zeta,t)$ on ζ, it is not possible to evaluate the Fourier transform in (4.6.13) analytically. However, $p_{y,\tau=0}(y,t)$ can be computed numerically using a discrete Fourier transform.

If one sets all λ_k to the same positive value $\lambda_k = \sigma^2$, then $\Phi_{y,\tau=0}$ equals the characteristic function of a noncentral chi-square distribution [48, Eq. (20)], [50, Eq. (2-1-117)], and its inverse Fourier transform, the corresponding pdf, is

$$p_{\chi^2}(y) = \frac{1}{2\sigma^2} \left(\frac{y}{y_{\text{nf}}}\right)^{(n-2)/4} \exp\left(-\frac{y+y_{\text{nf}}}{2\sigma^2}\right) I_{n/2-1}\left(\frac{\sqrt{y\, y_{\text{nf}}}}{\sigma^2}\right), \qquad y \geq 0, \tag{4.6.14}$$

where I_l is the l-th order modified Bessel function of the first kind [61]. This function depends on three parameters: the noncentrality parameter y_{nf}, σ^2, and the degree of freedom $n = 2N$. The pdf $p_{\chi^2}(y)$ is the pdf of the squared sum y of n independent Gaussian-distributed random variables with the means $m_k = \sigma Q_k$ and identical variances $\sigma^2 = \lambda_k$. In this sense, (4.6.12) is the characteristic function of a chi-square distribution generalized to unequal and possibly negative "variances" λ_k. The reason that the λ_k also appear in the means $m_k = \sigma Q_k = \sqrt{\lambda_k} Q_k$ is due to my choice of diagonalizing $\mathcal{K}^{-1}$ and $\mathcal{W}$ according to (4.6.7a) and (4.6.7b). An alternative would be the diagonalization $\mathcal{K}^{-1} = \mathcal{C}^T \Lambda \mathcal{C}$ and $\mathcal{W} = \mathcal{C}^T \mathcal{C}$, but this choice requires a matrix $\mathcal{W}$ that is positive definite, implying electrical filters whose impulse response is positive.

I note that the mean of $p_{\chi^2}(y)$ is $\langle y \rangle = y_{\text{nf}} + n\sigma^2$, where the term $n\sigma^2$ is the expected current from the noise-noise terms in the receiver, its variance is $\sigma_y^2 = 4\sigma^2 y_{\text{nf}} + 2n\sigma^4$, where the terms are due to signal-noise and noise-noise beating in the receiver respectively [50], and the third-order central moment $\left\langle (y - \langle y \rangle)^3 \right\rangle$ equals $24\sigma^4 y_{\text{nf}} + 8n\sigma^6$. The limit of (4.6.14) for $y_{\text{nf}} \longrightarrow 0$ is the central chi-square pdf [50, Eq. (2-1-110)]

$$p_{\chi^2}(y)\Big|_{y_{\text{nf}}=0} = \frac{1}{\sigma^2 2^{n/2} \Gamma(n/2)} y^{(n/2-1)} \exp\left(-\frac{y}{2\sigma^2}\right). \tag{4.6.15}$$

The relationship of the final pdf p_y that includes timing jitter with $p_{y,\tau=0}$ is

$$p_y(y,t) = \int_{-\infty}^{\infty} p_{y,\tau=0}(y, t-\tau')\, p_\tau(\tau')\, d\tau', \tag{4.6.16}$$

where p_τ is the pdf of the timing jitter. I will show examples of p_τ and $p_y(y,t)$ in the next two chapters. Because y is phase-independent, the phase variation does not contribute to $\Phi_{y,\tau=0}$. Note that the integral in (4.6.16) is a convolution with respect to τ and could be expressed in terms of characteristic functions; however, $\Phi_{y,\tau=0}$ is a characteristic function with respect to y, and one cannot use it to simplify (4.6.16). Since the timing jitter is Gaussian distributed, we use a numerical Gauss-Hermite integration [61] to solve (4.6.16) as I will show in Appendix A.2. Although the λ_k are not identical in the simulations, $p_y(y)$ can sometimes be approximated by the chi-square pdf $p_{\chi^2}(y)$ with the parameters $\sigma^2 = \sigma_y^2/2(\langle y\rangle + y_{\text{nf}})$ and $n = 2(\langle y\rangle^2 - y_{\text{nf}}^2)/\sigma_y^2$.

Equation (4.6.16) is only valid if τ and $\boldsymbol{r}$ are uncorrelated. For example, if the average pulse shape in a transmission system depends on τ, then τ would be correlated with the residual noise $\boldsymbol{r}$, and τ would have to be incorporated into an extended covariance matrix and undergo the diagonalization procedure in order to decouple it from the noise continuum. I discuss this issue further in Section 6.1. We have found that these correlations can be safely neglected in optical transmission systems.

4.7 Pattern Dependences

So far, the entire treatment of noise linearization in this dissertation has dealt with the signal-noise interactions in the optical fiber. The previous sections show how an accurate pdf can be computed for a given bit sequence. The function $p_y(y,t)$ yields the probability density of receiving the filtered current y at time t in the presence of the noise-free signal $u_0(t)$. However, there are other physical effects in optical fiber systems besides ASE noise that lead to fluctuations in the current y. One example is polarization mode dispersion (PMD), which I do not discuss in this dissertation. A second example is phase noise that is introduced by the transmitter. Another example is the interaction of neighboring bits. In single-channel systems, neighboring pulses can overlap and interact nonlinearly [62], [63]. This interaction leads to a distortion of the pulse chirp, which in turn causes the pulses to walk off, inducing timing jitter and amplitude jitter [64]. I will show an example of signal distortion in a quasilinear CRZ system in Chapter 6. Signal-signal interactions are deterministic and can easily be computed for a given signal using a numerical simulation. The bit pattern of $u_0(t)$ has an important influence on the signal-signal interactions and I elaborate on this issue in the following. I restrict my discussion to single-channel systems. In WDM systems, all channels interact, leading to more complicated interactions. However, these interactions are independent of ASE noise and can in principle be treated separately.

The existence of pattern dependences affects the noise distributions. There is no longer a single pdf for the marks and a single pdf for the spaces. Instead, there is a different conditional pdf for each possible mark and each possible space in the system. The total pdfs for the marks and the spaces are determined by summing the conditional probabilities. As the number of channel in a WDM system grows, the number of possible patterns grows explosively, making a calculation of the complete pdfs impractical. A complete resolution of this problem is beyond the scope of this dissertation. However, work to date indicates that it is sufficient to focus on some number of the worst patterns.

A computational eye diagram is traditionally produced by overlaying the signal traces in all the bit slots in a given channel. I introduce the ***average pdf*** $p_{y,\text{eye}}(y,t_\text{eye})$ for the current y in the eye diagram in the time range $0 \le t_\text{eye} \le T_\text{bit}$. This pdf depends on $u_0(t)$ and on $p_y(y,t)$ at all points in time $t_k = t_\text{eye} + kT_\text{bit}$ that are overlaid in the eye diagram. One then obtains

$$p_{y,\text{eye}}(y,t_\text{eye}) = \frac{1}{n}\sum_{k=1}^{n} p_y(y,t_\text{eye}+kT_\text{bit}), \tag{4.7.1}$$

where $n = T/T_\text{bit}$ is the number of bits in the eye diagram. I will call the $p_y(y,t_\text{eye} + kT_\text{bit})$ ***partial pdfs***. Equation (4.7.1) can be refined by dividing $p_{y,\text{eye}}(y,t_\text{eye})$ into two pdfs $p_{y,\text{eye}}(y,t_\text{eye}) = \left[p_{y,\text{eye},0}(y,t_\text{eye}) + p_{y,\text{eye},1}(y,t_\text{eye})\right]/2$, one containing all the spaces (0's) and the other containing the marks (1's). Each of the pdfs $p_{y,\text{eye},0}(y,t_\text{eye})$ and $p_{y,\text{eye},1}(y,t_\text{eye})$ at $t_\text{eye} \approx T_\text{bit}/2$ near the center of the eye usually consist of very similar partial pdfs. In other words, $p_y(y,t)$ for a given $u_0(t)$ mainly depends on the noise-free current y_nf at time t, defined in (4.6.6). To make this more concrete, I conjecture that if $a = y_\text{nf}(t_2)/y_\text{nf}(t_1)$ is the ratio of the noise-free currents at the times t_1 and t_2, the approximation

$$p_y(y,t_2) \approx a p_y(ay,t_1) \tag{4.7.2}$$

holds in the case of $a \approx 1$. Note that $\int_{-\infty}^{\infty} p_y(y',t_2)\,dy' = \int_{-\infty}^{\infty} a p_y(ay',t_1)\,dy' = 1$.

Pattern dependences are important in single-channel systems, just like in WDM systems. However, each bit only interacts with a limited number of neighboring bits, thus limiting the number of relevant bit patterns. Hence, pattern dependences are simpler in single-channel systems than in WDM systems, and it is reasonable to begin the study of pattern dependences by focusing on single-channel systems.

How many relevant bit patterns are there in a single-channel system? Each bit is directly influenced by neighboring bits that overlap with it in time at any point during the optical transmission. The neighbors are in turn influenced by their neighbors and so forth; so, the interaction time is theoretically unlimited. However, the coupling between bits decays rapidly with time and one can define a minimum influential bit pattern length d. First, I define a minimum noise-free eye opening for patterns of length l,

$$\Delta y_{\mathrm{nf}}(l,t) = \min_{C_1[l]} [y_{\mathrm{nf}}(t)] - \max_{C_0[l]} [y_{\mathrm{nf}}(t)] , \tag{4.7.3}$$

where $C_1[l]$ is the set of all bit patterns of length l surrounded by infinitely many zero bits whose central bit is a mark, *e.g.* $C_1[3] = \{\ldots 0\mathbf{0}\mathbf{1}\mathbf{0}0\ldots, \ldots 0\mathbf{1}\mathbf{1}\mathbf{0}0\ldots, \ldots 0\mathbf{0}\mathbf{1}\mathbf{1}0\ldots, \ldots 0\mathbf{1}\mathbf{1}\mathbf{1}0\ldots\}$, and the time t lies within this central bit slot. The sets $C_0[l]$ are defined analogously with a space in the central position instead of a mark. Note that $\Delta y_{\mathrm{nf}}(l,t)$ can become negative. I define $\Delta y_{\mathrm{nf,max}}(t) = \lim_{n\to\infty} \Delta y_{\mathrm{nf}}(l,t)$ for $0 \le t \le T_{\mathrm{bit}}$ for a single-channel system. A maximum relevant pattern length d can then be defined to be the smallest l with

$$\max_t \left| \Delta y_{\mathrm{nf,max}}(t) - \Delta y_{\mathrm{nf}}(l,t) \right| < y_0, \tag{4.7.4}$$

where y_0 is a current that is small compared to the average current induced by noise.

Anderson and Lyle studied a system where only nearest-neighbor optical pulse-to-pulse interactions occur and found $d = 3$ [65]. In the CRZ system that I study in Section 6.2, $d \approx 5$.

For any d, there is a set of cyclic bit strings of length $n = 2^d$ that contain all 2^d bit patterns of length d, called ***deBruijn sequences*** [66]. DeBruijn sequences are not unique in general. The well-known pseudorandom bit sequences (PRBS) of length $2^d - 1$ are derived from the deBruijn sequences by removing one zero bit from the substring that consists of d consecutive zeros. For example, a deBruijn sequence for $d = 3$ is the cyclic string 11101000 which contains all eight patterns 000, 001, 010, 100, 011, 110, 101, and 111, and a PRBS sequence of length $2^3 - 1 = 7$ is 1110100. DeBruijn sequences are important because they allow one to study a signal that contains all bit patterns of length d while requiring the minimum bit string length $n = 2^d$.

As noted earlier, it is useful if $p_{y,\mathrm{eye}}(y,t_{\mathrm{eye}})$ in (4.7.1) can be approximated by just computing a few partial pdfs, since that would reduce the computational work. I call a partial pdf *dominant* if

$$p_y(y_{\text{opt}}, t_{\text{eye}} + mT_{\text{bit}}) \gg \sum_{k=1, k\neq m}^{n} p_y(y_{\text{opt}}, t_{\text{eye}} + kT_{\text{bit}}), \tag{4.7.5}$$

where y_{opt}, defined in the next chapter, is near the optimum decision level at which $p_{y,\text{eye}}(y, t_{\text{eye}})$ assumes its minimum. If one of the partial pdfs becomes dominant, the other partial pdfs can be neglected.

Chapter 5

Monte Carlo Simulations

5.1 Dispersion-Managed Soliton System Design

In this section, I describe Monte Carlo simulations that my colleagues and I carried out to compute the covariance matrix $\mathcal{K}^{(r)}$. We simulated a dispersion-managed soliton (DMS) system with a transmission distance of 24,000 km. This system is very well characterized both in simulation and in experiment [28], [67]–[69]. The system does not operate in the quasilinear DMS regime [70], but at an optical peak power of about 8 mW, which makes the transmission significantly nonlinear. The evolution of the pulse shape is approximately periodic with the period given by the dispersion map. We verified that the phase jitter obeys a Jacobi-Θ distribution, which is the periodic analogue of a Gaussian distribution, and we also verified that the timing jitter is Gaussian distributed. Finally, we verified that the real and imaginary parts of the residual noise Fourier coefficients, after the jitter is separated, are Gaussian distributed.

The simulated transmission line is shown in Fig. 5.1 and consists of 225 periods of a dispersion map of length 106.7 km [28]. Each map contains a fiber span of length 4×25 km long with a normal dispersion of -1.03 ps/nm-km and a span of length 6.7 km with an anomalous dispersion of 16.7 ps/nm-km, denoted by the circles labeled N and A respectively. The path average dispersion equals 0.08 ps/nm-km, which is larger than in [28]. Third-order dispersion is not relevant in this system [28] and is set to zero. The carrier wavelength is 1551.49 nm, matching the experimental value. The fiber loss is compensated by five EDFAs. One EDFA follows each of the four 25-km segments of normal-dispersion fiber, and the fifth follows the segment of anomalous-dispersion fiber. There is a 2.8 nm (350 GHz) optical bandpass filter (OBF) in each map period to reduce the amount of noise.

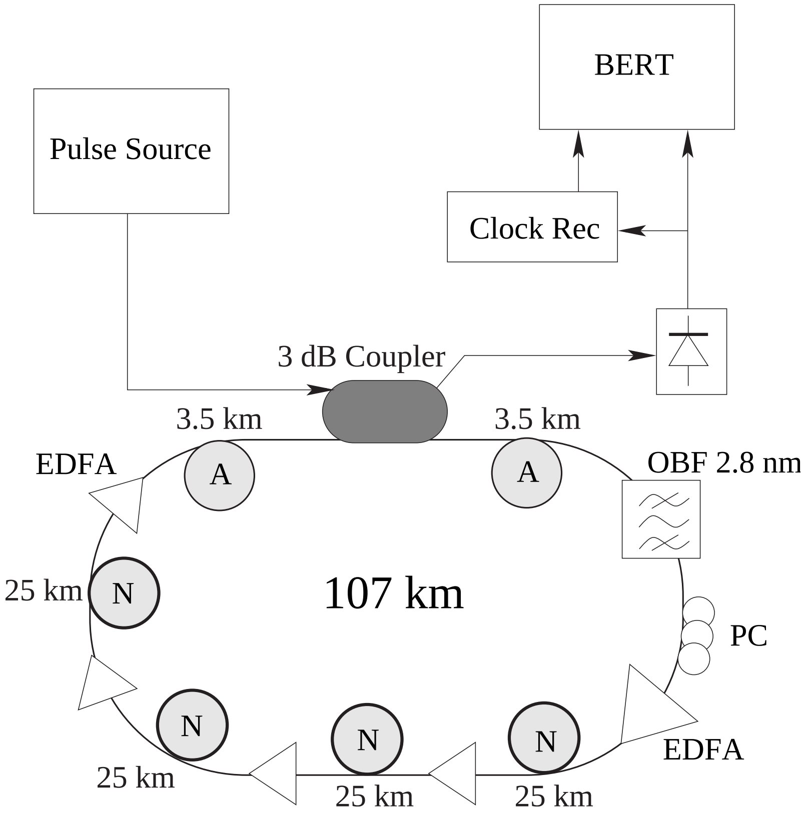

Figure 5.1: Cartoon of the DMS system. PC denotes the manual polarization controller, BERT the bit error ratio tester, ClockRec the clock recovery circuit, and OBF the optical bandpass filter.

As described in Section 4.5, I model the amplifiers as EDFAs with static gain, as opposed to explicitly including gain saturation. I carefully adjust the static gains so that they equal the effective gains one would obtain using EDFAs with a saturation time of 1 ms and a saturation power of 10 mW, similar to [28]. The spontaneous emission factor is $n_{\text{sp}} = 1.4$. In each amplifier, I add a random amount of lumped noise separately to the real and imaginary parts of the signal in the Fourier domain. This noise input by each amplifier is Gaussian-distributed with zero means and variances $\rho_m^2/2$, where $\rho_m^2 = (G_m - 1)\eta$, and G_m is the power gain associated with the m-th amplifier [28]. All other parameters are defined in Section 4.2. I choose a Box-Mueller generator [71] to obtain the Gaussian-distributed random variables; the generator takes its inputs from a 48-bit random number generator. The launched pulses have a Gaussian shape with a FWHM duration of 9 ps, and the signal is injected and received at the chirp-free midpoint of the anomalous span. The transmission distance of 24,000 km is 400 times larger than nonlinear scale

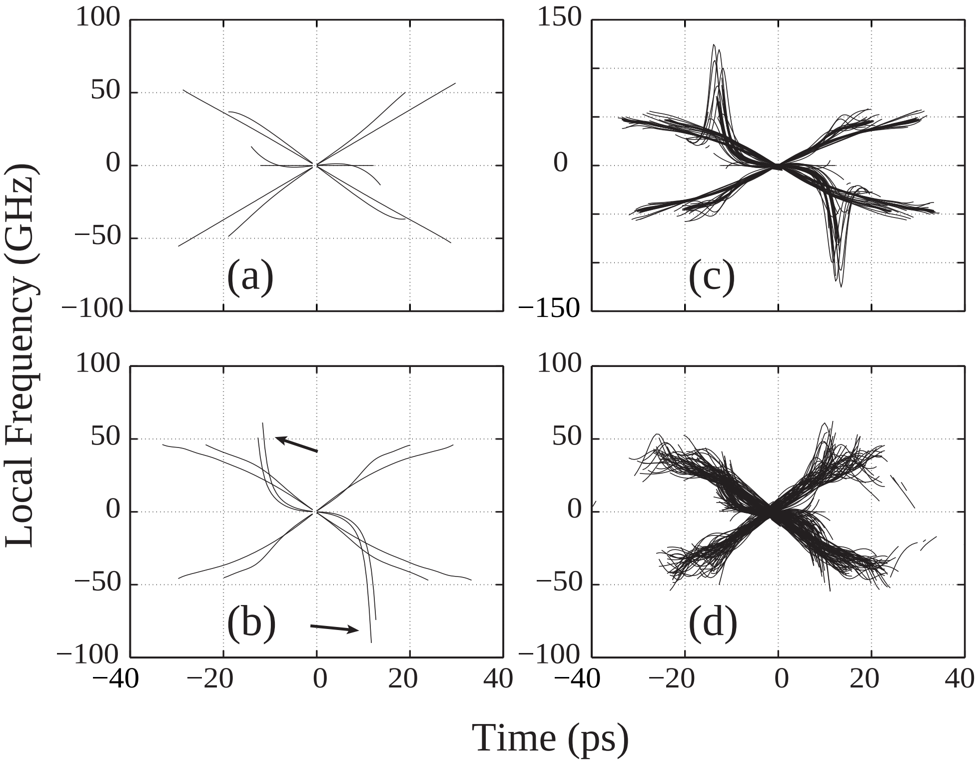

Figure 5.2: Phase space portraits for the DMS system at the beginning of each fiber span. (a) First map period; (b) Last map period, arrows: middle of the anomalous span; (c) All 225 periods; (d) Same as (b), but with added ASE noise.

length L_{nl} (2.3.5). I transmit the 4-bit sequence 1000 in a total simulation time window of $T = 400$ ps; hence, there are no optical interpulse interactions. The receiver is modeled as in Fig. 2.1 as an ideal square law detector followed by an electrical low-pass 5th-order Bessel filter with a one-sided 3-dB bandwidth of 4.3 GHz. This bandwidth is much smaller than the commonly used bandwidth of 70–80% of the data rate, but it was shown to be advantageous in this experiment to suppress the effects of timing jitter [28].

I use the split-step Fourier method to solve the scalar nonlinear Schrödinger equation, which only takes into account one optical polarization. In the recirculating loop that I am modeling, the polarization dependent loss (PDL) is large and the polarization controllers are optimized to pass the signal with minimum loss. Consequently, the signal is dominated by one polarization, and the orthogonal polarization can be neglected. The nonlinear propagation equation in all simulations in this dissertation is solved using a third-order

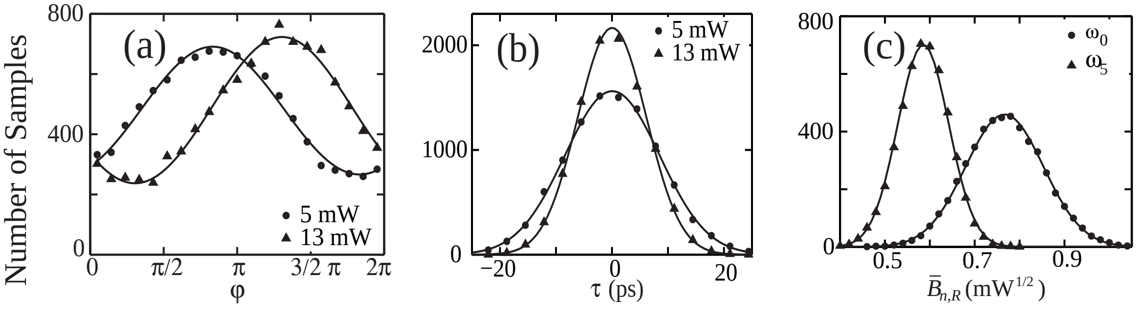

Figure 5.3: (a) Histogram of the phase offset φ and (b) histogram of the time offset τ for two different simulations with signal peak powers $P_{\text{peak}} = 5$ mW and $P_{\text{peak}} = 13$ mW. (c) Histogram of the two real Fourier coefficients $\bar{B}_{0,R}$ and $\bar{B}_{5,R}$ at $\omega_0 = 0$ and $\omega_5 = 2\pi \times 25$ GHz, respectively, after the phase and time offsets are removed ($P_{\text{peak}} = 5$ mW). The solid lines are fits of the Jacobi-Θ function in (a) and Gaussians in (b) and (c).

split-step algorithm [72].

I present a phase space portrait of the DMS system with the purpose of demonstrating some characteristic properties of a periodically stationary DMS system and for later comparison with the CRZ system. Phase space portraits contain more information than time- or frequency domain plots; for instance they are a good way of visualizing chirp. Fig. 5.2 shows four different phase portraits, by which I mean trajectories in the space of local frequency versus time. Related plots have been produced experimentally for various systems, using a technique called Frequency Resolved Optical Gating (FROG) [73]. The following discussion applies to noise-free signals. The local frequency of an optical signal $u(t)$ is defined as the derivative of the local phase with respect to time

$$f_{\text{loc}}(t) = \frac{1}{2\pi} \frac{d\arg[u(t)]}{dt}, \tag{5.1.1}$$

where $\arg[u(t)] = \arctan[\operatorname{Im} u(t)/\operatorname{Re} u(t)] = \varphi(t)$ is the local phase of $u(t)$. To produce the graphs, I transmit a single pulse centered at $t = 0$ in a 400 ps time window without adding ASE noise. Six times during each dispersion map, at the beginning of each fiber span and in the middle of the anomalous span, I iterate over each sample point in time and save the pairs (t, f_{loc}) if the local pulse power $|u(t)|^2$ is larger than 0.5% of the peak pulse power $\max(|u(t)|^2)$.

Fig. 5.2(a) shows the six phase portraits during the first period of the dispersion map. The straight horizontal line represents the launched signal, which is an unchirped Gaussian. Chirp is defined as the second time derivative of the local phase, $d^2 \arg[u(t)]/dt^2 = 2\pi\, df_{\text{loc}}/dt$. Consequently, the portrait of a chirped pulse will have a nonzero slope. The oblique lines show the pulse at later points in the dispersion map. Fig. 5.2(b) shows the same portrait as Fig. 5.2(a), but for the final map period (period 225). The difference between Figs. 5.2(a) and (b) demonstrates that the pulse chirp of the dispersion-managed soliton evolves. The curve that is marked by the arrows shows the pulse portrait at the middle of the anomalous span. The chirp is small for $-5\,\mathrm{ps} < t < +5\,\mathrm{ps}$; however, the pulse tails are strongly chirped because the chirp-free point in a map is only located in the middle of the anomalous span if the fiber is assumed to be lossless, but it moves closer to the beginning of the anomalous span in a lossy system [74].

Fig. 5.2(c) shows the pulse portrait at the beginning of each fiber span for the entire transmission over 225 map periods, and Fig. 5.2(d) shows the same, but with added ASE noise and a power cutoff of 3%. In all four graphs, one can see that the solitons never extend beyond $-40\,\mathrm{ps} < t < +40\,\mathrm{ps}$ with a significant pulse power, and hence the overlap of the pulse tails is negligible. Note that phase portraits are not directly related to the least-squares phase fit of (4.3.5) since the phase portraits show phase in the time domain, not in the frequency domain.

5.2 Jitter Separation

Using the least-squares method outlined in Section 4.3, I now show that the central time offsets of the pulses τ are individually Gaussian-distributed. The same holds for the $\overline{B}_{n,R}$ and $\overline{B}_{n,I}$, where $\overline{B}_n = \overline{B}_{n,R} + i\,\overline{B}_{n,I}$, and the $\overline{B}_n$ are the transformed Fourier coefficients defined in (4.3.4). Furthermore, I show that the distribution of the phase offset φ is a Jacobi-Θ function, which is the periodic analogue of a Gaussian [61], defined by

$$\Theta(\mu_\varphi, \sigma_\varphi^2, 2\pi) = \sum_{k=-\infty}^{\infty} N(\mu_\varphi + 2\pi k, \sigma_\varphi^2), \tag{5.2.1}$$

where $N(\mu, \sigma^2)$ is a Gaussian (normal) distribution of mean μ and variance σ^2. The Θ-function is the natural choice for the phase fit since φ at the receiver is only determined modulo 2π unless one tracks its evolution.

Figs. 5.3(a) and (b) show histograms of φ and τ for the two different signal peak powers P_{peak} of 5 mW and 13 mW. The two histograms are approximations to $p_\varphi(\varphi)$ and $p_\tau(\tau)$, respectively. The simulation consists of 10,000 Monte Carlo runs. The phase distribution $p_\varphi(\varphi)$ converges to the Jacobi-Θ function, and $p_\tau(\tau)$ converges to a Gaussian distribution. The number of samples per bin in each histogram in Fig. 5.3(a) never falls below 200, and hence the phase wrapping is relevant here. However, one could in principle track the evolution of the phases and retain the full information, yielding a phase in the range $-\infty < \varphi < \infty$. The corresponding pdf $p'_\varphi(\varphi)$ would be Gaussian.

Fig. 5.3(c) shows histograms of $\overline{B}_{0,R}$ and $\overline{B}_{5,R}$ at the angular frequencies $\omega_0 = 0$ and $\omega_5 = 2\pi \times 25$ GHz respectively. The simulated data agree very well with the Gaussian fit. The algorithm for removing the linear part of the signal phase causes the imaginary parts of the $\overline{B}_k$ to be close to zero; so, they are not shown here. Using a chi-square statistical test [45], I verified that all the $\overline{B}_{k,R}$ in the simulation are Gaussian distributed .

Adding the effect of timing jitter when the electrical pdfs are computed by employing the convolution in (4.6.16) relies on the independence of the residual noise $\boldsymbol{r}$, or at least the electrical current y, on the timing jitter τ. I compute the normalized cross-correlations

$$C[u,v] = \frac{\langle uv \rangle - \langle u \rangle \langle v \rangle}{\text{std}(u)\,\text{std}(v)}, \tag{5.2.2}$$

where u and v are random variables and $\text{std}(z) = \sqrt{\langle z^2 \rangle - \langle z \rangle^2}$ is the standard deviation of a quantity z. My simulations show that both $C[\overline{B}_{k,R}, \varphi]$ and $C[\overline{B}_{k,I}, \varphi]$, as well as $C[\overline{B}_{k,R}, \tau]$ and $C[\overline{B}_{k,I}, \tau]$ are significantly nonzero, where $\overline{B}_k = \overline{B}_{k,R} + i\overline{B}_{k,I}$ are the Fourier coefficients of the noisy signal after the separation of phase and timing jitter, and φ and τ are the phase and timing offsets of a given noise realization respectively, according to (4.3.5). These correlations seem to pose a problem in the computation of f_y, if not accounted for.

The physical reason for these correlations is that noise realizations that increase the pulse power cause a faster phase rotation which tends to lead to extreme phase excursions. Analogously, noise realizations with extreme frequency offsets tend to produce extreme timing offsets, the Gordon-Haus effect [17], [19]. However, I find that $C[\overline{B}_{k,R}, y]$ and $C[\overline{B}_{k,I}, y]$ are zero to the limit of the numerical precision, implying that the correlations of the $\overline{B}_k$ with φ and τ have no effect on the receiver current y. This paradox is explained by the nature of the square law detector that both neglects the optical phase and works as a demodulator, removing shifts in the central frequency.

5.3 Computation of the Covariance Matrix

My colleagues and I employed Monte Carlo simulations to compute the reduced covariance matrix $\mathcal{K}^{(r)}$ from (4.3.6). I drop the superscript (r) in the following. In this section, I

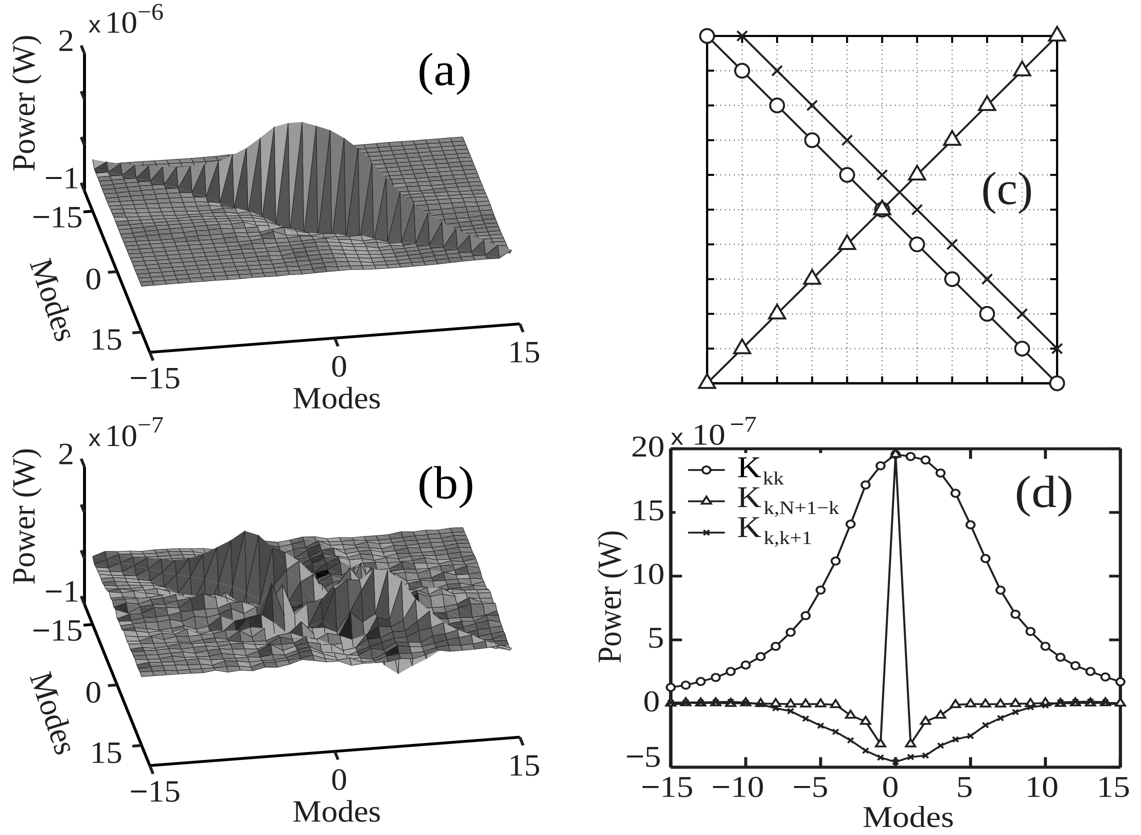

Figure 5.4: (a) Two-dimensional plot of the RR block of $\mathcal{K}$ for ω_k in the range $k = [-15, 15]$. (b) Same for the RI block of $\mathcal{K}$. (c) Cartoon of a square matrix with three slices. Open circles: principal diagonal; hatch marks: first parallel to the principal diagonal; triangles: secondary diagonal. (d) The values of the RR block along these slices.

show $\mathcal{K}$ for the dispersion-managed soliton system that we study, and I compare the resulting pdfs for the marks and the spaces in the following section. The point of characterizing $\mathcal{K}$ is to reveal the significant difference between Marcuse's approach of assuming optical white noise at the receiver [48] and the covariance matrix method. Marcuse's method sets the principal diagonal elements to a constant value and sets all other elements to zero.

According to (4.2.12), $\mathcal{K}$ is a block matrix consisting of 4 blocks of size $N \times N$ each. I label the four blocks RR, RI, IR and II, where RR is the block $\langle \boldsymbol{\alpha}_R \boldsymbol{\alpha}_R^T \rangle$, RI is $\langle \boldsymbol{\alpha}_R \boldsymbol{\alpha}_I^T \rangle$, and so on. Since $\mathcal{K}$ is symmetric, $(\mathrm{RI})^T = \mathrm{IR}$. In the following, the elements of the matrix $\mathcal{K}$ are indexed by the frequencies whose covariance is located at each element, so that the upper left matrix element is $\mathcal{K}_{-N/2,-N/2}$, the lower right one is $\mathcal{K}_{N/2-1,N/2-1}$, and the center element is $\mathcal{K}_{0,0}$.

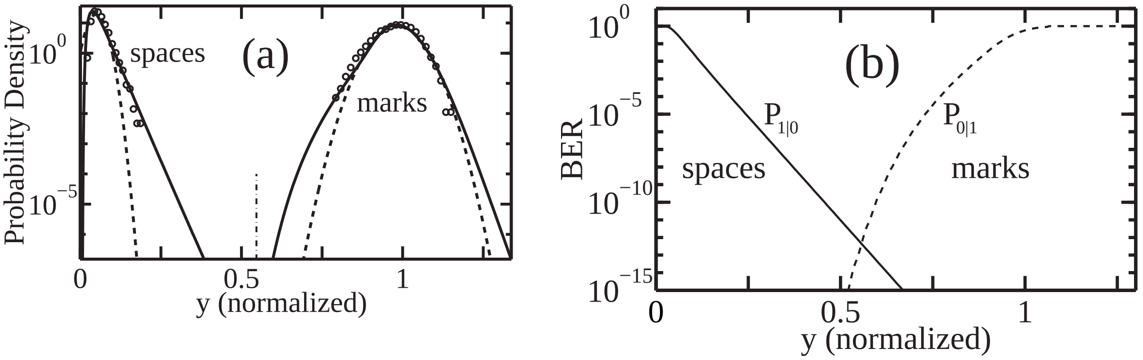

Figure 5.5: (a) Probability density function of the filtered current y after square-law detection and a 4.3 GHz narrow-band filter. The solid lines are the pdfs calculated using linearization at the center of the bit window of the marks (ONES) and the spaces (ZEROS); the dashed lines are Gaussian fits. Note the bump on the left tail of the pdf of the marks. The circles are direct results from the Monte Carlo simulation. Particularly for the spaces, the agreement with the pdf obtained by linearization is much better than with the Gaussian fit. (b) The function $P_{1|0}(y)$ is shown as a solid line, and $P_{0|1}(y)$ is shown dashed. The BER is defined as their mean. The optimum decision level lies at $y = 0.55$ near the intersection of the graphs, shown as the vertical dash-dotted line in (a), and it yields a BER of 5×10^{-13}. From the Gaussian fits, one obtains a Q-factor of 13.5 which would imply an optimal BER of 10^{-41}.

Fig. 5.4(a) displays the block RR in a two-dimensional form. The ridge along the principal diagonal represents the variances $\mathcal{K}_{kk}$ and all other elements correspond to cross-covariances. The cw entry lies at $k = 0$. Fig. 5.4(b) shows the block RI whose maximum is about one order of magnitude smaller than that in RR and II. The RI and IR blocks are point-symmetric around their central elements ($\mathrm{RI}_{k,l} = \mathrm{RI}_{-k,-l}$), while RR and II blocks are symmetric ($\mathrm{RR}_{k,l} = \mathrm{RR}_{l,k}$). Fig. 5.4(c) shows a cartoon of a square matrix with the principal diagonal (solid line), a parallel to the principal diagonal (crosses), and the secondary diagonal (circles), starting from the indices $(-15, 15)$ and ending at $(15, -15)$. Fig. 5.4(d) shows the values of the RR block along the slices in (c). The shape of the graph of $\mathcal{K}_{kk}$ is a consequence of the optical inline filtering in the recirculating loop, as well as optical signal-noise interactions. Because of the inline filter, $\mathcal{K}_{kk}$ vanishes for $|k| \gg 15$. In the absence of inline filters, $\mathcal{K}_{kk}$ would converge for large $|k|$ to a finite value. At small frequencies $|k| < 15$, the signal-noise interactions lead to parametric gain and emerge as a peak whose shape is similar to the signal power spectrum $|A_k|^2$. The slices along the circles and crosses reveal that the values of $\mathcal{K}_{k,k\pm l}$ at small $|k|$ and $|l|$ are actually negative, leading to two elongated troughs along both sides of the principal diagonal.

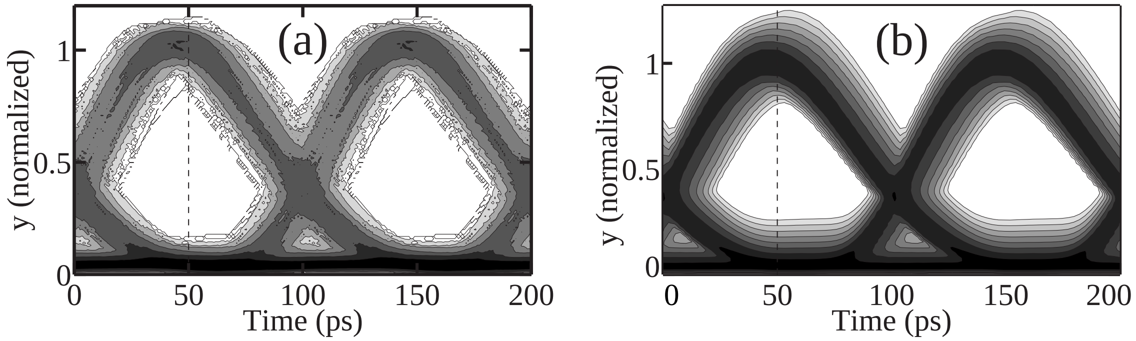

Figure 5.6: (a) Eye diagram, generated from the Monte Carlo simulation of the dispersion-managed soliton. The probability density of the current y is displayed as a contour plot. The dashed line at $t = 50$ ps shows the location of the pdf in Fig. 5.5. The logarithm of the pdf is displayed as different shades of gray. (b) Accurate eye diagram produced by the linearization approach. To obtain a more readable diagram, I only plot probability densities in the range $[10^{-4}, 10^{1}]$. However, the approach allows me to find the probability density at any point (t, y), thereby enabling me to accurately calculate BERs.

5.4 Calculating the BER

We calculated the BER as a function of the decision threshold, and we determined its value at the optimum decision point. Fig. 5.5(a) shows the pdf f_y of the filtered current $y(t)$ defined in (4.6.5a) that corresponds to the output of the electrical receiver. The current $y(t)$ is normalized to the mean of the pdf in the marks. The calculation is performed as explained in the previous chapter in two different bit slots, corresponding to the 1 and the central 0 in the 1000 bit pattern that we simulated, so that we obtained the pdfs for the marks and the spaces separately. The effect of the timing jitter is included in the calculation of the pdfs, using (4.6.16). The Gaussian fit is a good approximation over about two orders of magnitude, but it deviates strongly at low probability densities. At small values of y, the pdf of the marks is dominated by the timing jitter, leading to a visible bump. This bump exists because the current y at the center of the eye is lower for a strongly jittered pulse. Without this bump, the left tail of the accurate pdf for the marks would cross the Gaussian fit, and then run inside, so that the error probability density would be lower there than for a Gaussian pdf, in agreement with [48]. I conclude that strong timing jitter can lead to an increased eye penalty that must be taken into account. Note that the path average dispersion of 0.08 ps/nm-km enhances the timing jitter and differs from the value that was used in Section 6.1.

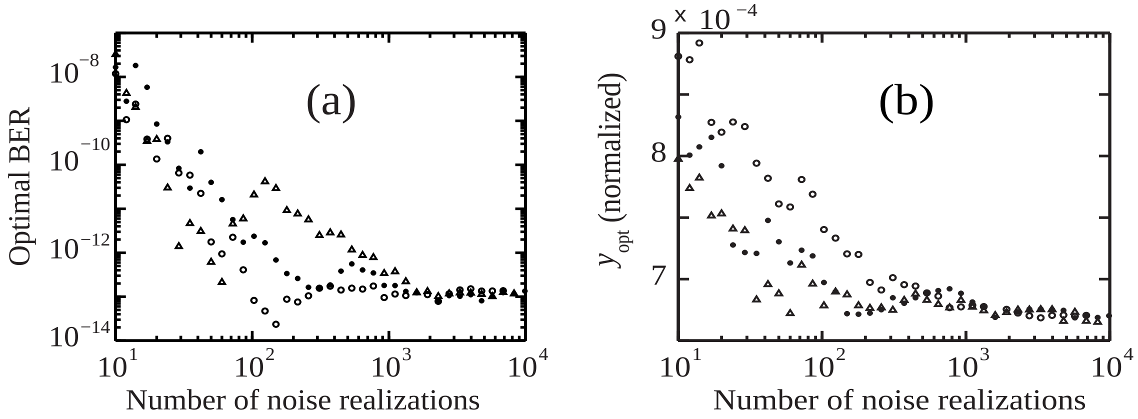

Figure 5.7: (a) Convergence of optimal BER values, resulting from $\mathcal{K}^{(r)}$ and A_k obtained from a Monte Carlo simulation, as a function of the number of noise realizations. The symbols $(\triangle)$, $(\bullet)$, and $(\circ)$ pertain to three simulations that were started with different random seeds. (b) Convergence of the optimal decision level y_{opt}. Both BER and y_{opt} converge after about 2,000 Monte Carlo realizations.

The knowledge of the separate pdfs of the marks and spaces allows one to calculate the bit error probabilities. I define the two probabilities

$$P_{1|0}(y) \;=\; \int_{y}^{\infty} f_y(y', t_1)\, dy', \tag{5.4.1a}$$

$$P_{0|1}(y) \;=\; \int_{0}^{y} f_y(y', t_0)\, dy'. \tag{5.4.1b}$$

The quantity $P_{1|0}(y)$ is the probability of detecting a mark when a space was transmitted, using the decision level y, and analogously $P_{0|1}(y)$ is the probability of detecting a space when a mark was sent. The function $f_y(y, t_1)$ is the pdf of y, taken at the central time in a bit window t_1 when a mark is received, and $f_y(y, t_0)$ is taken at the central time in a bit window when a space is received, t_0. I define the bit error probability at the decision level y as [48]

$$\mathrm{BER}(y) = \tfrac{1}{2}\left[P_{1|0}(y) + P_{0|1}(y)\right]. \tag{5.4.2}$$

I define the optimum decision level y_{opt} to be the decision threshold at which the BER becomes minimal. Fig. 5.5(b) shows $P_{1|0}(y)$ and $P_{0|1}(y)$, as well as the BER as a function of y. In the DMS system, we find that at $y_{\mathrm{opt}} = 0.55$ the BER is 5×10^{-13}. From the Gaussian fit, we obtain a Q-factor of 13.5 implying an optimal BER of 10^{-41}. This large difference in the BERs is again a consequence of the large timing jitter.

Fig. 5.6 shows the corresponding eye diagram, where y is normalized the same way as in Fig. 5.5. It is a contour plot of the pdf for $y(t)$ at each time t. This way of plotting the eye diagram is closer to what is measured experimentally than the common practice of simply superimposing discrete traces of $y(t)$ in the different bit windows. The optimal decision point for the system lies close to 0.5. Gaussian extrapolations usually yield optimal decision levels that are much smaller than 0.5 in normalized units [48].

I now consider the off-diagonal elements in $\mathcal{K}$. Although $\mathcal{K}$ is diagonally dominant, we find that the off-diagonal elements have a large impact on the BER. We compared the optimal BER to a computation in which we set the off-diagonal elements to zero. It turns out that the spread in the pdf of the spaces is reduced, as one might expect, while the spread in the pdf of the marks is increased, leading to a lower y_{opt} and an optimal BER that deviates from the true value by orders of magnitude. I therefore conclude that the nonlinearity can lead to a substantial cross-correlation in $\mathcal{K}$ and it is in fact necessary to take account of off-diagonal elements.

I next turn to the question of the accuracy of the BER and the decision level that I obtain by employing a Monte Carlo simulation to compute $\mathcal{K}$. Fig. 5.7 shows the convergence of the BER and the optimal decision level y_{opt} as the simulation proceeds. Both $\mathcal{K}$ and y_{opt} converge as we average over more noise realizations. The convergence requires on the order of a few thousand realizations.

I note that the values for the optimal BER and the decision level both drop as the averaging proceeds. The statistical fluctuations in $\mathcal{K}$, which vanish as the number of realizations increases, tend to decrease the BER, irrespective of their signs. Thus, the square law detector with its dependence of $f_y(y)$ on $\mathcal{K}$ corresponds to a biased estimator [75].

In conclusion, the computation of the covariance matrix by averaging over Monte Carlo simulations is robust, but it requires a large number of noise realizations in order for the BER to converge. Computing the pdfs and the BER using the covariance matrix in essence applies a multivariate-Gaussian fit to the optical noise, instead of a one-dimensional Gaussian fit to the receiver current histogram. In traditional Monte Carlo simulations, the only result of each noise realization that is used in the calculation of the electrical BER is the narrow-band filtered current $y(t)$, a scalar variable for each point in time. By contrast, the covariance matrix contains much more information and makes much better use of the available information.

Chapter 6

Covariance Matrix Method: Results

In this chapter, I apply the deterministic method described in Section 4.4 to propagate the noise covariance matrix, rather than running Monte Carlo simulations. I consider two different systems: the DMS system introduced in Chapter 5 with the reduced path average dispersion of 0.02 ps/nm-km and a chirped return to zero (CRZ) system with a total transmission distance of 6,100 km. In the following, I refer to the reduced covariance matrix $\mathcal{K}^{(r)}$ as $\mathcal{K}$, dropping the superscript.

6.1 The Dispersion-Managed Soliton System

The calculation that I present here is completely deterministic in contrast to Chapter 5, where I reported on the calculation of $\mathcal{K}$ based on Monte Carlo simulations [76]. This deterministic approach requires substantially less CPU time while producing a much higher degree of accuracy. For the 4-bit sequence considered in the previous chapter, the approximation of the covariance matrix with 5,000 Monte Carlo realizations required 72 hours of CPU time, while the deterministic covariance matrix method required only 5 hours of CPU time on a 400 MHz Pentium III PC. As in the last chapter, the nonlinear propagation equation is solved by a third-order split-step algorithm [72]. I transmit the 8-bit deBruijn sequence 11010001 in a time window of 800 ps in a single channel; all pulses are co-polarized. The Fourier vector length N_{FFT} is 2048.

Fig. 6.1(a) shows the optical power $|u_0(t,L)|^2$ of the noise-free 8-bit signal at the end of the transmission in the time domain. The tic marks indicate the boundaries of the bit slots. Fig. 6.1(b) shows the optical power spectrum at the end of the transmission as the circles. The 10 GHz tones and their harmonics are clearly visible. The dots show the average power spectrum of the noise as obtained from $|a_k|^2 + |a_{k+N}|^2 = \mathcal{K}_{k,k} + \mathcal{K}_{k+N,k+N}$

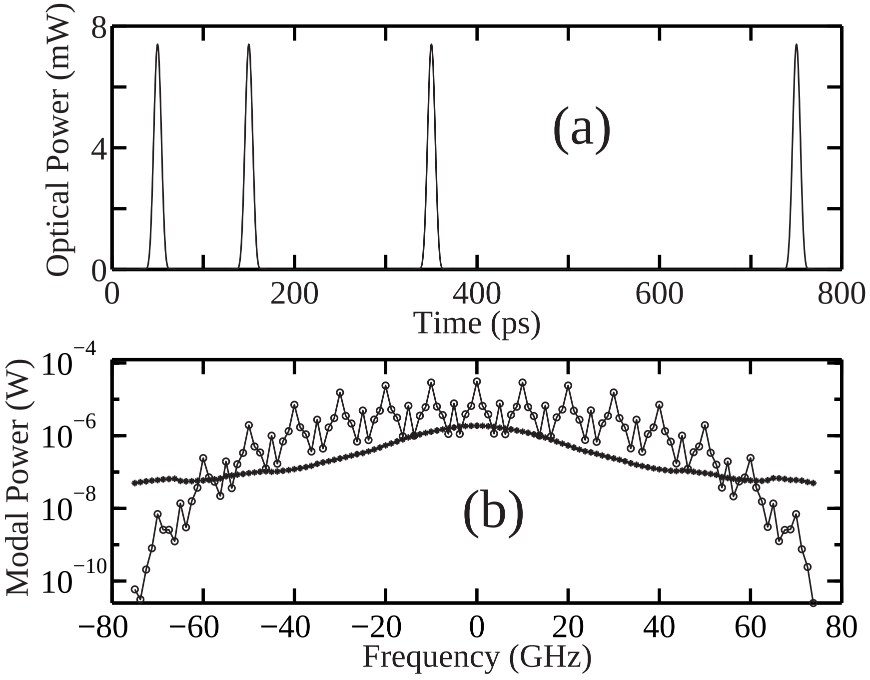

Figure 6.1: DMS system. (a) Noise-free optical power in the time domain $|u_0(t, L)|^2$ of the noise-free 8-bit signal at the end of the transmission. (b) Circles: optical power spectrum $|A_k|^2$ of the signal in logarithmic scale. The 10 GHz tones and their harmonics are clearly visible. Dots: average power spectrum of the noise from $\mathcal{K}$. The OSNR is 8.44 dB and $N = 120$.

. The number of modes in $\mathcal{K}$ is $N = 120$. The optical signal-to-noise ratio (OSNR), which I define as the ratio of total signal power to total noise power in the shown bandwidth, is 6.98, or 8.44 dB. The bandwidth of 160 GHz in this definition of the OSNR seems very large, but the OSNR would not change very much by reducing the bandwidth because of the inline filter that attenuates both the signal and the noise at high frequencies.

Fig. 6.2(a) shows three slices through the RR part of the reduced covariance matrix $\mathcal{K}$, where the inset gives a pictorial representation of a matrix whose elements are located at the grid points. The diagonal lines show the location of the three slices. The solid line hence runs along the principal diagonal, the circles run on the secondary diagonal, and the asterisks show elements on the nearest parallel to the principal diagonal. The large ratio of the noise power at frequency $f = 0$ to the power at $f = 75$ GHz is primarily due to the inline optical filter in the recirculating loop and only secondarily to parametric gain.[1]

Fig. 6.2(b) shows the eigenvalue spectrum of $\mathcal{K}$, where the the eigenvalues are sorted

---

[1] See Section 3.1 for a discussion of parametric gain.

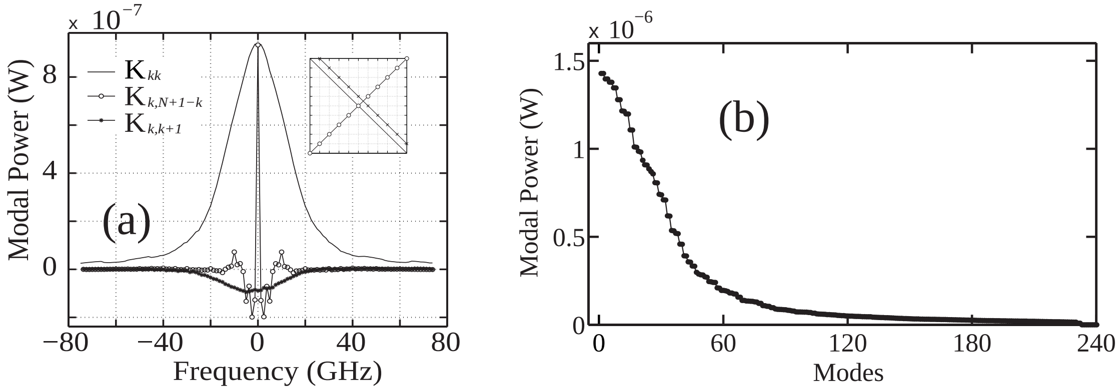

Figure 6.2: DMS system. (a) Three slices through the RR part of the reduced covariance matrix $\mathcal{K}$, indicated by the inset cartoon; (b) Eigenvalue spectrum of $\mathcal{K}$. The eigenvalues are ordered by magnitude.

by decreasing magnitude. The eigenfunctions connected to the spectrum are related to the *optical* Karhunen-Loève modes, namely the modes in which the accumulated noise can be expanded as independent Gaussian noise neglecting the receiver. If most of the eigenvalues were very small, one could make the matrix propagation more efficient by transforming the noise into the optical Karhunen-Loève basis and focusing on the modes with the largest noise power, but as Fig. 6.2(b) shows, the magnitude of the eigenvalues does not fall off very steeply.

Fig. 6.3 compares the average pdf in the marks and spaces, as defined in (4.7.1), of the narrow-band filtered receiver current from the linearization method with the histogram of a standard Monte Carlo simulation. Fig. 6.3(b) shows the corresponding eye diagram as a contour plot of the logarithm of the pdf as a function of time, analogous to Fig. 5.6(b).

6.2 Submarine CRZ System

6.2.1 System Setup

The simulated transmission line of the chirped return-to-zero (CRZ) system is shown in Fig. 6.4. It consists of 34 dispersion map periods each of length 180 km, for a total distance of 6,120 km. Each map period contains a 160 km span of normal dispersion fiber with -2.5 ps/nm-km, indicated by the circles labeled N, preceded and followed by two 10 km spans of anomalous dispersion fiber with 16.5 ps/nm-km, labeled A. Third-order

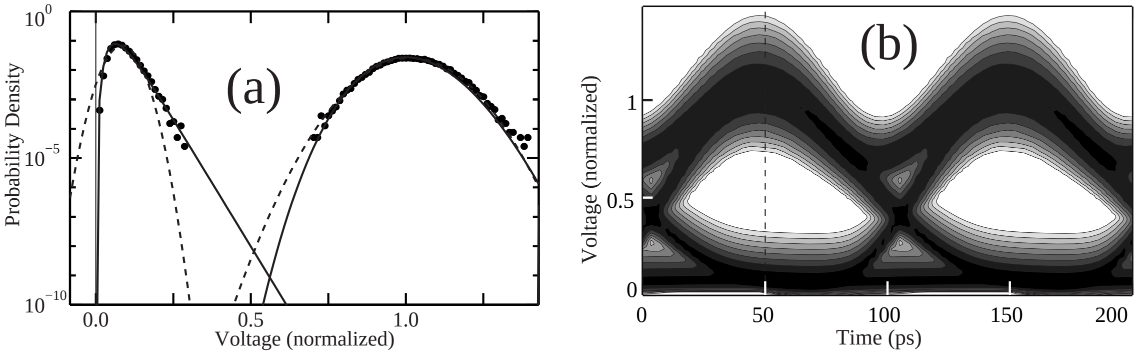

Figure 6.3: DMS system. (a) Histogram from a traditional Monte Carlo simulation (dots) with Gaussian fits of the data points in the marks and spaces based on their mean and variance (dashed lines) and the result of the covariance matrix method (solid line). (b) Eye diagram as a contour plot of the logarithm of the pdf as a function of time. The pdfs in (a) are computed at $t = 50$ ps (dashed line).

dispersion is of minor importance and I neglect it in this simulation, as I did in the DMS system. The fibers have an attenuation of 0.2 dB/km and an effective area of $A_{\text{eff}} = 50\,\mu$m. The loss is compensated every 45 km by an erbium-doped fiber amplifier (EDFA) with a spontaneous emission factor of $n_{\text{sp}} = 2.0$. I use pre- and post-compensating fiber spans, labeled C, each of which has a total dispersion of 916 ps/nm. The signal pulses are co-polarized and have a FWHM duration of 45 ps with a bit-synchronously chirped raised-cosine shape of the form

$$u(t) = \left(\frac{P_{\text{peak}}}{2} \left\{ 1 + \cos\left[\pi \sin\left(\frac{\Omega t}{2} \right) \right] \right\} \right)^{1/2} \exp\left(iA\pi \cos \Omega t \right), \tag{6.2.1}$$

where $u(t)$ is the envelope of the optical field at time t, $\Omega = 2\pi/T_{\text{bit}}$ with the bit spacing $T_{\text{bit}} = 100$ ps, and the chirp parameter is $A = -0.6$ [77]. The initial optical peak power is $P_{\text{peak}} = 1$ mW before entering the pre-compensating fiber. I transmit 32 bits, corresponding to a pseudorandom bit sequence of $2^5 - 1 = 31$ bits, plus an additional zero bit, thereby exhausting all possible bit patterns of length 5, with a Fourier vector length N_{FFT} equals to 4096. At the receiver, I model an ideal square law detector followed by a 5-th order Bessel filter with a one-sided 3-dB bandwidth of 4.3 GHz. Fig. 6.8 shows the narrow-band filtered noise-free receiver current after 6,100 km of transmission.

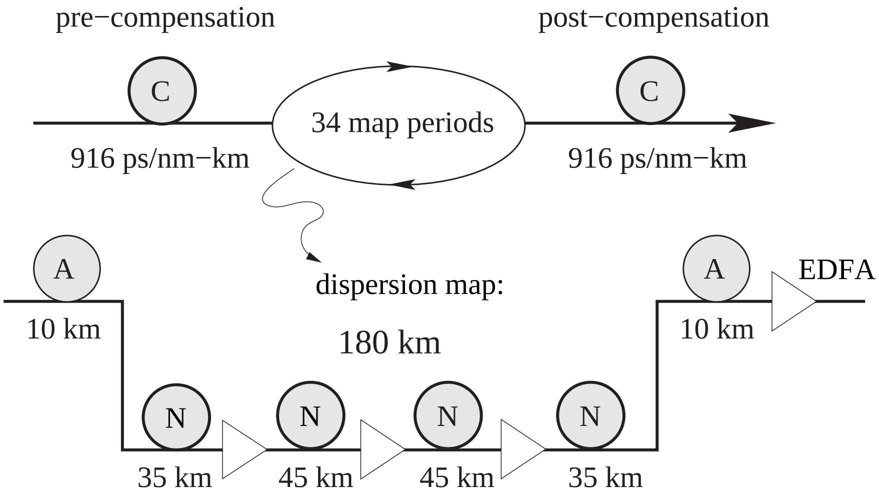

Figure 6.4: CRZ system, schematic illustration.

6.2.2 Results and Discussion

Fig. 6.5 shows the phase-space portraits of the CRZ system analogous to Fig. 5.2, shown in Section 5.1 on p. 52. All graphs are produced by a simulation of a single CRZ pulse centered at $t = 0$ in an 800 ps time window. Fig. 6.5(a) shows the six lines of the pulse portraits at the beginning of each fiber span in the first dispersion map period. It also shows the phase portrait of the pre-compensation. The latter is strongly chirped and has an S-shape, indicated by the arrow. Fig. 6.5(b) is the same as (a), except that I have added four copies of the lines in (a), each offset in time by a multiple of 100 ps. The copies correspond to pulses centered at 100 ps, 200 ps and so on. Note that the phase-space portrait of a signal consisting of five adjacent pulses would look completely different from Fig. 6.5(b), since the phase of the sum of two signals does not equal the sum of the phases. The point of showing multiple copies of the portraits in one graph is to demonstrate that the individual portraits of different pulses at a given transmission distance never overlap. Although pulses overlap in the time domain, they are still separated in phase space. Therefore, the phase space picture contains more information. A vertical line at $t = 200$ ps cuts through the portraits of five pulses, and the portraits of adjacent pulses in the same fiber span are separated by about 10 GHz. As the pulses evolve, their portraits rotate counter-clockwise and contract, as shown in Fig. 6.5(c) for the maps 16–19. In the middle of the total transmission distance near map 17, the pulses are maximally compressed and separated in time, but strongly chirped. After map 17, they expand again and the final state at map 34 in

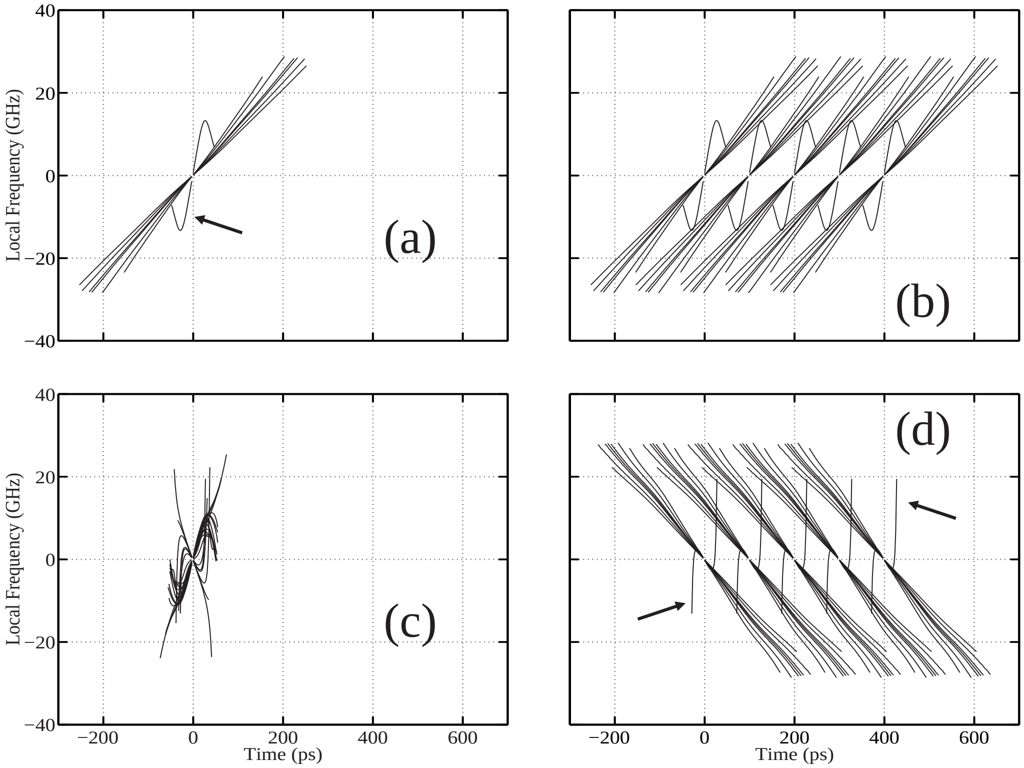

Figure 6.5: Phase space portraits for the CRZ system at the beginning of each fiber span. (a) Arrow: pre-compensation portrait, oblique lines: portrait during the first map period; (b) Same as (a), but five portraits are superimposed with time offsets of 100 ps; (c) Portraits of the maps 16–19; (d) Same as (b), but for the last map. Arrows: post-compensation portrait.

Fig. 6.5(d) has a mirror symmetry with Fig. 6.5(b). The almost vertical curves, indicated by the arrows, correspond to the pulse after post-compensation which is slightly narrower than the launched pulse.

Fig. 6.6(a) shows the optical power $|u_0|^2$ of the noise-free 32-bit signal at the end of the transmission in the time domain. The tic marks indicate the boundaries of the bit slots. Fig. 6.6(b) shows the optical power spectrum at the end of the transmission as the circles. The 10 GHz tones and their harmonics are clearly visible. The dots show the average power spectrum of the noise as obtained from $|a_k|^2 + |a_{k+N}|^2 = \mathcal{K}_{k,k} + \mathcal{K}_{k+N,k+N}$; the number of modes in $\mathcal{K}$ is $N = 140$. The ratio of the noise power at frequency $f = 0$ to the power at $f = 22$ GHz is 2.1 and is due to parametric gain. The OSNR, defined as the ratio of total signal power over total noise power in the shown bandwidth, is 14.83, or 11.7 dB.

Fig. 6.7(a) shows three slices through the RR part of the covariance matrix $\mathcal{K}$ similar

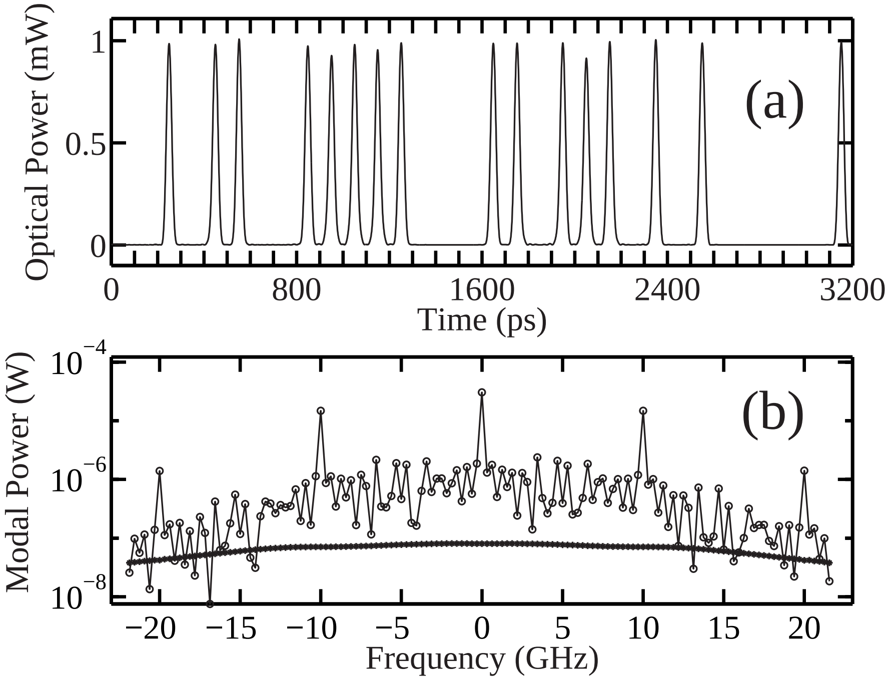

Figure 6.6: CRZ system. (a) Optical power $|u_0|^2$ of the noise-free 32-bit signal at the end of the transmission in the time domain. (b) Dots: optical power spectrum on a logarithmic scale. Asterisks: average power spectrum of the noise from $\mathcal{K}$. The ratio of the noise power at $f = 0$ to the power at $f = 22$ GHz is 2.1. The OSNR is 11.7 dB.

to Fig. 6.2(a), where the inset gives a pictorial representation of a matrix whose elements are located at the grid points. The oblique lines show the location of the three slices. The solid line runs along the principal diagonal, the circles run on the secondary diagonal, and the asterisks show elements that lie on a parallel to the principal diagonal where the cross-correlations are particularly large. As in the DMS system, most cross-correlations are negative.

Fig. 6.7(b) shows the eigenvalue spectrum of $\mathcal{K}$, analogous to Fig. 6.2(b). Due to numerical imprecisions a few eigenvalues are vanishingly small and must be increased before the pdfs are computed, as described in Appendix A.4.

Fig. 6.8 shows the narrow-band filtered noise free receiver current after transmission over 6,120 km. The 32 partial pdfs are computed at the points in time indicated by the dots. The variation in the peak power is due to nonlinear pulse-to-pulse interactions during the transmission, highlighting the importance of bit patterns.

Fig. 6.9(a) shows the average pdfs as defined in (4.7.1) of the receiver current that re-

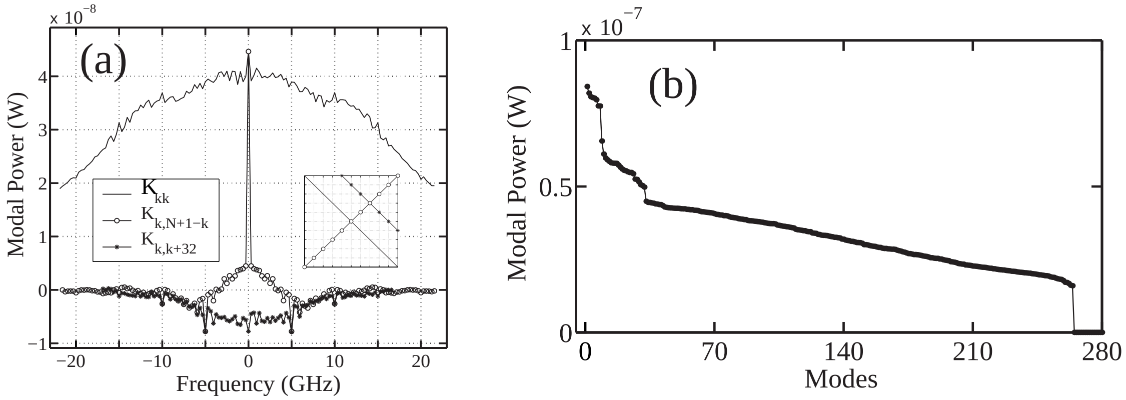

Figure 6.7: CRZ system. (a) Three slices through the RR part of the reduced covariance matrix $\mathcal{K}$, indicated by the inset cartoon; (b) Eigenvalue spectrum of $\mathcal{K}$. The eigenvalues are sorted by magnitude.

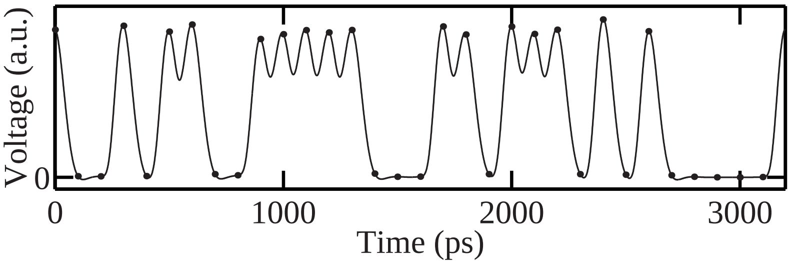

Figure 6.8: CRZ system. Narrow-band filtered noise-free current $y_{\mathrm{nf}}(t)$ of the signal after transmitting 6,120 km. The 32 partial pdfs are computed at the points in time indicated by the dots.

sult from the covariance matrix method as solid lines in comparison with a histogram from a traditional Monte Carlo simulation, consisting of 86,000 noise realizations represented by the dots. The current is normalized to the mean of the pdf of the marks. The dashed lines show a Gaussian fit to the Monte Carlo data, using the mean and variance. The large deviation between the solid and dashed curves is obvious, especially in the spaces. On the other hand, the agreement between the covariance matrix method and the Monte Carlo results is excellent. By integrating the pdfs, one obtains optimal BERs of 1.7×10^{-12} from the covariance matrix method and 9.9×10^{-12} from the Gaussian fit of the Monte Carlo data. The latter corresponds to a Q-factor of 6.71. Note that the relatively small difference between these two BERs occurs because the Gaussian fit overestimates the pdf of the marks and underestimates it in the spaces, and hence the Q-factor method relies on

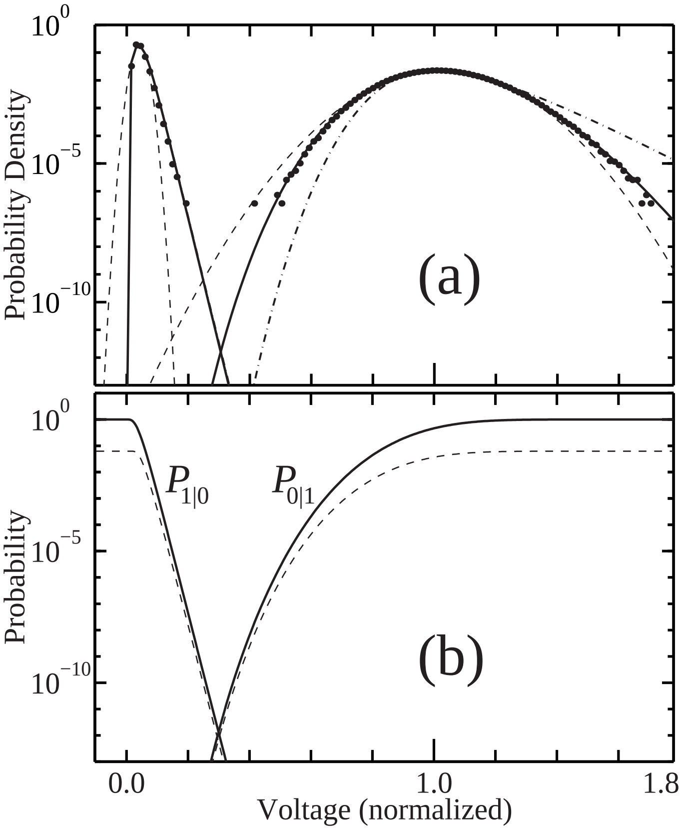

Figure 6.9: CRZ system. (a) Solid lines: average pdfs from the covariance matrix method; dots: histogram from a Monte Carlo simulation; dashed lines: Gaussian fit to the dots using their mean and variance; dash-dotted line: average pdf from the covariance matrix method without phase jitter separation. (b) Solid lines: error probabilities $P_{1|0}$ and $P_{0|1}$ corresponding to the solid lines in (a); dashed lines: $P_{1|0}$ and $P_{0|1}$ for the worst noise-free mark and space only.

the accidental partial cancellation of two errors. The dash-dotted line shows the pdf of the marks that one obtains if the phase jitter is not separated, which is clearly wrong.

The left solid curve in Fig. 6.9(b) shows the error probability $P_{1|0}$ of detecting a “1” when a “0” was sent, using a given decision level as defined in (5.4.1a). This curve corresponds to the left pdf in Fig. 6.9(a). The right solid curve is the probability $P_{0|1}$ of detecting a “0” when a “1” was sent. The dashed lines show the same, except that only the mark with the lowest current *in the noise-free signal* and the space with the highest current are taken into account. The bit errors near the optimum decision level are dominated by the worst mark and space. This result indicates that in the CRZ system it is sufficient to apply the linearization method only to the patterns that exhibit the worst behavior in the

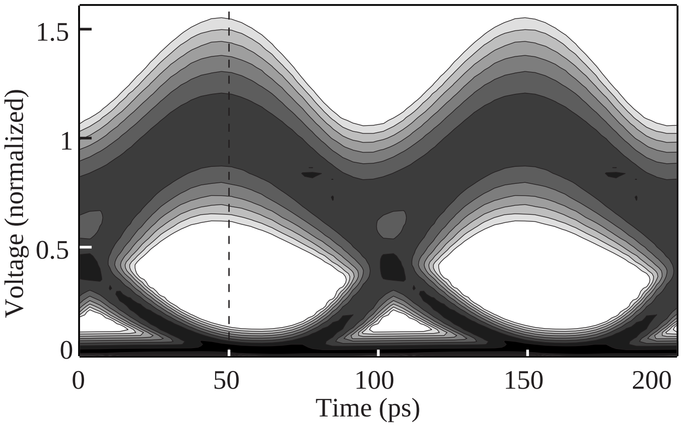

Figure 6.10: CRZ system. A contour plot of the logarithm of the pdf as a function of time showing the eye diagram. The pdfs in Fig. 6.9 were taken at $t = 50$ ps (dashed line).

absence of noise to obtain a good approximation of the average pdfs and the BER.

Fig. 6.10 shows the corresponding eye diagram of all 32 bits. The probability density is displayed as a contour plot. The pdf is only plotted over a range of about four orders of magnitude to make it look like an eye diagram, although I am able to accurately compute the probability density at any point in the diagram. The pdfs in Fig. 6.9 are taken at $t = 50$ ps, indicated by the dashed line.

6.2.3 WDM System

I now discuss the application of the covariance matrix method to a WDM CRZ system. This system is an extension of the single-channel CRZ system that I described in the previous section. We launch identical bit sequences in 5 channels, spaced 50 GHz apart. This channel spacing is narrow, compared to the channel bandwidth of about ±23 GHz, and hence corresponds to a dense wavelength-division multiplexed (DWDM) system. While real transmission systems have many more channels [78], work by Yu, *et al.* [79] shows that it is possible to simulate a dense WDM system with a limited number of channels.

Fig. 6.11 shows the optical power spectrum at the end of the transmission line. The channel spacing is 50 GHz and the channels are indicated by indices running from -2 to $+2$.

In the following, I focus on the central channel (channel 0). The system parameters $N_{\text{FFT}} = 4096$, $N = 140$, and $T = 3200$ ps, as well as all pulse parameters are identical to the single-channel system. The propagation algorithm of the covariance matrix, as shown

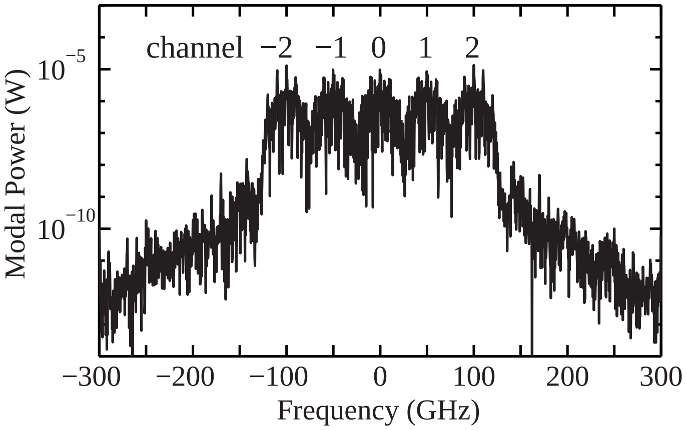

Figure 6.11: CRZ DWDM system with 5 channels. Noise-free optical power spectrum $|A_k|^2$ at the end of the transmission line. The channel spacing is 50 GHz, and the channels are indicated by numbers.

in Appendix A.4, is identical to the single-channel case, except that after step 1, I apply an artificial optical bandpass filter with a square shape and a bandwidth of ±25 GHz that only passes channel 0. The artificial filtering is necessary because the neighboring channels would otherwise disturb the phase jitter separation. However, I simulate the transmission of u_0 with all 5 channels. Thus, I model the signal-noise interactions of the signal in any channel with the noise in the central channel, but I neglect the *inter-channel* noise-noise correlations. In other words, I neglect off-diagonal matrix blocks in an extended covariance matrix that spans multiple channels. This simplification is physically reasonable, since the noise-noise correlations decay with frequency separation, and it is validated by Monte Carlo simulations.

As in the single-channel case, I ran a Monte Carlo simulation to validate my results. Since I expected that the neighboring channels might distort the pdf slightly and the linearization method might yield inaccurate results, I ran 100,000 noise realizations in a simulation that took 51 days to complete.

Fig. 6.12 shows the average pdfs of the receiver current that result from the covariance matrix method as solid lines in comparison with a histogram from a traditional Monte Carlo simulation represented by the dots. The voltage is normalized to the mean of the pdf of the marks. The dashed lines show a Gaussian fit to the Monte Carlo data, using the mean and variance. The agreement between the linearization approach and the Monte Carlo results is excellent. By integrating the pdfs I obtain an optimal BER of 4.7×10^{-12}

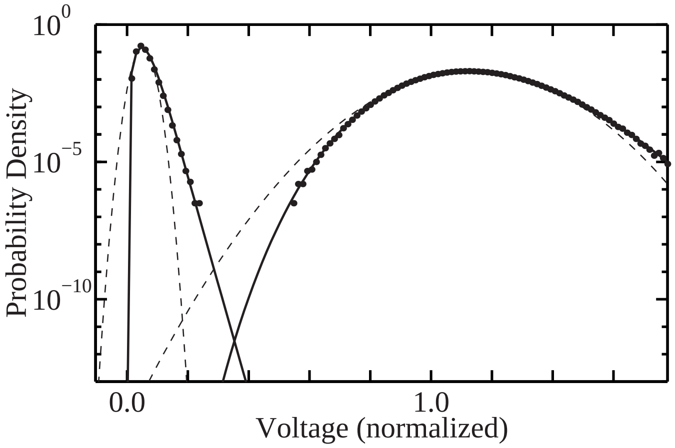

Figure 6.12: CRZ DWDM system. Solid lines: average pdfs from the linearization approach; dots: histogram from a Monte Carlo simulation; dashed lines: Gaussian fit to the dots using the mean and variance.

from the linearization approach as compared to 4.1×10^{-11} from the Gaussian fit of the Monte Carlo data; the latter corresponds to a Q-factor of 6.50.

As an additional test, I record the *inter-channel* regions of the extended covariance matrix, *i.e.*, in addition to averaging over the noise realizations to obtain $\mathcal{K}^{(r)} = \mathcal{K}^{(r)}_{[0,0]}$, I also compute the matrices $\mathcal{K}^{(r)}_{[1,0]} = \mathcal{K}^{(r)}_{[0,1]}$ and $\mathcal{K}^{(r)}_{[1,1]}$, where $\mathcal{K}^{(r)}_{[i,j]}$ refers to the inter-channel covariance matrix defined by a generalization of (4.3.6)

$$\mathcal{K}^{(r)}_{[i,j]} = \left\langle \boldsymbol{r}_{[i]} \boldsymbol{r}^T_{[j]} \right\rangle, \tag{6.2.2}$$

where $\boldsymbol{r}_{[k]} = (r_{-N/2+n,R}, \cdots, r_{N/2-1+n,R}, r_{-N/2+n,I}, \cdots, r_{N/2-1+n,I})^T$. Here, the offset $n = k|\Delta f_{\text{ch}}|/\Delta f_{\text{FFT}}$ is the number of frequency modes between the center of channel k and channel 0 with the channel spacing $|\Delta f_{\text{ch}}| = 50$ GHz, and $\Delta f_{\text{FFT}} = 1/T$ is the frequency resolution of the Fourier transform. I find that $\mathcal{K}^{(r)}_{[0,0]} = \mathcal{K}^{(r)}_{[1,1]}$ to within the computational accuracy set by the split-step algorithm and $\left\|\mathcal{K}^{(r)}_{[1,0]}\right\|_{\text{max}} \Big/ \left\|\mathcal{K}^{(r)}_{[0,0]}\right\|_{\text{max}} = 0.076$, where $\|\cdot\|_{\text{max}}$ is the matrix maximum norm. Inter-channel noise-noise correlations are of little relevance since only the noise in the central channel contributes to the eye diagram.

In summary, the generalization of the covariance matrix to WDM does not seem to pose major difficulties. The simulation time increases relative to a single-channel system, and the size of this increase depends exclusively on the step size in the split-step algorithm when simulating multiple channels.

Chapter 7

Conclusions

Accurately modeling bit error rates is crucial when planning, designing, or upgrading optical transmission systems. Bit errors are primarily caused by ASE noise from the optical amplifiers. This noise is white at the point where it is contributed. However, this ASE noise accumulates along the transmission and interacts nonlinearly with the signal, which influences the statistical distribution of the noise in a complicated way. On the other hand, bit errors are very rare events. This combination of circumstances is the reason why neither simple assumptions about the pdf of the optical noise at the receiver and the pdf of the electrical receiver current, nor traditional Monte Carlo simulations can yield accurate results. I show in this dissertation that the commonly used Q-factor method, based on a Gaussian extrapolation of the receiver current, is an unreliable approach and depends on the fortuitous partial cancellation of two errors.

I present a covariance matrix method that yields the pdf of the optical noise at any transmission distance, which is based on linearization. From the pdf of the optical noise, the accurate pdf of the narrow-band filtered receiver voltage can be computed, yielding the accurate BER and eye diagrams. I developed two different approaches to compute the key statistical quantity, which is the covariance matrix of the optical noise, and I show that both lead to a significant efficiency enhancement over traditional Monte Carlo simulations.

To my knowledge, a complete linearization of the optical transmission in a realistic transmission system has never been carried out before, as discussed in Chapter 3. My colleagues and I started our investigations in the fall of 1998 by checking the distribution of the optical noise at the receiver of a Monte Carlo simulation of a 10 Gb/s DMS system over 24,000 km. This system was very well characterized both in simulation and in experiment [28], [67]–[69]. We focused on the real and imaginary coefficients of the Fourier components of the noise, and we found their pdfs. Doob's theorem [46] states that the noise pdfs of a linear system that is driven by Gaussian noise will remain multivariate-Gaussian distributed. If one expands the noise in a specific basis, for example the Fourier basis, then each Fourier coefficient obeys a marginal distribution of the total noise distribution, and should be Gaussian distributed. Indeed, that proved to be the case for short transmission distances. However, over distances greater than about 2,000 km, we found that these components were no longer Gaussian distributed, indicating the breakdown of the linearization assumption for the Fourier basis. This breakdown is due to the phase and timing jitter in this highly nonlinear system. we find, however, that if we use a new basis set in which the phase and timing jitter are separated from the other noise components, then the linearization assumption remains valid over the entire propagation distance of 24,000 km.

I describe two approaches for calculating the covariance matrix. First, it is possible to run Monte Carlo simulations and average the results to approximate the covariance matrix. In essence, one fits a multivariate-Gaussian to the Monte Carlo simulation result, after separating the phase and timing jitter. This procedure is self-validating and relatively simple to program, but requires on the order of 2000 noise realizations, as shown in Chapter 5. However, this approach makes better use of the simulation data and yields more accurate results than traditional Monte Carlo simulations with the same number of noise realizations.

The second method, referred to in this dissertation as the covariance matrix method, propagates the covariance matrix from transmitter to receiver in a deterministic way. The propagation is accomplished by multiplying the covariance matrix with a propagator matrix $\boldsymbol{\Psi}$ that describes the evolution of a noise vector over a fiber span according to (4.4.2). In Chapter 4, I described a method for numerically computing the matrix $\boldsymbol{\Psi}$ based on the split-step algorithm. My colleagues and I applied the deterministic approach to both the DMS system and a 10 Gb/s single-channel CRZ system. In both cases, the result is efficient and accurate to the degree it can be compared to traditional Monte Carlo simulations, as demonstrated in Chapter 6. The computational cost is roughly $2N$ times that of a single Monte Carlo noise realization, where N is the number of complex Fourier modes stored in the matrix. In a simulation of eight bits in the DMS system, I used $N = 120$ and in the CRZ system with 32 bits I used $N = 140$, and hence this method solves the covariance matrix in a fraction of the time required by the Monte Carlo simulation, discussed in Chapter 5. Moreover, it is intrinsically far more accurate, since the Monte Carlo methods have an intrinsic statistical inaccuracy, while the covariance matrix method, being deterministic, is only limited by the accuracy of the numerical solver.

Modern propagation formats such as CRZ exhibit large pulse overlap, leading to nonlinear interactions of the signal with itself. The phase and timing jitter separation becomes more complicated in this case, but it can still be carried out after applying artificial dispersion compensation [80] as discussed in Section 4.4. Another important result is that the worst bit patterns in the noise-free signal can dominate the BER and hence it is sufficient to apply the linearization approach to these patterns.

In addition to single-channel DMS and CRZ systems, my colleagues and I studied WDM systems. We focused on a WDM system with five channels and a channel spacing of 50 GHz. While five channels is small compared to the number of channels in commercial systems, earlier work shows that it is possible to accurately model WDM systems with a limited number of channels [79]. The question was not if the method worked in principle for a WDM system, but rather what was the scaling behavior of the computational time with the number of channels. I found for the 5-channel system that computing the pdf of a single channel still only requires $2N$ times as much CPU time as one noise realization in a traditional Monte Carlo simulation. This result, combined with the decomposition mentioned in the previous paragraph, bodes well for the application of the noise linearization approach to massive WDM systems with a lot of pulse overlap, such as are currently used in commercial systems.

It is well known that noise-noise beating cannot be neglected in the receiver, in contrast to the transmission. The implementation of accurate receiver models that take this beating into account is critical when calculating the eye diagrams and bit error rates. We have implemented an accurate receiver model for a receiver with a square law photodetector, followed by an electrical filter. In this case, the current obeys a generalized chi-square distribution when phase and timing jitter can be neglected. The phase jitter has no effect on the received current because of the square law detector. By contrast, the timing jitter does affect the distribution, which no longer obeys a generalized chi-square distribution. We found an integral expression for the new distribution, and, in this dissertation, I show how to calculate it.

In conclusion, my colleagues and I achieved the goals 1–4 that I described in the Introduction: We derived a covariance matrix approach that works in two very different transmission systems, one of which is highly nonlinear. We computed the pdf of the optical noise and the narrow-band filtered receiver current, and we obtained an efficient implementation of the algorithm.

What are the limits of the noise linearization approach? Highly nonlinear formats with optical pulse-to-pulse interactions such as 40 Gb/s DMS systems [64] are challenging to a linearization approach. It has been found that the timing jitter in some soliton systems with strong pulse-to-pulse interactions is not Gaussian distributed anymore [81]. If noise-induced jitter and nonlinear pulse-to-pulse interactions are not independent of each other, then the linearization assumption breaks down, although it is still an important starting point for a more complete theory. Another possible challenge is strong ASE noise. In the two systems that we have studied, the optimum BER is below 10^{-9}, but in many commercial systems, raw BERs of 10^{-3} or higher are acceptable with the use of forward error correction (FEC) [1], [6]. However, it is only the *probability* of noise realizations leading to raw bit errors that is increased in systems with FEC, not the current threshold at which these errors occur. Since the covariance matrix method can accurately compute the electrical pdf anywhere in the eye diagram, I do not anticipate problems in dealing with higher raw bit errors.

In the future, the covariance matrix method should be generalized and refined. It might be possible to improve the numerical efficiency of the algorithm by further reducing the number of frequency modes in the covariance matrix or using some interpolation scheme. The agreement between the pdfs from the covariance matrix method should be compared to Monte Carlo simulations all the way down to the intersection point of the marks and the spaces pdfs. The range of the pdfs from Monte Carlo simulations might be extended using the importance sampling technique [10], [11]. Moreover, I have not applied the covariance matrix method to systems in which two optical polarizations and hence PMD are relevant, systems that use Raman amplification, or modulation formats such as non-return-to-zero (NRZ) or differential phase shift keying (DPSK). Finally, a determination of the bit patterns that lead to the smallest eye openings should be investigated, in particular for WDM systems.

Appendix A

Appendix

A.1 Average Phase Definition

If the phase and timing jitter are small, then the averages of the noisy signal A_k and of the renormalized signal $\overline{A}_k$ defined after (4.3.4) will be approximately identical, $A = \langle B \rangle \approx \left\langle \overline{B} \right\rangle = \overline{A}$. In this case, one can decompose the renormalized noise $a_k \exp[-i(\varphi + \omega_k \tau)]$ as

$$\begin{aligned} a_k \exp[-i(\varphi + \omega_k \tau)] &= \overline{A}_k - A_k \exp[-i(\varphi + \omega_k \tau)] + r_k \\ &\approx \overline{A}_k \Big\{ 1 - \exp\big[-i(\varphi + \omega_k \tau)\big] \Big\} + r_k \\ &\approx \varphi c_k + \tau d_k + r_k, \end{aligned} \tag{A.1.1}$$

where

$$c_k = i\overline{A}_k, \tag{A.1.2a}$$

$$d_k = i\omega_k \overline{A}_k. \tag{A.1.2b}$$

The term φc_k is responsible to first order for a phase shift $u(t) = u_0(t)\exp(i\varphi)$, while the component τd_k produces the time shift $u(t) = u_0(t+\tau)$ and thus leads to timing jitter. The residual noise r_k is orthogonal to the c_k and d_k with the scalar product Re $\sum_k v_k r_k^* = 0$, where $v_k = c_k$ or $v_k = d_k$. The vectors c_k and d_k will only be orthogonal to each other if the signal is an even function in time. I note that the decomposition (A.1.1) is similar to (3.2.10).

At the chirp-free maximum pulse compression point in the loop, one may define the average signal phase by (4.3.3c), which can be written in the frequency domain as

$$\alpha_{\text{ave}} = \sum_{k=-N/2}^{N/2-1} |B_k|^2 \arctan\left(\frac{B_{k,I}}{B_{k,R}}\right) \Big/ \sum_{k=-N/2}^{N/2-1} |B_k|^2, \tag{A.1.3}$$

where $B_k = B_{k,R} + iB_{k,I} = A_k + a_k$ are the Fourier modes of $u(t)$. First, one would like to determine the relationship of α_{ave} in (A.1.3) and α in (4.3.5) to φ in (A.1.1), and analogously the relationship of β to τ, for small noise. The Taylor expansion of the signal phase at frequency mode k, using the identity $d\arctan x/dx = (1+x^2)^{-1}$, is

$$\arctan\frac{B_{k,I}}{B_{k,R}} = \alpha_0 + \frac{a_{k,I}A_{k,R} - a_{k,R}A_{k,I}}{|A_k|^2} + O(a_k^2), \tag{A.1.4}$$

where $\alpha_0 = \arctan(A_{k,I}/A_{k,R})$. Using (A.1.4), one may linearize the least-squares criterion (4.3.5) to obtain

$$H = \min_{\alpha,\beta} \sum_{k=-N/2}^{N/2-1} \frac{1}{|A_k|^2}\left[a_{k,I}A_{k,R} - a_{k,R}A_{k,I} - (\delta\alpha + \beta\omega_k)|A_k|^2\right]^2, \tag{A.1.5}$$

where $\delta\alpha = \alpha - \alpha_0$. The stationary points of the sum in (A.1.5) with respect to α and β are given by

$$\frac{\delta H}{\delta\alpha} = -2\sum_k \left[a_{k,I}A_{k,R} - a_{k,R}A_{k,I} - (\delta\alpha + \beta\omega_k)|A_k|^2\right] = 0, \tag{A.1.6a}$$

$$\frac{\delta H}{\delta\beta} = -2\sum_k \omega_k \left[a_{k,I}A_{k,R} - a_{k,R}A_{k,I} - (\delta\alpha + \beta\omega_k)|A_k|^2\right] = 0. \tag{A.1.6b}$$

One may express the orthogonality relations $(\boldsymbol{c},\boldsymbol{r}) = (\boldsymbol{d},\boldsymbol{r}) = 0$ as $\sum_k \left[r_{k,I}A_{k,R} - r_{k,R}A_{k,I}\right] = 0$ and $\sum_k \omega_k \left[r_{k,I}A_{k,R} - r_{k,R}A_{k,I}\right] = 0$. We infer

$$\sum_k \left[a_{k,I}A_{k,R} - a_{k,R}A_{k,I}\right] = \sum_k (\varphi + \tau\omega_k)|A_k|^2, \tag{A.1.7a}$$

$$\sum_k \omega_k \left[a_{k,I}A_{k,R} - a_{k,R}A_{k,I}\right] = \sum_k \omega_k (\varphi + \tau\omega_k)|A_k|^2. \tag{A.1.7b}$$

Comparing the last two identities with (A.1.6a) and (A.1.6b), one concludes that the sums in (A.1.6a) and (A.1.6b) vanish with the choice $\delta\alpha = \varphi$ and $\beta = \tau$. This result shows that the definition of an average phase (A.1.3) is reasonable and consistent with the least-squares fit (4.3.5).

A.2 Gauss-Hermite Integration

We chose a Gauss-Hermite integration technique as described by Abramowitz and Stegun [61] to solve the convolution (4.6.16). This technique approximates the integral

$$\int_{-\infty}^{\infty} g(x)\,\mathrm{exp}(-x^2)\,dx \approx \sum_{i=1}^{n} w_i\,g(x_i), \tag{A.2.1}$$

where $g(x)$ is any function with which the integral in (A.2.1) exists, w_i are weights, and the x_i are (unevenly spaced) abscissas. Both w_i and x_i depend only on n. The values are tabulated and can be determined using numerical routines [82]. One must approximate the integral

$$\begin{aligned} f_y(y,t) &= \int_{-\infty}^{\infty} f_{y,\tau=0}(y,t-\tau')\,f_\tau(\tau')\,d\tau' \\ &= \frac{1}{\sqrt{2\pi}\sigma_t}\int_{-\infty}^{\infty} f_{y,\tau=0}(y,t-\tau')\,\mathrm{exp}\left(-\frac{\tau'^2}{2\sigma_\tau^2}\right)d\tau' \\ &= \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} f_{y,\tau=0}\left(y,\,t-\sqrt{2}\sigma_\tau x\right)\mathrm{exp}\left(-x^2\right)dx \\ &\approx \sum_{i=1}^{n} w_i\,g_i, \end{aligned} \tag{A.2.2}$$

where

$$g_i = \frac{1}{\sqrt{\pi}} f_{y,\tau=0}(y,t_i) \tag{A.2.3a}$$

$$t_i = t - t_{GH,i} = t - \sqrt{2}\sigma_\tau x_i. \tag{A.2.3b}$$

We chose $n=5$. If $f_y(y,t)$ must be computed for many evenly spaced times t, for instance for the use in an eye diagram with a high temporal resolution, it might be faster to solve the convolution (4.6.16) by Fourier transforming f_y and f_τ with respect to t and take advantage of the convolution theorem. The abscissas and weights for $n=5$ are given in the following table:

i	x_i	w_i
1	−2.020	0.01995
2	−0.9586	0.3936
3	0	0.9453
4	0.9586	0.3936
5	2.020	0.01995

A.3 Simultaneous Diagonalization

The theorem of simultaneous diagonalization (see [59, p. 106]) holds for $\mathcal{K}^{-1}$ (real symmetric and positive definite) and $\mathcal{W}$ (symmetric) and states that in this case there is a real $2N \times 2N$ matrix $\mathcal{C}$ with $\det \mathcal{C} \neq 0$ and

$$\begin{aligned}
(\mathcal{C}^{-1})^T \mathcal{K}^{-1} \mathcal{C}^{-1} &= \mathcal{I} \qquad \text{and} && \text{(A.3.1a)}\\
(\mathcal{C}^{-1})^T \mathcal{W} \mathcal{C}^{-1} &= \Lambda = \text{diag}(\lambda_1, \ldots, \lambda_{2N}), && \text{(A.3.1b)}\\
\mathcal{C}^{-1} &= \mathcal{U}\mathcal{D}\mathcal{V} \qquad \text{with} && \text{(A.3.1c)}\\
\mathcal{U}^T \mathcal{K}^{-1} \mathcal{U} &= \text{diag}(\epsilon_1, \ldots, \epsilon_{2N}), \qquad \mathcal{U}^T\mathcal{U} = \text{I}, && \text{(A.3.1d)}\\
\mathcal{D} &= \text{diag}(\epsilon_1^{-1/2}, \ldots, \epsilon_{2N}^{-1/2}), && \text{(A.3.1e)}\\
\mathcal{V}^T \mathcal{H} \mathcal{V} &= \Lambda, \qquad \mathcal{V}^T \mathcal{V} = \mathcal{I}, && \text{(A.3.1f)}\\
\mathcal{H} &= \mathcal{D}^T \mathcal{U}^T \mathcal{W} \mathcal{U} \mathcal{D}, && \text{(A.3.1g)}
\end{aligned}$$

where all the λ_i are real numbers. This diagonalization is equivalent to solving the generalized real eigenvalue problem $\mathcal{W}\mathcal{C}^{-1} = \mathcal{K}^{-1}\mathcal{C}^{-1}\Lambda$. Note that there is no need to ever compute the matrix $\mathcal{C}$ explicitly from $\mathcal{C}^{-1}$. We found that the implementation of (A.3.1a)–(A.3.1g) is preferable to using a general-purpose generalized eigensystem solver because the matrix $\mathcal{U}\mathcal{D}$ only depends on $\mathcal{K}$ and has only to be calculated once, while the matrix $\mathcal{H}$ depends on $\mathcal{W}$ and hence on the time t. In particular when calculating eye diagrams that require many pdfs, this approach yields a computational gain.

A.4 Algorithms

The computation of the electrical pdfs can be broken down into the optical transmission part, yielding the covariance matrix $\mathcal{K}$, the computation of the electrical pdf from $\mathcal{K}$, and additional post-processing such as drawing eye diagrams and computing the optimum BER. Without going into the details of the programming, I just state that the transmission part is implemented as a stand-alone C++ program, and the computation of the BER and any post-processing is implemented as Matlab code. I will now describe three parts of the algorithm, namely

1. The main loop of the transmission simulation routine,
2. the propagation of the covariance matrix, and

3. the computation of the electrical pdf.

The nonlinear propagation equation (4.2.1a) in all simulations in this dissertation is solved with a third-order split-step algorithm [72].

A.4.1 The Main Loop of the Transmission Simulation

The transmission program can be operated in four different modes:

1. Noise-free propagation, yielding $u_0(t)$,
2. Traditional Monte Carlo simulation, yielding noise signal realizations $u(t)$ with options to generate traditional eye diagrams and compute jitter averages,
3. Extended Monte Carlo: same as in mode 2., but with the additional computation of $\mathcal{K}$ by averaging over noise realizations according to (4.3.6),
4. Deterministic simulation: compute $u_0(t)$ and $\mathcal{K}$, using the method described in Section (4.4).

The goal in setting up the simulation this way was to facilitate the comparison of the different pdfs resulting from modes 2–4. The algorithm in the main loop of the simulation can be itemized as follows:

1. Setup: Read in a parameter file that contains all physical and technical parameters, such as the number of bits, the pulse durations, the peak power, the random generator seed, etc. Check the consistency of the parameters. Allocate and initialize all arrays.
2. In Monte Carlo modes 2 and 3, start a loop over different noise realizations:
    (a) Launch the optical signal $u_0(t, 0)$.
    (b) In all modes, start a loop over the dispersion map periods, possibly preceded by a pre-compensating fiber span:
        i. Compute the signal evolutions element by element of the dispersion map such as fiber spans, amplifiers, optical filters. Add ASE noise in amplifiers in the Monte Carlo modes 2 and 3. In mode 4, propagate $\mathcal{K}$ through all dispersion map elements.

        ii. At regular intervals during the transmission, for instance every five dispersion maps, draw eye diagrams, compute various fluctuations such as timing jitter, and accumulate data to compute $\mathcal{K}(z)$ at the local point.

    (c) Write all data to the disk every 30 minutes.

3. Write the final field $u_0(t)$, the matrix $\mathcal{K}$, and all eye diagrams and statistics to the disk.

A.4.2 Deterministic Propagation of the Covariance Matrix

The deterministic propagation of the covariance matrix $\mathcal{K}$ is described in Section (4.4) and is implemented in the C++ routine `propagator()`. At every amplifier, Gaussian white noise power is added to the principal diagonal of $\mathcal{K}$. The action of an optical filter on $\mathcal{K}$ is given by $\mathcal{K}_{\text{out},kl} = \mathcal{K}_{k,l} H_k H_l$, where H_i is the real filter function at frequency mode i.

I now summarize the algorithm for propagating $\mathcal{K}$ over one fiber span from $z = 0$ to $z = L$, followed by an EDFA. First, propagate the field $u_0(t, 0)$ using a standard Fourier split-step algorithm, yielding $u_0(t, L)$. Save $u_0(t, L)$, return to $z = 0$ and repeat the following for each k:

1. Compute $u^{(k)}(t, L)$ by perturbing $u_0(t, 0)$ in the k-th frequency mode and propagating it to $z = L$.

2. Separate the pulses in the signal by passing $u^{(k)}(t, L)$ through a linear and lossless fiber with total dispersion $-D(L)$.

3. Compute the perturbation vector $\boldsymbol{a}^{(k)}$ and apply the phase jitter separation (4.4.5a), (4.4.5b) individually to each pulse, yielding the vector $\widetilde{\boldsymbol{a}}^{(k)}$ of the dispersion compensated signal.

4. Invert step 2 to compute $\widetilde{\boldsymbol{a}}^{(k)}$ at point L. Evaluate the propagator matrix elements $\boldsymbol{\Psi}_{jk}$ for all j.

Finally, compute $\mathcal{K}(L)$ according to (4.4.2).

A.4.3 Computation of the Electrical pdf from the Covariance Matrix

The computation of the electrical pdf from the covariance matrix is implemented in the Matlab routine `noise.m.`

1. Read in $\overline{\boldsymbol{A}}$, the noise-free optical field vector in frequency domain, and $\mathcal{K}$, the covariance matrix at the receiver.

2. To increase the numerical accuracy, scale $\overline{\boldsymbol{A}}$ and $\mathcal{K}$ so that the maximum element of $\mathcal{K}$ becomes 1.

3. As Figs. 6.2(b) and 6.7(b) show, some of the eigenvalues of $\mathcal{K}$ can become very small or even negative during the propagation, due to numerical inaccuracies. These small eigenvalues would inhibit the inversion of $\mathcal{K}$. If there are any eigenvalues of $\mathcal{K}$ that are smaller than $\epsilon = 1/300$ times the maximum eigenvalue, diagonalize $\mathcal{M}^T\mathcal{K}\mathcal{M} = \Lambda_{\mathcal{K}}$ where $\Lambda_{\mathcal{K}}$ is the diagonal matrix of the eigenvalues, increase the small eigenvalues to the threshold value, and compute a corrected covariance matrix $\mathcal{K}' = \mathcal{M}\Lambda'_{\mathcal{K}}\mathcal{M}^T$. I verified that the resulting pdf is independent of ϵ over a wide range.

4. Invert $\mathcal{K}'$.

5. Compute the matrix $\mathcal{U}\mathcal{D}$ by applying (A.3.1d) and (A.3.1e).

6. Compute the self-adjoint matrix $\mathcal{W}_{kl}$ as described in (4.6.3).

7. Start a loop over all points in time $t = t_k$ for which the pdf must be computed:

   (a) If timing jitter was separated and must be added again by applying (A.2.1), start a loop over the n Gauss-Hermite sample times $t_{k,i} = t_k - t_{GH,i}$ with $1 \leq i \leq n$:

      i. Compute $\mathcal{W}(t_{k,i})$ from F.

      ii. Compute the matrix $\mathcal{H}$ by evaluating (A.3.1g), then compute $\mathcal{V}$ and Λ.

      iii. Compute the values $s_l(t_{k,i}) = \lambda_l Q_l^2 = \sum_m \overline{\boldsymbol{A}}_l \mathcal{W}_{lm}(t_{k,i})\overline{\boldsymbol{A}}_m$ and $y_{\text{nf}}(t_{k,i}) = \sum_l s_l$.

      iv. Invert step 2.

      v. Evaluate the characteristic function $\boldsymbol{\Phi}_{\tau=0}(\zeta, t_{k,i})$ for zero timing jitter according to (4.6.12), using s_k.

      vi. Prepare the discrete Fourier transform to compute $f_{y,\tau=0}(y, t_{k,i})$: If $i = 1$, optimize the grid increment $\Delta\zeta_k$ to yield the best resolution in $f_{y,\tau=0}(y, t_{k,i})$ This $\Delta\zeta_k$ is then used for all i.

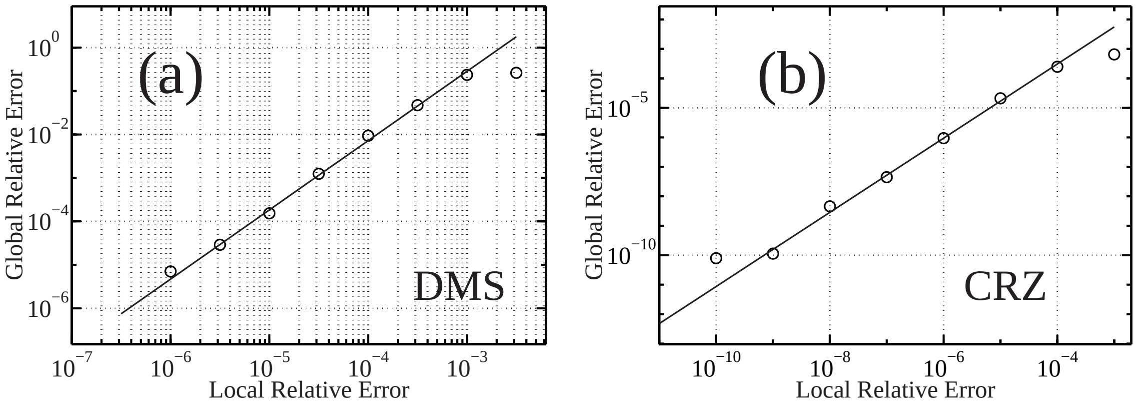

Figure A.1: Simulated values of global versus local relative error $\epsilon(\delta)$ shown by the circles for the (a) DMS system and (b) CRZ system. The lines are fits to a power law.

   vii. Compute $f_{y,\tau=0}(y, t_{k,i})$ by taking the discrete Fourier transform of $\mathbf{\Phi}_{\tau=0}(\zeta, t_{k,i})$.
   viii. Verify that $f_{y,\tau=0}(y, t_{k,i})$ is normalized and $f_{y,\tau=0}(0, t_{k,i}) = 0$, then save it.

  (b) Compute $f_y(y, t_k)$ from the $f_{y,\tau=0}(y, t_{k,i})$ by applying (A.2.2).

  (c) Save $y_{\text{nf}}(t_k)$, $\Delta\zeta_k$, and $f_y(y, t_k)$.

8. Write all $y_{\text{nf}}(t_k)$, $\Delta\zeta_k$, and $f_y(y, t_k)$ to the disk.

It turns out that that the steps 2, 3, and 7(a)vi significantly improve the numerical accuracy of the resulting pdf. The computation can fail altogether if step 3 is omitted. Both steps 2 and 7(a)vi improve $f_y(y, t)$ by lowering the roundoff floor. According to my experience, $f_y(y, t)$ will never span more than about 16 orders of magnitude when using `double` floating point arithmetic (64 bits). I use a Fourier transform vector length of 1024 in step 7(a)vi.

A.5 Split-Step Accuracy

All simulation in this dissertation are based on the third-order split-step algorithm [72]. In order to compare traditional Monte Carlo simulations to the deterministic method, one has to verify that the numerical accuracy is sufficient. We propagate one single Monte Carlo noise realization in both the 8-bit DMS system over 24,000 km as described in Section 6.1 and in the 32-bit CRZ system over 6,100 km as described in Section 6.2, while varying the local relative error bound $\delta = \left\|u_c - u_f\right\| / \left\|u_f\right\|$, where u_c and u_f are the coarse and the fine solutions respectively [72]. The noise is only added in the N lowest frequency modes and the random generator's seed is the same in all simulations.

In Fig. A.1 I show the global relative error $\epsilon = \left\|u_{(\delta)} - u\right\| / \|u\|$ versus the local relative error δ in double-logarithmic scale, where $u_{(\delta)}$ is the optical field at the receiver that was obtained using the error bound δ, and u is the solution for $\delta = 10^{-12}$. The circles in Fig. A.1(a) show simulation values of $\epsilon(\delta)$ for the DMS system, and the line is a fit with the function $q\delta^m$, where $q = 1.737 \times 10^4$ and $m = 1.594$. Fig. A.1(b) displays the same for the CRZ system, where $q = 33.19$ and $m = 1.258$. For a global error goal of $\epsilon \leq 10^{-5}$ that I used in the Monte Carlo simulations, the fit yields the local relative errors of $\delta = 1.6 \times 10^{-6}$ and $\delta = 6.6 \times 10^{-6}$ for the DMS and CRZ systems, respectively. For $\epsilon = 10^{-4}$, one finds $\delta = 6.8 \times 10^{-6}$ and $\delta = 4.0 \times 10^{-5}$.

Bibliography

[1] S. Bigo and W. Idler, *et al.*, “Transmission of 125 WDM channels at 42.7 Gbit/s (5 Tbit/s capacity) over 12×100 km of TeraLight Ultra fibre,” in *Proc. ECOC’01*, postdeadline paper PDM11, (Amsterdam, Netherlands), 2001.

[2] J.-X. Cai, M. Nissov, and A. N. Pilipeskii, *et al.*, “Transmission of thirty-eight 40 Gb/s channels ($>$ 1.5 Tb/s) over transoceanic distance,” in *Proc. OFC’02*, postdeadline paper FC4, (Anaheim, CA), 2002.

[3] E. Desurvire, *Erbium-Doped Fiber Amplifiers: Principles and Applications*. New York, NY: Wiley, 1994.

[4] K. Rottwitt and H. D. Kidorf, “A 92 nm bandwidth Raman amplifier,” in *Proc. OFC’98*, paper PD6, (Baltimore, MD), 1998.

[5] N. S. Bergano, “WDM in long-haul transmission systems,” in *Proc. OFC’00*, short course 108, (Baltimore, MD), 2000. The original quote is “Nonlinearity in optical fibers is always bad, except for the careers of engineers.”.

[6] J.-X. Cai, M. Nissov, and A. N. Pilipeskii, *et al.*, “1.28 Tb/s (32 × 40 Gb/s) transmission over 4,500 km,” in *Proc. ECOC’01*, postdeadline paper PDM12, (Amsterdam, Netherlands), 2001.

[7] N. S. Bergano, F. W. Kerfoot, and C. R. Davidson, “Margin measurements in optical amplifier systems,” *IEEE Photon. Technol. Lett.*, vol. 5, pp. 304–306, 1993.

[8] G. P. Agrawal, *Nonlinear Fiber Optics*. London: Academic Press, 2nd ed., 1995.

[9] N. Metropolis and S. Ulam, “The Monte Carlo method,” *J. Am. Stat. Assoc.*, vol. 44, pp. 335–341, 1949.

[10] P. M. Hahn and M. C. Jeruchim, “Developments in the theory and application of importance sampling,” *IEEE Trans. Commun.*, vol. 35, pp. 706–714, 1987.

[11] G. Biondini, W. L. Kath, and C. R. Menyuk, “Importance sampling for polarization-mode dispersion,” *IEEE Photon. Technol. Lett.*, vol. 14, no. 3, pp. 310–312, 2002.

[12] R. Hui, D. Chowdhury, M. Newhouse, M. O’Sullivan, and M. Poettcker, “Nonlinear amplification of noise in fibers with dispersion and its impact in optically amplified systems,” *IEEE Photon. Technol. Lett.*, vol. 9, pp. 392–394, 1997.

[13] R. Hui, M. O’Sullivan, A. Robinson, and M. Taylor, “Modulation instability and its impact in multispan optical amplified IMDD systems: Theory and experiments,” *J. Lightwave Technol.*, vol. 15, pp. 1071–1081, 1997.

[14] A. N. Pilipetskii, V. J. Mazurczyk, and C. J. Chen, “The effect of dispersion compensation on system performance when nonlinearities are important,” *IEEE Photon. Technol. Lett.*, vol. 11, pp. 284–286, 1999.

[15] C. E. Shannon, “A mathematical theory of communication,” *Bell System Technical Journal*, vol. 27, pp. 379–423 and 623–656, 1948.

[16] E. E. Narimanov and P. Mitra, “The channel capacity of a fiber optics communication system: perturbation theory,” *J. Lightwave Technol.*, vol. 20, pp. 530–537, 2002.

[17] J. P. Gordon and H. A. Haus, “Random walk of coherently amplified solitons in optical fiber transmission,” *Opt. Lett.*, vol. 11, pp. 665–667, 1986.

[18] D. J. Kaup, “Perturbation theory for solitons in optical fibers,” *Phys. Rev. A*, vol. 42, pp. 5689–5694, 1990.

[19] H. A. Haus, “Quantum noise in solitonlike repeater systems,” *J. Opt. Soc. Am. B*, vol. 8, pp. 1122–1126, 1991.

[20] T. Georges, “Perturbation theory for the assessment of soliton transmission control,” *Opt. Fiber Technol.*, vol. 1, pp. 97–116, 1995.

[21] G. Bosco, A. Carena, V. Curri, R. Gaudino, P. Poggiolini, and S. Benedetto, “A novel analytical method for the BER evaluation in optical systems affected by parametric gain,” *IEEE Photon. Technol. Lett.*, vol. 12, pp. 152–154, 2000.

[22] A. Carena, V. Curri, R. Gaudino, P. Poggiolini, and S. Benedetto, "New analytical results on fiber parametric gain and its effects on ASE noise," *IEEE Photon. Technol. Lett.*, vol. 9, pp. 535–537, 1997.

[23] V. S. Grigoryan, C. R. Menyuk, and R.-M. Mu, "Calculation of timing and amplitude jitter in dispersion-managed optical fiber communications using linearization," *J. Lightwave Technol.*, vol. 17, pp. 1347–1356, 1999.

[24] A. Messiah, *Quantum Mechanics*, vol. I and II. Amsterdam: North Holland Publ., 1964.

[25] H. Heffner, "The fundamental limit of linear amplifiers," *Proc. IRE*, vol. 50, pp. 1604–1608, 1962.

[26] W. S. Wong, H. A. Haus, L. A. Jiang, P. B. Hansen, and M. Margalit, "Photon statistics of amplified spontaneous emission noise in a 10-Gbit/s optically preamplified direct-detection receiver," *Opt. Lett.*, vol. 23, pp. 1832–1834, 1998.

[27] C. R. Menyuk, "Application of multiple-length-scale methods to the study of optical fiber fiber transmission," *J. Eng. Math.*, vol. 36, pp. 113–136, 1999.

[28] R.-M. Mu, V. S. Grigoryan, C. R. Menyuk, G. M. Carter, and J. M. Jacob, "Comparison of theory and experiment for dispersion-managed solitons in a recirculating fiber loop," *IEEE J. Sel. Topics Quant. Electron.*, vol. 6, pp. 248–257, 2000.

[29] R.-M. Mu, T. Yu, V. S. Grigoryan, and C. R. Menyuk, "Convergence of the CRZ and DMS formats in WDM systems using dispersion management," in *Proc. OFC'00*, paper FC1, (Baltimore, MD), 2000.

[30] V. E. Zakharov and A. B. Shabat, "Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media," *Sov. Phys.-JETP*, vol. 34, pp. 62–69, 1972.

[31] C. R. Menyuk, G. M. Carter, W. L. Kath, and R.-M. Mu, "Dispersion managed solitons & chirped return to zero: What is the difference?," in *Optical Fiber Telecommunications* (I. P. Kaminow and T. Li, eds.), vol. IVB, ch. 7, pp. 305–328, San Diego, CA: Academic, 2002.

[32] L. F. Mollenauer, R. H. Stolen, and J. P. Gordon, “Experimental observation of picosecond pulses narrowing and solitons in optical fibers,” *Phys. Rev. Lett.*, vol. 45, pp. 1095–1098, 1980.

[33] N. J. Smith, N. J. Doran, F. M. Knox, and W. Forysiak, “Energy-scaling characteristics of solitons in strongly dispersion-managed fibers,” *Opt. Lett.*, vol. 21, pp. 1981–1983, 1996.

[34] N. J. Smith, N. J. Forysiak, and W. Doran, “Reduced Gordon-Haus jitter due to enhanced power solitons in strongly dispersion managed systems,” *Electron. Lett.*, vol. 32, pp. 2085–2086, 1996.

[35] C. D. Poole and J. Nagel, “Polarization effects in lightwave systems,” in *Optical Fiber Telecommunications* (I. P. Kaminow and T. L. Koch, eds.), vol. IIIA, ch. 6, pp. 114–161, San Diego, CA: Academic, 1997.

[36] N. S. Bergano, C. R. Davidson, and F. Heismann, “Bit-synchronous polarization and phase modulation scheme for improving the performance of optical amplifier transmission systems,” *Electron. Lett.*, vol. 32, pp. 52–54, 1996.

[37] G. M. Carter and J. M. Jacob, “Dynamics of solitons in filtered dispersion-managed systems,” *IEEE Photon. Technol. Lett.*, vol. 10, pp. 546–548, 1998.

[38] E. Iannone, F. Matera, A. Mecozzi, and M. Settembre, *Nonlinear Optical Communication Networks*. Microwave and Optical Engineering, New York: Wiley, 1998.

[39] L. E. Reichl, *A Modern Course in Statistical Physics*. Austin: Univ. of Texas Press, 1980.

[40] M. Yu, G. P. Agrawal, and C. J. McKinstrie, “Pump-wave effects on the propagation of noisy signals in nonlinear dispersive media,” *J. Opt. Soc. Am. B*, vol. 12, pp. 1126–1132, 1995.

[41] M. Midrio, “Statistical properties of noisy propagation in normally dispersive fibers,” *J. Opt. Soc. Am. B*, vol. 14, pp. 2910–2914, 1997.

[42] A. Yariv, *Optical Electronics in Modern Communcations*. New York: Oxford University Press, 1997.

[43] A. Mecozzi, "Long-distance transmission at zero dispersion: Combined effect of the Kerr nonlinearity and the noise of the in-line amplifiers," *J. Opt. Soc. Am. B*, vol. 11, pp. 462–469, 1994.

[44] L. Arnold, *Stochastic Differential Equations: Theory and Applications*. Malabar, FL: Krieger Publishing Company, 1992. Reprinted from New York: Wiley, 1974.

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

[46] J. L. Doob, "The Brownian movement and stochastic equations," in *Selected Papers on Noise and Stochastic Processes* (N. Wax, ed.), New York: Dover, 1954.

[47] R. Holzlöhner, V. S. Grigoryan, C. R. Menyuk, and W. L. Kath, "Accurate calculation of eye diagrams and bit error rates in long-haul transmission systems using linearization," *J. Lightwave Technol.*, vol. 20, pp. 389–400, 2002.

[48] D. Marcuse, "Derivation of analytical expressions for the bit-error probability in lightwave systems with optical amplifiers," *J. Lightwave Technol.*, vol. 8, pp. 1816–1823, 1990.

[49] L. Dieci and T. Eirola, "Positive definitenness in the numerical solution of Riccati differential equations," *Num. Math.*, vol. 67, pp. 303–313, 1994.

[50] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1995.

[51] H. Risken, *The Fokker-Planck Equation*. Berlin: Springer, 2nd ed., 1989.

[52] Y. Kodama, M. Romagnoli, and S. Wabnitz, "Chirped nonlinear pulse propagation in a dispersion-compensated system," *Opt. Lett.*, vol. 22, pp. 1689–1691, 1997.

[53] G. Bennetin, L. Galgani, and J. M. Strelcyn, "Kolgomorov entropy and numerical experiments," *Phys. Rev. A*, vol. 14, pp. 2338–2345, 1976.

[54] G. Strang, *Linear Algebra and Its Applications*. New York, NY: Academic Publishing, 1980.

[55] R.-M. Mu and C. R. Menyuk, "Symmetric slope compensation in a long-haul WDM system using the CRZ format," *IEEE Photon. Technol. Lett.*, vol. 13, pp. 797–799, 2001.

[56] S. D. Cohen and A. C. Hindmarsh, *CVODE User Guide*. Lawrence Livermore National Laboratory, 1994.

[57] J.-S. Lee and C.-S. Shim, "Bit-error-rate analysis of optically preamplified receivers using an eigenfunction expansion method in optical frequency domain," *J. Lightwave Technol.*, vol. 12, pp. 1224–1229, 1994.

[58] E. Forestieri, "Evaluating the error probability in lightwave systems with chromatic dispersion, arbitrary pulse shape and pre- and postdetection filtering," *J. Lightwave Technol.*, vol. 18, pp. 1493–1503, 2000.

[59] J. Franklin, *Matrix Theory*. New Jersey: Prentice-Hall, 1968.

[60] L. Schwartz, *Théorie des Distributions I*. Paris: Hermann, 1966.

[61] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*. New York: Dover, 1965.

[62] P. V. Mamyshev and N. A. Mamysheva, "Pulse-overlapped dispersion-managed data transmission and intrachannel four-wave mixing," *Opt. Lett.*, vol. 24, pp. 1454–1456, 1999.

[63] M. J. Ablowitz and T. Hirooka, "Resonant nonlinear intrachannel interactions in strongly dispersion-managed systems," *Opt. Lett.*, vol. 25, pp. 1750–1752, 2000.

[64] R. Holzlöhner, H. N. Ereifej, G. M. Carter, and C. R. Menyuk, "Experimental and theoretical characterization of a 40 Gb/s long-haul single-channel transmission system," *J. Lightwave Technol.*, vol. 20, pp. 1124–1131, 2002.

[65] C. J. Anderson and J. A. Lyle, "Technique for evaluating system performance using Q in numerical simulations exhibiting intersymbol interference," *Electron. Lett.*, vol. 30, pp. 71–72, 1994.

[66] M. C. Jeruchim, P. Balaban, and K. S. Shanmugan, *Modeling, Methodology and Techniques*. Simulation of Communication Systems, San Diego: Kluwer Academic/Plenum Publishers, 2000.

[67] E. A. Golovchenko, J. M. Jacob, A. N. Pillipetskii, C. R. Menyuk, and G. M. Carter, "Dispersion-managed solitons in a fiber loop with in-line filtering," *Opt. Lett.*, vol. 22, pp. 289–291, 1997.

[68] G. M. Carter, J. M. Jacob, C. R. Menyuk, E. A. Golovchenko, and A. N. Pillipetskii, “Timing-jitter reduction for a dispersion-managed soliton system: experimental evidence,” *Opt. Lett.*, vol. 22, pp. 513–515, 1997.

[69] T. Yu, R.-M. Mu, V. S. Grigoryan, and C. R. Menyuk, “Energy enhancement of dispersion-managed solitons in optical fiber transmission systems with lumped amplifiers,” *IEEE Photon. Technol. Lett.*, vol. 11, pp. 75–77, 1999.

[70] F. Le Guen, S. Del Burgo, M. L. Moulinard, D. Grot, M. Henry, F. Favre, and T. Georges, “Narrow band 1.02 Tbit/s (51×20 Gbit/s) soliton DWDM transmission over 1000 km of standard fiber with 100 km amplifier spans,” in *Proc. OFC’99*, paper PD4, (Washington, DC), 1999.

[71] D. E. Knuth, *Seminumerical Algorithms*. The Art of Computer Programming, Reading: Addison Wesley, 3rd ed., 1998.

[72] O. V. Sinkin, R. Holzlöhner, J. Zweck, and C. R. Curtis, “Optimization of the split-step Fourier method in modeling optical fiber communications systems,” *J. Lightwave Technol.*, vol. 21, pp. 61–68, 2003.

[73] K. W. DeLong, R. Trebino, and D. J. Kane, “Ultrashort pulse frequency-resolved optical gating traces for three common beam geometries,” *J. Opt. Soc. Am. B*, vol. 11, pp. 1595–1608, 1994. http://www.physics.gatech.edu/frog/.

[74] B. Marks, *Dispersion Management in Fiber Optic Communications*. PhD thesis, Northwestern University, Chicago, IL, 2000.

[75] H. Stark and J. E. Woods, *Probability, Random Processes, and Estimation Theory for Engineers*. New Jersey: Prentice-Hall, 1994.

[76] R. Holzlöhner, C. R. Menyuk, V. S. Grigoryan, and W. L. Kath, “Accurate calculation of eye diagrams and error rates in long-haul transmission systems,” in *Proc. OFC’01*, talk MF3, (Anaheim, CA), 2001.

[77] T.-S. Yang, W. L. Kath, and S. G. Evangelides Jr., “Optimal prechirping for dispersion-managed transmission of return-to-zero pulses,” in *Proc. OFC’99*, paper ThQ4, (Washington, D.C.), 1999.

[78] M. Vaa, B. Bakhshi, E. A. Golovchenko, Y. Chai, F. Heismann, H. Li, M. Arend, W. W. Patterson, A. L. Simons, G. T. Harvey, R. L. Maybach, and N. Bergano, "Demonstration of a 640 Gbit/s $\times$ 7000 km submarine transmission system technology ready for field deployment," in *Proc. OFC'01*, paper WF5, (Anaheim, CA), 2001.

[79] T. Yu, W. M. Reimer, V. S. Grigoryan, and C. R. Menyuk, "A mean field approach for simulating wavelength-division multiplexed systems," *IEEE Photon. Technol. Lett.*, vol. 12, pp. 443–445, 2000.

[80] R. Holzlöhner, C. R. Menyuk, W. L. Kath, and V. S. Grigoryan, "Efficient and accurate calculation of eye diagrams and bit-error rates in a single-channel CRZ system," *IEEE Photon. Technol. Lett.*, vol. 14, pp. 1079–1081, 2002.

[81] C. R. Menyuk, "Non-Gaussian corrections to the Gordon-Haus distribution resulting from soliton interactions," *Opt. Lett.*, vol. 20, pp. 285–287, 1995.

[82] W. H. Press and W. T. Vetterling, *et al.*, *Numerical Recipes in C.* New York: Cambridge University Press, 2nd ed., 1988.

Index