Simulation Methods for the Temporal and Frequency Dynamics of Optical Communication Systems

by

Michael Andrew Reimer

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Physics

Waterloo, Ontario, Canada, 2012

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

I examine two methods for modeling the temporal dynamics of optical communication networks that rapidly and accurately simulate the statistics of unlikely but physically significant system configurations. First, I implement a fiber emulator based upon a random uniform walk over the Poincaré sphere that reproduces the expected polarization temporal autocorrelation statistics with a small number of emulator sections. While easy to implement numerically, the increased computational efficiency afforded by this approach allow simulations of the PMD temporal dynamics to be preferentially biased towards regions of low probability using standard multicanonical methods for the first time. Then, in a subsequent study, I present a general transition matrix formalism that additionally applies to other time-dependent communication systems. I compare the numerical accuracy of several transition matrix sampling techniques and show that straightforward modifications of the acceptance rule can significantly increase computational efficiency if the numerical parameters are chosen to ensure a small self-transition probability within each discretized histogram bin. The general applicability of the transition matrix method is then demonstrated by calculating the outage dynamics associated with the hinge model of polarization evolution and, separately, fading in wireless communication channels.

Further, I develop a Magnus expansion formalism for the rapid and accurate estimation of the frequency dynamics of optical polarization that extends the work of Ref. [94] to systems with PMD and PDL. My approach reproduces the power-series expansion and differential equation solution techniques of previous authors while also preserving the required symmetries of the exact solution in every expansion order. This significantly improves the bandwidth of high estimation accuracy, making this method well-suited to the stochastic analysis of PMD and PDL induced system penalty while also yielding physically realizable operator expansions applicable to the joint compensation of PMD and PDL.

Finally, I employ high-speed polarimetery to demonstrate experimentally that low-amplitude mechanical excitations of commercially available dispersion compensation modules can excite high-frequency, $> 75 \times 10^3$ rotations/s, polarization transients that are

nearly invariant between successive measurements. I extend this procedure to measurements of the transient evolution of PMD.

Acknowledgments

I would like to thank my supervisor, Prof. David Yevick, for the guidance and support that made this work possible, and Dr. Maurice O'Sullivan of Ciena Inc. for many helpful discussions. Financial support for this thesis was provided by Ciena Inc., the Natural Sciences and Engineering Research Council of Canada (NSERC), the Centre for Electrophotonic Materials and Devices (CEMD) and the Ontario Research and Development Challenge Fund (ORDCF).

Dedication

For Kerry and Hope.

Contents

List of Figures

List of Symbols

Symbol	Definition	Page
$\mathrm{E}\{\cdots\}$	Ensemble average	26
$\mathbf{I}_N$	$N \times N$ identity matrix	14
$\mathbf{R}$	Matrices are shown in bold font	14
R_{nm}	n, m:th element of the matrix $\mathbf{R}$	15
$\boldsymbol{\sigma}_n$	n:th Pauli spin matrix, $n = 1, 2, 3$	13
$\vec{\boldsymbol{\sigma}}$	Pauli spin vector	14
$\mathbf{R}^{\mathrm{T}}$	Transpose of $\mathbf{R}$	7
$\mathbf{R}^*$	Complex conjugate of $\mathbf{R}$	8
$\mathbf{R}^\dagger$	Hermitian conjugate of $\mathbf{R}$	8
δ_{nm}	Kronecker delta function	15
$\|s\rangle$	2×1 complex polarization Jones vector	7
$\langle s\|$	1×2 Hermitian conjugate of $\|s\rangle$	8
$\vec{s} = \langle s\|\vec{\boldsymbol{\sigma}}\|s\rangle$	3×1 Stokes vector corresponding to $\|s\rangle$	14
$\|\vec{s}\|$	Vector norm of $\vec{s}$	14
$\hat{s}$	3×1 unit vector	13
$\vec{s}\times$	3×3 skew-symmetric cross product matrix	16
$\vec{W}$	3×1 complex principal state vector	42
$\vec{\Omega}$	3×1 polarization mode dispersion vector	20

Glossary

Item	Definition	Page
DCM	Dispersion compensation module	77
DGD	Differential group delay	2
PDG	Polarization dependent gain	34
PDL	Polarization dependent loss	1
PMD	Polarization mode dispersion	1
PSP	Principal state of polarization	20
SOP	State of polarization	7
pdf	Probability density function	4
rms	Root mean square	26

Chapter 1

Introduction

An overview of polarization mode dispersion (PMD) and polarization dependent loss (PDL) in optical communication systems is presented in Section 1.1 while Section 1.2 discusses the material contributions of this thesis. The remaining Sections 1.3-1.5 review the linear algebra of polarization optics and present a fiber emulator model for the stochastic analysis of PMD and PDL.

1.1 PMD in optical communication systems

Optical communication fibers are often deployed in uncontrolled environments, buried beside highways or rail lines, mounted to bridge crossings or strung in aerial cables. The fiber is continuously subjected to the time-varying random mechanical forces of wind, vibrations from passing vehicles and trains or from the periodic thermal variations of daytime heating and central-office air conditioning. These applied external forces together with intrinsic fiber stress tend to modulate the refractive index of the silica glass through the elasto- and thermo-optic effects [12], resulting in slight variations in the propagation constant between

two preferential optical axes aligned with and perpendicular to the principal direction of the mechanical perturbation. To first order, polarized light at the input of the fiber is resolved along each optical axis, or polarization mode, with light coupled to each mode propagating with a different group velocity. For a fiber of length L, the delay

$$\tau = \frac{L}{v_g^{slow}} - \frac{L}{v_g^{fast}} \tag{1.1}$$

between polarization modes [1,12,42,47] is termed the differential group delay (DGD) of the optical fiber, with differing group velocities v_g along each of the orthogonal *fast* and *slow* optical polarization axes, respectively. This effect, termed *polarization mode dispersion* (PMD), varies randomly in time with both the DGD and orientation of the fast and slow axes driven by the stochastic nature of the fiber's immediate physical environment.

At the low ≤ 2.5 Gb/s data rates common in the early 1990s however the transmission impairments due to PMD were typically negligible, with the notable exception of long undersea applications in which cable lengths often exceed 2000 km [9]. Towards the mid 1990s, PMD was recognized as a severe system impairment limiting all-optical transmission distances to < 1000 km and increasing network deployment costs through a requirement for a larger number of expensive electrical repeaters as optical systems operating at 10 Gb/s began widespread deployment. Recent work has even identified PMD as an important mechanism for the decoherence of polarization-entangled photon pairs transmitted over fiber-based quantum communication systems [17,77]. To date more than 1500 articles have been written on PMD, its stochastic properties and strategies for its optical or electronic compensation, providing some indication of its relative importance to the optical engineering and research community.

PMD is particularly problematic in high bit rate intensity modulated systems based upon the direct detection of the received signal power. Here, the relative propagation delay between principal polarization axes and associated broadening of the received optical pulse can significantly increase the probability of detection error and, accordingly,

significant research effort has been devoted to the active tracking and compensation of PMD [49,68,71,74,120]. Compensating PMD in the optical domain however requires hardware that must adaptively track high speed polarization transients occurring on millisecond time scales. Polarization transients in aerial fiber for example can occur at frequencies ranging from several hundred hertz to kilohertz [18, 113] while, as I will demonstrate in later chapters, mechanically induced transients can exceed 10^5 rotations of the Poincaré sphere per second. Optically based PMD compensators operating at these rates have proven to be both complex and expensive to construct. Electronic compensation of the received signal on the other hand offers the high speed response required for transient polarization control, however, the optical phase information required for complete compensation of the DGD is lost as a result of intensity detection, reducing the efficiency of many electronic compensation schemes [85]. Currently, no commercially viable PMD compensation solutions have found widespread adoption in direct detection systems.

Recently there has been significant renewed interest in transmission systems based upon coherent heterodyne detection, facilitated by the introduction of electronic analog to digital converters (ADC) operating at sampling rates exceeding 20 GSamples/s. Since in a coherent lightwave system the received signal is proportional to the optical field and not its intensity, recent research [35, 84, 100] has focused on the implementation of complex digital modulation formats, such as quadrature phase shift keying (QPSK), that exploit both the phase and amplitude of the propagating electric field to increase the transmit symbol rate, while maintaining a spectral width equivalent to a lower baud rate intensity modulated signal. Multiplexing the transmission onto each of the fiber's two orthogonal polarization modes further doubles the spectral efficiency. In this approach, linear transmission impairments such as chromatic dispersion and PMD are effectively compensated electronically using parallel banks of linear tap-delay filters operating on the optical field following coherent detection. Several authors have in fact demonstrated error free transmission over several thousands of kilometers of fiber with mean DGD values exceeding 50

ps [83]. Though less sensitive to PMD induced pulse distortion, the receive electronics must actively adapt to rapid changes in the polarization state to ensure correct polarization demultiplexing during high speed transient events, increasing overall system complexity and cost [82].

The temporal dynamics of strongly mode coupled fiber can be subdivided into two broad regions of polarization activity in which (1) time-dependent effects are assumed to arise from a small number of locations in the fiber subjected to time varying physical environment changes or (2) variations of the fiber's physical environment act uniformly along its length. The so-called "hinge-model" of polarization activity of case (1) is motivated by long-term measurements of PMD in some systems [4, 5] that indicate the presence of long, stable sections of fiber whose birefringence varies on the time scale of several hours to days modulated by rapidly varying, localized regions of polarization activity. Conversely, in case (2) measurements of aerial and buried fiber links [18, 61] often exhibit a stochastic variation of the output polarization and, by extension, the PMD, that is well approximated by a diffusion of the local fiber birefringence over the Poincaré sphere [60]. While these temporal characteristics have been previously measured in a variety of systems and devices [4, 61, 129], the numerical modeling of polarization evolution has been mainly limited to qualitative analyses based upon fiber emulators [57], approximate methods involving a random walk over the DGD magnitude and its frequency derivative [129], or semi-analytic approaches based upon the longitudinal evolution of PMD in a continuum limit [3, 80].

In communication system simulations, biased Monte-Carlo methods such as the multicanonical or importance sampling techniques have been widely employed to determine static quantities such as the probability distribution functions (pdfs) of the DGD or of the system penalty [8, 74, 123]. To model the temporal dynamics of an optical system, however, the distribution of the times required to transition between two groups of states with different physical properties must be calculated. In these cases, as I will demonstrate in subsequent chapters, the multicanonical algorithm can be reformulated as a "transition

matrix" method [34, 110, 115], to adapt such calculations efficiently towards physically unlikely but practically important configurations. This method is applied to the evaluation of the relative probability for transitions between any two states of an optical system over a single time step and, from the resulting formalism, I will compute the pdf of the time duration of outages in an optical system affected by PMD. This method can extract information from the low-probability regions of the pdf far more efficiently than conventional Monte-Carlo techniques.

1.2 Contributions of this thesis

For long lengths of transmission fiber, both the orientation of the fast and slow polarization axes and the DGD vary stochastically with optical wavelength. In this case, the first-order model of PMD as a temporal delay between two orthogonal polarization modes only holds over a finite and limited range of optical frequencies. Consequently, several authors [36, 38, 51, 64] have developed models relating the frequency variation of the Jones polarization transfer matrix to Taylor orders of the PMD about the optical carrier frequency. Here, I will advance the Magnus expansion formalism introduced previously within Refs. [94–98] by analyzing the frequency variation of both PMD and polarization dependent loss (PDL) in optical networks. I will demonstrate that the Magnus formalism both improves the frequency interval of high PMD estimation accuracy relative to previously published works [30,36,51], and provides for the first time a direct inverse mapping between arbitrary Taylor orders of PMD and PDL and the Jones/Mueller polarization matrix. The resulting formalism has since been adapted by other groups [25] to the experimental determination of higher order PMD effects in optical fiber while also finding interest in the wider research community [11]. In addition, I will present an exact recursive formula for arbitrary orders of the frequency dependent PMD and PDL for a finite section fiber emulator that eliminates the need for direct, numerical differentiation of the Jones or Mueller matrix. The reduction

in computation time afforded by this approach makes it well suited to statistical analyses of PMD and PDL induced pulse distortion in optical networks [52].

While the temporal characteristics of the state of polarization and PMD has been previously measured in a variety of systems and devices [4, 61, 129] the numerical modeling of polarization evolution has been limited to qualitative numerical analyses based upon fiber emulators [57], approximate methods involving a random walk over the DGD magnitude and its frequency derivative [129], or semi-analytic approaches based upon PMD in the continuum limit [3, 80]. These methods have typically been difficult to implement numerically, while the former two only approximate the expected PMD temporal autocorrelation statistics. In this thesis, I will develop a fiber emulator model for the temporal evolution of PMD based upon a random walk over the unit sphere that accurately simulates the temporal dynamics of PMD in the distributed limit with only a few emulator sections, potentially providing a significant increase in computational efficiency relative to previously published results. This model, while straightforward to implement, further allows a complete analytic characterization of the autocorrelation properties of optical fiber emulator. Finally, I will present a novel modification of the fiber emulator that allows the PMD temporal dynamics to be biased towards low probability but physically significant regions of polarization activity using the standard multicanonical algorithm.

Transition matrix algorithms have received significant recent attention due to their improved statistical qualities as compared to standard Markov chain Monte-Carlo methods [34, 105, 115]. In this thesis, I adapt the basic transition matrix procedure to the temporal dynamics of PMD by associating each transition in the Markov chain with a simulated temporal step, leading to the accurate reproduction of the temporal dynamics of PMD induced system outages using orders of magnitude fewer samples as compared to traditional methods. While similar in context to the time-dependent PMD emulator introduced previously, the transition matrix method is more generally applicable to other physical systems whose global observable, $\vec{E}$, evolves according to a Markov walk over its

internal state space. In later sections I will demonstrate the wider applicability of the approach by applying the transition matrix method to the estimation of the duration of fading events in wireless communication systems [117]. I will then discuss a novel modification of the transition matrix technique that can significantly improve the efficiency of the sampling of low probability regions as compared to standard numerical simulations based upon the multicanonical algorithm. This can be particularly advantageous in experimental contexts, as in Refs. [75, 76], where the data acquisition time limits the number of samples that can be practically measured, or in system simulations where each determination of the system output observable may require time intensive calculations [94, 112].

1.3 Linear algebra of optical polarization

This section provides an overview of the Pauli spin vector description of electric field polarization, using the notation adopted in Refs. [22, 40, 47], relevant for the analysis of PMD and PDL in subsequent sections. More detailed discussions can be found in Refs. [22, 47, 94] or in quantum mechanics texts dealing specifically with spin-1/2 systems.

1.3.1 Jones vectors in bra-ket notation

In weakly guiding optical fiber the state of polarization (SOP) of a propagating electric field can be decomposed into a linear combination of two orthogonal polarization basis vectors, typically the horizontal, $|x\rangle = [1, 0]^{\mathrm{T}}$, and vertical, $|y\rangle = [0, 1]^{\mathrm{T}}$, linearly polarized states. An arbitrary linear combination of these two basis vectors results in a general elliptical SOP described by a complex 2×1 Jones vector denoted using the "ket" notation

$|A\rangle$ commonly encountered in polarization literature [21, 22] where

$$|A\rangle = \begin{bmatrix} A_x \\ A_y \end{bmatrix} = E_0\, \mathrm{e}^{i\delta} \begin{bmatrix} \cos(\theta/2)\, \mathrm{e}^{+i\phi/2} \\ \sin(\theta/2)\, \mathrm{e}^{-i\phi/2} \end{bmatrix}. \tag{1.2}$$

Here, the signs of the complex phase and factors of two are chosen such that the angles θ and ϕ specify the orientation of the state of the polarization on the Poincaré sphere, cf. Section 1.3.2, while δ represents an overall phase common to both polarization modes. E_0 is a real valued quantity that is often normalized such that E_0^2 represents the power of the optical field.

Associated with the Jones vector $|A\rangle$ is the 1×2 "bra" vector $|A\rangle^\dagger$ which is written in bra-ket notation as the symbol

$$\langle A| = |A\rangle^\dagger = \begin{bmatrix} A_x^*, & A_y^* \end{bmatrix} = E_0\, \mathrm{e}^{-i\delta} \begin{bmatrix} \cos(\theta/2)\, \mathrm{e}^{-i\phi/2}, & \sin(\theta/2)\, \mathrm{e}^{+i\phi/2} \end{bmatrix}, \tag{1.3}$$

in which $\dagger$ represents hermitian conjugation. It follows that the magnitude squared of $|A\rangle$, $\langle A|A\rangle = |A_x|^2 + |A_y|^2 = E_0^2$, is the power of the optical field. The inner-product of two Jones vectors $|A\rangle$ and $|B\rangle$ is written as $\langle A|B\rangle$ with $\langle A|B\rangle = \langle B|A\rangle^*$ while $|A\rangle$ and $|B\rangle$ are orthogonal if $\langle A|B\rangle = 0$.

1.3.2 Stokes vectors and the Poincaré sphere

Aside from an ambiguity in the phase common to both polarization modes, seven measurements are required to determine the Jones vector $|A\rangle$, although the number can be

reduced to four using a suitable linear combination of the results [22]. The first of these determines the total optical intensity while the remaining six measure the optical power following variously oriented polarizers. The goal of this section is to introduce the measurements defining the state of polarization, and relate the measurement results to a convenient three-dimensional "Stokes vector" representation. Apart from being fundamental to the analytic treatment of PMD and PDL developed in later sections, the Stokes vector description of polarization is the basis of operation for most commercially available polarimeters and proves to be extremely useful for visualizing the state of polarization.

For simplicity, the following discussion will be limited to the case of a totally polarized electric field produced for example by the coherent output of a laser. Generalizing the Stokes vector description to include partially polarized fields such as those produced by incoherent broadband sources (LEDs), white light, or amplified spontaneous emission, etc., instead involves a four-dimensional view of polarization that will be deferred to Chapter 2. The measurement procedure outlined in the following discussion however applies to both totally- and partially-polarized electric fields.

The first measurement is that of the total optical power

$$\begin{aligned} P_{tot} &= \langle A|A\rangle \\ &= \langle A| \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} |A\rangle \\ &= E_0^2, \end{aligned} \tag{1.4}$$

where the identity matrix was inserted to motivate the subsequent discussion. The next two measurements determine the optical power following linear polarizers oriented at 0 and 90 deg, i.e. horizontal and vertical polarizations, which are specified in bra-ket notation

by the inner-products

$$P_0 = \langle A| \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} |A\rangle$$
$$P_{90} = \langle A| \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} |A\rangle. \tag{1.5}$$

The first Stokes parameter, A_1, is defined as the difference,

$$\begin{aligned} A_1 &= P_0 - P_{90} \\ &= \langle A| \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} |A\rangle \\ &= E_0^2 \cos\theta, \end{aligned} \tag{1.6}$$

between the two power measurements, where the last line of Eq. (1.6) was obtained by inserting the explicit component form of $|A\rangle$ from Eq. (1.2). The optical power is then measured following linear polarizers oriented at 45 and 135 deg,

$$P_{45} = \frac{1}{2}\langle A| \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} |A\rangle$$
$$P_{135} = \frac{1}{2}\langle A| \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} |A\rangle, \tag{1.7}$$

with the second Stokes parameter defined as

$$\begin{aligned} A_2 &= P_{45} - P_{135} \\ &= \langle A| \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} |A\rangle \\ &= E_0^2 \sin\theta\cos\phi. \end{aligned} \tag{1.8}$$

The final two measurements determine the "circularity" of the polarization state by observing the power following right- and left- circular polarizers,

$$P_{\text{right}} = \frac{1}{2}\langle A| \begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix} |A\rangle$$
$$P_{\text{left}} = \frac{1}{2}\langle A| \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix} |A\rangle, \tag{1.9}$$

with the third Stokes parameter

$$\begin{aligned} A_3 &= P_{\text{right}} - P_{\text{left}} \\ &= \langle A| \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} |A\rangle \\ &= E_0^2 \sin\theta \sin\phi. \end{aligned} \tag{1.10}$$

Consequently, combining Eqs. (1.6), (1.8), (1.10) the angles

$$\begin{aligned} \phi &= \text{atan}\left(\frac{A_3}{A_2}\right) \\ \theta &= \text{atan}\left(\sqrt{\left(\frac{A_2}{A_1}\right)^2 + \left(\frac{A_3}{A_1}\right)^2}\right) \end{aligned} \tag{1.11}$$

specifying the orientation of the Jones vector in Eq. (1.2) are found to be completely determined through simple measurements involving only the optical power following left- and right-circular polarizers and linear polarizers aligned at 0, 45, 90, and 135 deg. Notice that this measurement procedure is invariant with regards to the overall phase, δ, common to both polarization modes.

Collecting the individual Stokes parameters into a 3×1 Stokes vector $\vec{A} = [A_1, A_2, A_3]^{\text{T}} =$

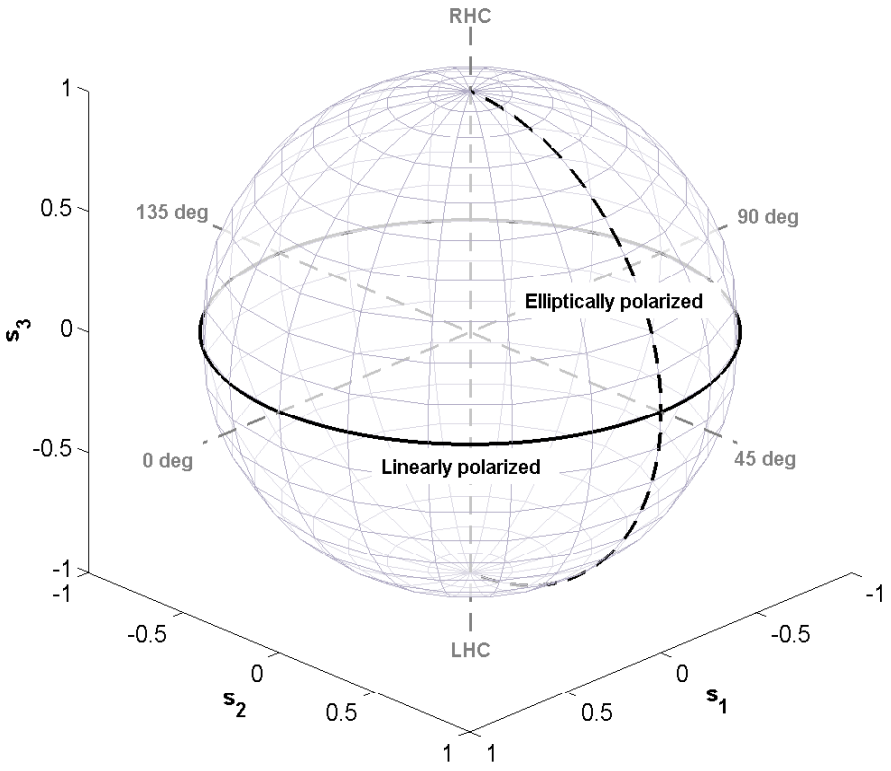


Figure 1.1: An illustration of the Poincaré sphere representation of optical polarization. The Jones vectors corresponding to 0 deg, 45 deg and right hand circular (RHC) polarization states map to the Stokes vectors $s_1 = [1, 0, 0]^{\mathrm{T}}$, $s_2 = [0, 1, 0]^{\mathrm{T}}$ and $s_3 = [0, 0, 1]^{\mathrm{T}}$, forming an orthonormal basis of the Poincaré sphere. Linearly polarized states correspond to points on the equator of the sphere, while points off the equator represent general states of elliptical polarization.

$E_0^2\,\hat{A}$ results in a unit vector

$$\hat{A} = \begin{bmatrix} \cos\theta \\ \sin\theta\cos\phi \\ \sin\theta\sin\phi \end{bmatrix} \tag{1.12}$$

that resides on the surface of the *Poincaré sphere* with coordinates θ and ϕ, respectively. As illustrated in Fig. 1.1, the Jones vectors corresponding to horizontal, $|A\rangle = [1,0]^{\mathrm{T}}$, 45 deg, $|A\rangle = [1,1]^{\mathrm{T}}/\sqrt{2}$, and right hand circular (RHC), $|A\rangle = [1,i\,]^{\mathrm{T}}/\sqrt{2}$, polarization states map to the Stokes vectors $s_1 = [1,0,0]^{\mathrm{T}}$, $s_2 = [0,1,0]^{\mathrm{T}}$ and $s_3 = [0,0,1]^{\mathrm{T}}$, respectively, and form an orthogonal basis for the Poincaré sphere. Notice the factor of two relating polarization angles in Jones and Stokes space. This implies, for example, that orthogonal vectors in Jones space map to antipodal Stokes space vectors, as in the vertically polarized (90 deg) Jones vector $|A\rangle = [0,1]^{\mathrm{T}}$ mapping to the negative s_1 Stokes space axis $[-1,0,0]^{\mathrm{T}}$. In general, the equatorial plane of the Poincaré sphere represents linearly polarized fields while points off the equator correspond to general states of elliptical polarization.

The matrix representations of the three Stokes parameters $A_1 = P_0 - P_{90}$, $A_2 = P_{45} - P_{135}$ and $A_3 = P_{right} - P_{left}$ are further simplified by introducing the *Pauli-spin matrices*,

$$\boldsymbol{\sigma}_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad \boldsymbol{\sigma}_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \boldsymbol{\sigma}_3 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \tag{1.13}$$

so that $A_1 = \langle A|\boldsymbol{\sigma}_1|A\rangle$, $A_2 = \langle A|\boldsymbol{\sigma}_2|A\rangle$ and $A_3 = \langle A|\boldsymbol{\sigma}_3|A\rangle$ (note that throughout this thesis I will refer to "Pauli spin matrix" and "Pauli matrix" interchangeably). The three-component Stokes vector $\vec{A}$ is then

$$\vec{A} = \begin{bmatrix} \langle A|\boldsymbol{\sigma}_1|A\rangle \\ \langle A|\boldsymbol{\sigma}_2|A\rangle \\ \langle A|\boldsymbol{\sigma}_3|A\rangle \end{bmatrix} = \langle A| \begin{bmatrix} \boldsymbol{\sigma}_1 \\ \boldsymbol{\sigma}_2 \\ \boldsymbol{\sigma}_3 \end{bmatrix} |A\rangle = \langle A|\vec{\boldsymbol{\sigma}}|A\rangle, \tag{1.14}$$

where the quantity $\vec{\boldsymbol{\sigma}} = [\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, \boldsymbol{\sigma}_3]^{\mathrm{T}}$ is referred to as the *Pauli-spin vector*. The relationship $\vec{A} = \langle A|\vec{\boldsymbol{\sigma}}|A\rangle$ linking the Jones and Stokes polarization vectors will be used extensively throughout this thesis. In practical calculations, it is typically more convenient to suppress the explicit component form of Eq. (1.2) and instead refer to the coordinate-free representations, $|A\rangle$ and $\vec{A}$, of the Jones and Stokes polarization vectors.

A large number of useful matrix identities are summarized in Refs. [22, 47, 94], each of which can be established directly from the definition of the Pauli matrices, $\boldsymbol{\sigma}_n$, and the Jones and Stokes vectors $|A\rangle$ and $\vec{A}$. A subset of these identities will be used frequently throughout this thesis and are listed here for clarity.

A 2×2 hermitian matrix can be formed by taking the dot product of a 3×1 Stokes vector $\vec{A}$ with the Pauli spin vector $\vec{\boldsymbol{\sigma}}$. This matrix is denoted by the symbol $\vec{A} \cdot \vec{\boldsymbol{\sigma}}$ and is calculated according to

$$\begin{aligned} \vec{A} \cdot \vec{\boldsymbol{\sigma}} &= A_1\boldsymbol{\sigma}_1 + A_2\boldsymbol{\sigma}_2 + A_3\boldsymbol{\sigma}_3 \\ &= \begin{bmatrix} A_1 & A_2 - iA_3 \\ A_2 + iA_3 & -A_1 \end{bmatrix}. \end{aligned} \tag{1.15}$$

It follows that $\left(\vec{A} \cdot \vec{\boldsymbol{\sigma}}\right)^2 = |\vec{A}|^2\,\mathbf{I}_2$, where $\mathbf{I}_2$ is the 2×2 identity matrix. For a unit-length Stokes vector $\hat{p}$, $\left(\hat{p} \cdot \vec{\boldsymbol{\sigma}}\right)^2 = \mathbf{I}_2$, while the eigenvalues of $\hat{p} \cdot \vec{\boldsymbol{\sigma}}$ are ± 1 with corresponding eigenvectors $|p\rangle$ and $|q\rangle$ where $\hat{p} = \langle p|\vec{\boldsymbol{\sigma}}|p\rangle$ and $|q\rangle$ is any Jones vector orthogonal to $|p\rangle$; that is, $\left(\hat{p} \cdot \vec{\boldsymbol{\sigma}}\right)|p\rangle = |p\rangle$ and $\left(\hat{p} \cdot \vec{\boldsymbol{\sigma}}\right)|q\rangle = -|q\rangle$ whenever $\langle p|q\rangle = 0$.

The matrix exponential

$$\begin{aligned} \mathbf{U} &= \exp\left(-\frac{i}{2}\psi\,\hat{p} \cdot \vec{\boldsymbol{\sigma}}\right) \\ &= \cos\left(\frac{\psi}{2}\right)\mathbf{I}_2 - i\sin\left(\frac{\psi}{2}\right)\hat{p} \cdot \vec{\boldsymbol{\sigma}} \end{aligned} \tag{1.16}$$

will be used extensively to describe the rotational transformation of an incident state of polarization in Jones space [22, 47, 94]. Since $\mathbf{U}^\dagger = \exp\left(+i\psi\,\hat{p} \cdot \vec{\boldsymbol{\sigma}}/2\right)$ and $\hat{p} \cdot \vec{\boldsymbol{\sigma}}$ necessarily

commutes with itself, $\mathbf{U}^\dagger\mathbf{U} = \mathbf{I}_2$ and the matrix $\mathbf{U}$ is unitary. In particular, when applied to an incident Jones vector $|A_{in}\rangle$, it can be shown that the output Stokes vector $\vec{A}_{out} = \langle A_{out}|\vec{\boldsymbol{\sigma}}|A_{out}\rangle$ with $|A_{out}\rangle = \mathbf{U}|A_{in}\rangle$ precesses about $\hat{p}$ by the angle ψ on the Poincaré sphere.

Next, I will derive an important relationship between rotation matrices acting on Jones and Stokes polarization vectors. Suppose that the unitary Jones matrix $\mathbf{U} = \exp\left(-i\psi\,\hat{p}\cdot\vec{\boldsymbol{\sigma}}/2\right)$ is applied to an incident Jones vector $|A_{in}\rangle$ yielding $|A_{out}\rangle = \mathbf{U}|A_{in}\rangle$ with corresponding input and output Stokes vectors $\vec{A}_{in} = \langle A_{in}|\vec{\boldsymbol{\sigma}}|A_{in}\rangle$ and $\vec{A}_{out} = \langle A_{out}|\vec{\boldsymbol{\sigma}}|A_{out}\rangle = \langle A_{in}|\mathbf{U}^\dagger\vec{\boldsymbol{\sigma}}\mathbf{U}|A_{in}\rangle$. Since the matrices $\mathbf{U}^\dagger\vec{\boldsymbol{\sigma}}\mathbf{U}$ in this last expression are hermitian and have zero trace they can be decomposed as in Appendix A into a linear combination of the Pauli matrices,

$$\begin{aligned}
\mathbf{U}^\dagger\boldsymbol{\sigma}_1\mathbf{U} &= R_{11}\boldsymbol{\sigma}_1 + R_{12}\boldsymbol{\sigma}_2 + R_{13}\boldsymbol{\sigma}_3 \\
\mathbf{U}^\dagger\boldsymbol{\sigma}_2\mathbf{U} &= R_{21}\boldsymbol{\sigma}_1 + R_{22}\boldsymbol{\sigma}_2 + R_{23}\boldsymbol{\sigma}_3 \\
\mathbf{U}^\dagger\boldsymbol{\sigma}_3\mathbf{U} &= R_{31}\boldsymbol{\sigma}_1 + R_{32}\boldsymbol{\sigma}_2 + R_{33}\boldsymbol{\sigma}_3,
\end{aligned} \tag{1.17}$$

for real valued coefficients R_{ab}, $a, b = 1, 2, 3$. This last expression, which can be rewritten in matrix form as $\mathbf{R}\vec{\boldsymbol{\sigma}} = \mathbf{U}^\dagger\vec{\boldsymbol{\sigma}}\mathbf{U}$, defines a real valued 3×3 matrix $\mathbf{R}$ with components R_{ab} that acts upon the input Stokes vector $\vec{A}_{in}$ according to

$$\begin{aligned}
\vec{A}_{out} &= \langle A_{out}|\vec{\boldsymbol{\sigma}}|A_{out}\rangle \\
&= \langle A_{in}|\mathbf{U}^\dagger\vec{\boldsymbol{\sigma}}\mathbf{U}|A_{in}\rangle \\
&= \langle A_{in}|\mathbf{R}\vec{\boldsymbol{\sigma}}|A_{in}\rangle \\
&= \mathbf{R}\langle A_{in}|\vec{\boldsymbol{\sigma}}|A_{in}\rangle \\
&= \mathbf{R}\vec{A}_{in}.
\end{aligned} \tag{1.18}$$

That is, if $|A_{out}\rangle = \mathbf{U}|A_{in}\rangle$ then $\vec{A}_{out} = \mathbf{R}\vec{A}_{in}$. Since $\text{Trace}\left\{\boldsymbol{\sigma}_a\boldsymbol{\sigma}_b\right\} = 2\delta_{ab}$, where δ_{ab} is the

Kronecker delta function, each element R_{ab} of $\mathbf{R}$ can be calculated from $\mathbf{U}$ according to

$$R_{ab} = \frac{1}{2}\text{Trace}\left\{\mathbf{U}^{\dagger}\boldsymbol{\sigma}_a\mathbf{U}\boldsymbol{\sigma}_b\right\}. \tag{1.19}$$

This last equation provides a useful analytical and numerical method for the conversion between Jones and Stokes matrices that will be generalized in the next chapter to arbitrary optical systems. Further, for $\mathbf{U} = \exp\left(-i\psi\,\hat{p}\cdot\vec{\boldsymbol{\sigma}}/2\right)$ and $\hat{p} = [p_1, p_2, p_3]^{\mathrm{T}}$, we have

$$\begin{aligned}\mathbf{R} &= \exp\left(\psi\,\hat{p}\times\right) \\ &= \cos\psi\,\mathbf{I}_3 + (1-\cos\psi)\,\hat{p}\hat{p}^{\mathrm{T}} + \sin\psi\,\hat{p}\times,\end{aligned} \tag{1.20}$$

where $\mathbf{R}$ is an orthogonal 3×3 matrix and

$$\hat{p}\times = \begin{bmatrix} 0 & -p_3 & p_2 \\ p_3 & 0 & -p_1 \\ -p_2 & p_1 & 0 \end{bmatrix} \tag{1.21}$$

is defined such that $\left(\hat{p}\times\right)\vec{r} = \hat{p}\times\vec{r}$ is the vector cross product for any $\vec{r}$. Here, I have derived Eq. (1.20) by expanding the matrix exponential in its power series and applying $\left(\hat{p}\times\right)^2 = \hat{p}\hat{p}^{\mathrm{T}} - \mathbf{I}_3$ together with the matrix product identities of Appendix A. Summarizing this relationship, the Jones and Stokes matrices $\mathbf{U}$ and $\mathbf{R}$ satisfy

$$\mathbf{U} = \exp\left(-\frac{i}{2}\psi\,\hat{p}\cdot\vec{\boldsymbol{\sigma}}\right) \iff \mathbf{R} = \exp\left(\psi\,\hat{p}\times\right), \tag{1.22}$$

where $\mathbf{U}$ and $\mathbf{R}$ act to rotate the incident polarization vectors $|A_{in}\rangle$ and $\vec{A}_{in}$ by the angle ψ about the unit vector $\hat{p}$ on the Poincaré sphere [94].

The following identities involving similarity transformations of the matrix $\vec{A}\cdot\vec{\boldsymbol{\sigma}}$ will be used to analyze the accumulation of PMD and PDL in optical networks and can be derived using the results of Appendix A. If $\mathbf{U} = \exp\left(-i\psi\,\hat{p}\cdot\vec{\boldsymbol{\sigma}}/2\right)$ and $\mathbf{R} = \exp(\psi\,\hat{p}\times)$ it follows that

$$\mathbf{U}\left(\vec{A}\cdot\vec{\boldsymbol{\sigma}}\right)\mathbf{U}^{\dagger} = \left(\mathbf{R}\vec{A}\right)\cdot\vec{\boldsymbol{\sigma}}. \tag{1.23}$$

Analogously, a similarity transformation of the unitary matrix $\exp\left(-i\theta\,\hat{r}\cdot\vec{\boldsymbol{\sigma}}/2\right)$,

$$\mathbf{U}\exp\left(-\frac{i}{2}\theta\,\hat{r}\cdot\vec{\boldsymbol{\sigma}}\right)\mathbf{U}^{\dagger}=\exp\left(-\frac{i}{2}\theta\left(\mathbf{R}\hat{r}\right)\cdot\vec{\boldsymbol{\sigma}}\right), \tag{1.24}$$

rotates the axis $\hat{r}$ while leaving the eigenvalues $\exp(\pm i\theta/2)$ unaffected. As I will show later in Chapter 2, the above identities can be modified in a straightforward manner to describe optical systems with significant polarization dependent loss (PDL) in which case the Jones matrix $\mathbf{U}$ is no longer unitary.

1.4 Introduction to PMD in optical fiber

Optical fibers or devices are often broadly characterized according to the ratio of their physical length, L, to the birefringence correlation length, L_{corr}, associated with the characteristic stochastic fluctuations of the underlying birefringence, with $L_{\text{corr}} \approx 10$ to 100 m typical for standard single mode fiber [114]. In the "weak mode coupling" limit $L/L_{\text{corr}} \ll 1$, any intrinsic and extrinsic fiber stress or geometrical imperfections responsible for the fiber birefringence are approximately uniform along the fiber and the DGD increases linearly with the fiber length. Conversely, in the "strong mode coupling" limit $L/L_{\text{corr}} \gg 1$ typical of telecom transmission fiber the local orientation of the fiber's fast and slow polarization modes vary rapidly and randomly along the fiber length, leading to a rapid evolution of the state of polarization on the Poincaré sphere with increasing propagation distance. Somewhat surprisingly, in the strong mode coupling limit it was observed [29, 37, 90] that there exist two orthogonally polarized modes at the output of the fiber whose orientation on the Poincaré sphere is invariant with respect to small frequency deviations, $\Delta\omega$, relative to the optical carrier, ω_0. These two orthogonally polarized modes correspond to the instantaneous fast and slow axes of the transmission fiber.

An emulator model of optical transmission fiber is often developed by subdividing the fiber's underlying stochastic birefringence [54, 114] into a large number, N, of randomly

oriented linearly birefringent optical elements. As $N \to \infty$ the emulator statistics converge to those expected for a transmission fiber much longer than the characteristic birefringence correlation length [114]. For finite N, the method provides a convenient numerical algorithm for evaluating the frequency dependence of the SOP and the PMD vector in optical networks. Although the statistical properties of the emulator become independent of the particular stochastic model for each section as $N \to \infty$, I will show that for finite N the statistics of each emulator section can significantly affect both the probability distribution and autocorrelation of the emulated PMD.

The time-domain view of PMD as a DGD between two orthogonal polarization modes has an equivalent frequency-domain representation in which PMD acts to rotate an incident SOP on the Poincaré sphere with increasing optical frequency. For example, adjusting the wavelength of a broadband tunable laser at the input of a fiber or optical device with PMD causes the output SOP to rotate on the Poincaré sphere by an angle proportional to the DGD. This frequency domain representation of PMD is the basis for most PMD system measurements [50, 119], and will be used extensively in what follows to quantify the polarization frequency evolution in a system composed of a concatenation of many linearly birefringent optical devices. To establish the correspondence between the time- and frequency-domain representations, suppose that τ represents the DGD of the m:th emulator section while the fast and slow polarization axes are labeled $|p\rangle$ and $|q\rangle$, respectively, where $|q\rangle$ is orthogonal to $|p\rangle$. The time-domain Jones matrix $\mathbf{U}_m(t)$ implementing the DGD of a single emulator section can be written as

$$\mathbf{U}_m(t) = \delta\left(t - \frac{\tau}{2}\right)\ |p\rangle\langle p| + \delta\left(t + \frac{\tau}{2}\right)\ |q\rangle\langle q|, \tag{1.25}$$

where $\delta(t)$ is the Dirac delta function and the matrices $|p\rangle\langle p|$ and $|q\rangle\langle q|$ act to project an arbitrary incident SOP onto each of the fast and slow optical axes. Transforming into the

frequency domain,

$$\mathbf{U}_m(\omega) = \exp\left(-\frac{i\tau\omega}{2}\right)|p\rangle\langle p| + \exp\left(\frac{i\tau\omega}{2}\right)|q\rangle\langle q|, \tag{1.26}$$

and applying the identities of Appendix A according to $|p\rangle\langle p| = (\mathbf{I}_2 + \hat{p}\cdot\vec{\boldsymbol{\sigma}})/2$ and $|q\rangle\langle q| = (\mathbf{I}_2 - \hat{p}\cdot\vec{\boldsymbol{\sigma}})/2$, where $\hat{p} = \langle p|\vec{\boldsymbol{\sigma}}|p\rangle$ and $\hat{q} = -\hat{p}$ since $\langle p|q\rangle = 0$, I find

$$\begin{aligned}\mathbf{U}_m(\omega) &= \cos\left(\frac{\tau\omega}{2}\right)\mathbf{I}_2 - i\sin\left(\frac{\tau\omega}{2}\right)\hat{p}\cdot\vec{\boldsymbol{\sigma}} \\ &= \exp\left(-\frac{i}{2}\tau\omega\,\hat{p}\cdot\vec{\boldsymbol{\sigma}}\right),\end{aligned} \tag{1.27}$$

which is recognized as a unitary Jones matrix, Eq. (1.16). Finally, the *PMD vector*, $\vec{\tau}_m = \tau\hat{p}$, with magnitude and direction corresponding to the DGD and fast polarization axis, respectively, provides a convenient parametrization of the polarization frequency dynamics

$$\mathbf{U}_m(\omega) = \exp\left(-\frac{i}{2}\omega\,\vec{\tau}_m\cdot\vec{\boldsymbol{\sigma}}\right). \tag{1.28}$$

As a result, PMD acts to rotate an incident polarized field on the Poincaré sphere by an angle $\psi = \omega\,|\vec{\tau}_m|$ about the direction of $\vec{\tau}_m$. More generally, the concatenation of N linearly birefringent elements

$$\mathbf{T}_N(\omega) = \mathbf{U}_N(\omega)\,\mathbf{T}_{N-1}(\omega), \tag{1.29}$$

with $\mathbf{T}_0(\omega) = \mathbf{I}_2$ [24,47,55], characterizes the frequency domain Jones matrix of the optical fiber emulator. Unlike $\mathbf{U}_m(\omega)$ however, the polarization at the output of $\mathbf{T}_N(\omega)$ can exhibit complicated stochastic variations with optical frequency due to the random orientation of adjacent emulator sections.

The frequency evolution of $\mathbf{U}_m(\omega)$ can be recast as a differential equation involving the PMD vector $\vec{\tau}_m$. Since $\vec{\tau}_m$ is frequency independent, differentiating $\mathbf{U}_m(\omega)$ yields

$$\frac{\partial\mathbf{U}_m}{\partial\omega}\mathbf{U}_m^{-1} = -\frac{i}{2}\,\vec{\tau}_m\cdot\vec{\boldsymbol{\sigma}} \tag{1.30}$$

and, consequently, if $\mathbf{U}_m(\omega)$ is known for example from straightforward measurements involving a commercially available polarimeter, the magnitude and direction of the PMD vector can be determined through numerical differentiation. A more robust variation of this approach that instead evaluates the eigenvalues and eigenvectors of $\mathbf{U}_m(\omega+\Delta\omega)\mathbf{U}_m(\omega)^{-1}$, termed *Jones matrix eigenanalysis* (JME), has become a standardized measurement procedure of PMD in installed fiber optic systems [50]. Analogous to Eq. (1.30), the total PMD vector $\vec{\Omega}_N(\omega)$ of the optical system is defined according to [9, 47, 91]

$$\frac{\partial \mathbf{T}_N}{\partial \omega}\mathbf{T}_N^{-1} = -\frac{i}{2}\,\vec{\Omega}_N(\omega)\cdot\vec{\boldsymbol{\sigma}}, \tag{1.31}$$

with the magnitude and direction of $\vec{\Omega}_N(\omega)$ corresponding to the DGD and fast principal state of polarization (PSP) at optical frequency ω, respectively. Note that unlike $\vec{\tau}_m$ both the magnitude and direction of $\vec{\Omega}_N(\omega)$ vary stochastically with ω due to the random mode coupling between concatenated emulator sections.

Over small frequency intervals $\Delta\omega$ the frequency variation of $\vec{\Omega}_N(\omega)$ is often quantified in terms of an expansion in Taylor orders of the PMD vector about the optical carrier ω_0

$$\vec{\Omega}_N(\omega) \approx \vec{\Omega}_N(\omega_0) + \left.\frac{\partial\vec{\Omega}_N}{\partial\omega}\right|_{\omega_0}\Delta\omega + \left.\frac{\partial^2\vec{\Omega}_N}{\partial\omega^2}\right|_{\omega_0}\frac{\Delta\omega^2}{2!} + \ldots \tag{1.32}$$

with the n:th order PMD vector

$$\vec{\Omega}^{(n)} = \left.\frac{\partial^{n-1}\vec{\Omega}}{\partial\omega^{n-1}}\right|_{\omega_0} \tag{1.33}$$

responsible for what is typically referred to as "higher-order" PMD pulse distortion [30, 51, 64]. An analytic formula for the exact evaluation of the n:th order PMD vector in an N section fiber emulator will be presented in Chapter 2. The physical effects associated with second order PMD for example are often analyzed by decomposing $\vec{\Omega}_N(\omega)$ into its magnitude and direction $\vec{\Omega}_N = \tau\hat{p}$ and differentiating each term

$$\frac{\partial\vec{\Omega}_N}{\partial\omega} = \frac{\partial\tau}{\partial\omega}\,\hat{p} + \tau\frac{\partial\hat{p}}{\partial\omega}, \tag{1.34}$$

where $\partial\hat{p}/\partial\omega$ is orthogonal to $\hat{p}$ since $\hat{p}$ has unit length. The frequency derivative of the differential group delay, τ, is a differential chromatic dispersion between polarization modes, while the second term acts to "depolarize" an incident optical pulse by rotating the principal state of polarization with increasing optical frequency. Under normal operating conditions, second order PMD depolarization results in a system impairment approximately 8 times larger than the differential chromatic dispersion [51, 63]. Although PMD vectors of order ≥ 3 can be analyzed in a similar manner [64], the physical interpretation of the resulting PMD induced pulse distortion becomes increasingly complicated.

Further, it becomes increasingly difficult to relate Taylor orders of the PMD vector directly to the frequency variation of the Jones matrix $\mathbf{T}_N(\omega)$. Consequently, several authors [38, 86] have proposed matrix models of PMD that are applicable for relatively low values of the mean DGD, or attempt to solve Eq. (1.31) directly by truncating the Taylor expansion to either second or third order in the deviation $\Delta\omega$ in the optical frequency [36, 51]. Such solutions often do not preserve the unitarity of the Jones matrix $\mathbf{T}_N(\omega)$ reducing accuracy for large $\Delta\omega$ or alternatively large mean DGDs [20, 31, 50, 58, 89]. In Chapter 2, I will present a solution of $\mathbf{T}_N(\omega)$ that for a given mean DGD approximates the frequency evolution of the Jones matrix over a larger bandwidth than previously published results while preserving the relevant group properties of the exact solution. The resulting formalism not only extends to arbitrary orders of the PMD vector, but also applies to more general optical systems that additionally include polarization dependent loss.

To illustrate the impact of higher-order PMD, in Fig. 1.2 I show the variation of the output SOP $|A_{out}(\omega)\rangle = \mathbf{T}_N(\omega)|A_{in}\rangle$ with increasing optical frequency for a horizontally polarized input $|A_{in}\rangle = [1, 0]^{\mathrm{T}}$ for (a) a single emulator section with 60 ps DGD and (b) a concatenation of $N = 200$ randomly oriented emulator sections with an instantaneous DGD of 60 ps at the central frequency, where the fast polarization axis in each case was aligned at 45 deg to the incident SOP. Case (a) represents the worst case polarization variation, occurring with a probability of 10^{-5}, that might be expected for $\approx$ 1000 km

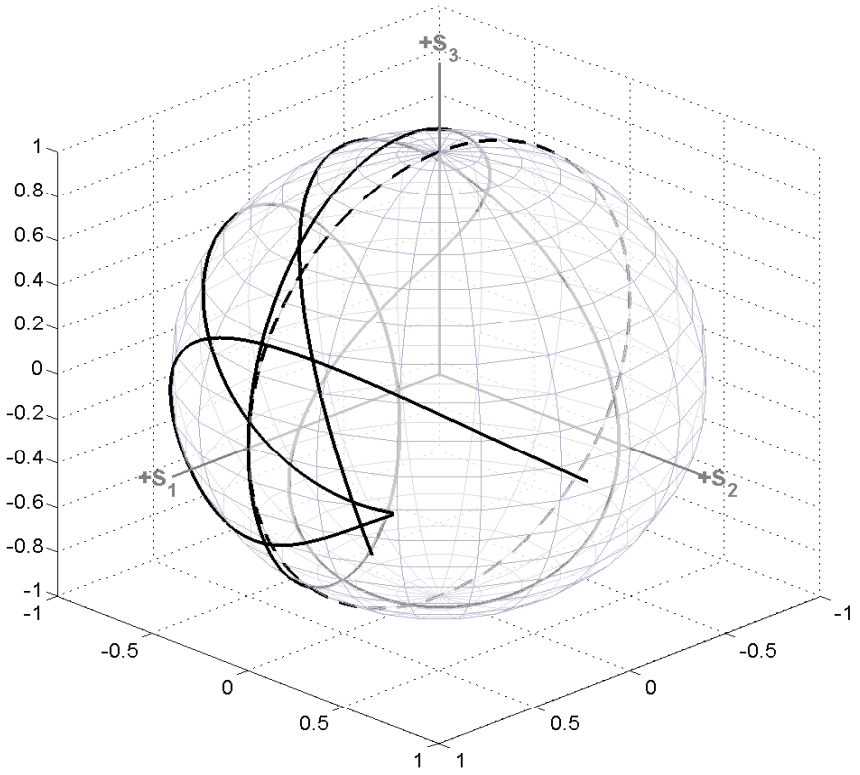


Figure 1.2: The variation of $|A_{out}(\omega)\rangle = \mathbf{T}_N(\omega)|A_{in}\rangle$ on the Poincaré sphere with increasing optical frequency for a horizontally polarized input $|A_{in}\rangle = [1, 0]^{\mathrm{T}}$ for (a) a single emulator section with 60 ps DGD (dashed line) and (b) a concatenation of $N = 200$ randomly oriented emulator sections with an instantaneous DGD of 60 ps at the central frequency (solid line). The fast polarization axis in each case was aligned at 45 deg to the incident SOP.

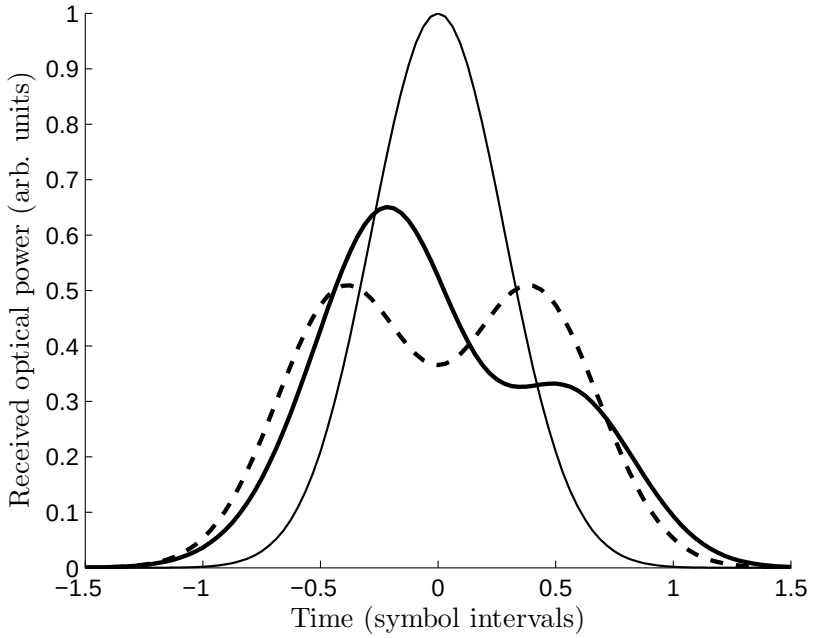


Figure 1.3: The received optical intensity for a 10 Gb/s horizontally polarized waveform for cases (a) (dashed line) and (b) (thick solid line) of Fig. 1.2. The input pulse is shown as the thin solid line.

of fiber over a 20 year anticipated life span of an older installed network with a PMD coefficient of 0.6 ps/$\sqrt{\text{km}}$. Case (b) on the other hand illustrates a long $\approx$ 120 m length of polarization maintaining fiber with the incident polarization offset by 45 deg relative to the polarization maintaining axis. Over a $\pm$30 GHz frequency interval, the incident SOP rotates about the direction of the fast polarization axis for case (a) (dashed line), while for case (b) the SOP varies stochastically over the Poincaré sphere (solid line). Notice that for small frequency offsets the local polarization activity for case (b) is reasonably well approximated by a rotation about the s_2 Stokes space axis.

The corresponding time-domain behaviour of a received optical pulse is shown in Fig. 1.3. Here, the Jones matrices corresponding to case (a) and (b) above are applied to a horizontally polarized 10 Gb/s incident optical pulse. In case (a), the output pulse results from the linear superposition of the optical fields coupled to each of the orthogonal fast and slow axes, while in (b) strong mode coupling between emulator sections leads to a complicated, stochastic temporal distortion of the received pulse. The results of several system studies [14, 27] have indicated the importance of including these higher order PMD distortions when assessing the impact of PMD on transmission performance.

In what follows, I will return to Eq. (1.29) to discuss a number of important properties of the PMD vector that follow from the recursion relation for Jones matrices. First, a powerful equation describing the accumulation of PMD in systems involving concatenations of two or more fiber segments, or in general, birefringent optical devices can be derived by combining Eq. (1.31) and Eq. (1.29). The PMD vector $\vec{\Omega}_N(\omega)$ associated with $\mathbf{T}_N(\omega)$ is

calculated as

$$\begin{aligned}
-\frac{i}{2}\vec{\Omega}_N \cdot \vec{\boldsymbol{\sigma}} &= \frac{\partial \mathbf{T}_N}{\partial \omega}\mathbf{T}_N^{-1} \\
&= \frac{\partial \mathbf{U}_N}{\partial \omega}\mathbf{U}_N^{-1} + \mathbf{U}_N\left[\frac{\partial \mathbf{T}_{N-1}}{\partial \omega}\mathbf{T}_{N-1}^{-1}\right]\mathbf{U}_N^{-1} \\
&= -\frac{i}{2}\vec{\tau}_N \cdot \vec{\boldsymbol{\sigma}} + \mathbf{U}_N\left[-\frac{i}{2}\vec{\Omega}_{N-1} \cdot \vec{\boldsymbol{\sigma}}\right]\mathbf{U}_N^{-1} \\
&= -\frac{i}{2}\vec{\tau}_N \cdot \vec{\boldsymbol{\sigma}} - \frac{i}{2}\left[\mathbf{R}_N\vec{\Omega}_{N-1}\right] \cdot \vec{\boldsymbol{\sigma}} \\
&= -\frac{i}{2}\left[\vec{\tau}_N + \mathbf{R}_N\vec{\Omega}_{N-1}\right] \cdot \vec{\boldsymbol{\sigma}}. \qquad (1.35)
\end{aligned}$$

Here, I have employed the identity $\mathbf{U}\big(\vec{A} \cdot \vec{\boldsymbol{\sigma}}\big)\mathbf{U}^{-1} = \big(\mathbf{R}\,\vec{A}\big) \cdot \vec{\boldsymbol{\sigma}}$ where, for a unitary Jones matrix $\mathbf{U}$, $\mathbf{R}$ is the corresponding Stokes rotation matrix. The cumulative PMD vector for a series of N concatenated Jones matrices is then

$$\vec{\Omega}_N(\omega) = \vec{\tau}_N(\omega) + \mathbf{R}_N(\omega)\vec{\Omega}_{N-1}(\omega), \qquad (1.36)$$

where $\vec{\Omega}_0(\omega) = 0$ and in general each segment PMD vector, $\vec{\tau}_m(\omega)$, can be frequency dependent. Eq. (1.36), often referred to as the *PMD concatenation rule*, is the basis for most statistical analyses of PMD [19, 60, 61, 108] and has found widespread application in studies of PMD induced pulse distortion [30, 92] and more recently in the hinge model of PMD temporal dynamics [4].

A number of important properties of the PMD vector follow immediately from the concatenation rule, Eq. (1.36). If the segment PMD vectors $\vec{\tau}_m$ are spherically symmetric and are further independent and randomly distributed between segments, then by the central limit theorem the probability density of each component of $\vec{\Omega}_N(\omega)$ tends towards a zero mean normal distribution for $N \gg 1$ with the variance increasing as $\sqrt{N}$. Consequently, the DGD $\tau = |\vec{\Omega}_N|$ of the emulator is Maxwellian distributed

$$f_\tau(x) = \frac{32}{\pi^2}\frac{x^2}{\tau_{\text{mean}}^3}\exp\left(-\frac{4}{\pi}\frac{x^2}{\tau_{\text{mean}}^2}\right), \qquad (1.37)$$

with mean, $\tau_{\text{mean}} = \mathrm{E}\{|\vec{\Omega}_N|\}$ and rms, $\tau_{\text{rms}} = \sqrt{\mathrm{E}\{|\vec{\Omega}_N|^2\}}$, increasing as the square root of the number of emulator segments (the probability density of the DGD for finite N is somewhat more complicated and has been considered in, for example, Refs. [4, 59]). Numerous field measurements [61, 90] have confirmed both the Maxwellian distribution of the DGD and the square root dependence of its mean value with increasing length in fibers significantly longer than the correlation length scales associated with fluctuations of the underlying birefringence. In fact, these statistical properties have led to the PMD coefficient, D_{pmd}, of optical fiber being specified in units of ps/$\sqrt{\text{km}}$ where $\tau_{\text{mean}} = D_{pmd}\sqrt{L}$ is the mean PMD of a fiber of length L. As fiber manufacturing technology improves, the PMD coefficient has steadily decreased from $D_{pmd} > 0.5$ ps/$\sqrt{\text{km}}$ in older fiber predating the year 2000 to the present value $D_{pmd} \leq 0.05$ ps/$\sqrt{\text{km}}$.

1.4.1 Frequency autocorrelation of the PMD vector

Next, due to its relevance in the time dependent systems analyzed in Chapter 4, I will discuss the frequency autocorrelation function of the fiber emulator that has been shown to coincide with measurement over a broad spectral range [61]. The results of this analysis characterize the expected stochastic evolution of the PMD vector with increasing optical frequency and provide a characteristic frequency interval, that I will refer to as the decorrelation bandwidth ω_d, that will be related in a simple manner to the rms DGD, τ_{rms}, of the fiber emulator. The PMD vector evolves into a largely independent stochastic realization for frequency offsets $\Delta\omega > \omega_d$, while for $\Delta\omega \ll \omega_d$ a simple first order PMD model is often sufficient for describing the polarization frequency evolution.

The frequency autocorrelation function of the fiber emulator follows from the PMD concatenation rule, Eq. (1.36), by considering

$$C_N(\omega;\Delta\omega) = \mathrm{E}\left\{\vec{\Omega}_N(\omega)\cdot\vec{\Omega}_N(\omega+\Delta\omega)\right\} \tag{1.38}$$

in which the segment PMD vector $\vec{\tau}_m$ is frequency independent while $\mathbf{R}_m(\omega) = \exp(\omega\vec{\tau}_m\times)$. If the $\vec{\tau}_m$ are spherically symmetric and independent and identically distributed between emulator sections then

$$C_N(\omega;\Delta\omega) = \mathrm{E}\left\{|\vec{\tau}_N|^2\right\} + \mathrm{E}\left\{\vec{\Omega}_{N-1}^{\mathrm{T}}(\omega)\mathbf{R}_N^{\mathrm{T}}(\omega)\mathbf{R}_N(\omega+\Delta\omega)\vec{\Omega}_{N-1}(\omega+\Delta\omega)\right\} \tag{1.39}$$

while the rotation autocorrelation

$$\mathrm{E}\left\{\mathbf{R}_N^{\mathrm{T}}(\omega)\mathbf{R}_N(\omega+\Delta\omega)\right\} = \mathrm{E}\left\{\exp\left(\Delta\omega\,\vec{\tau}_m\times\right)\right\} \tag{1.40}$$

must by symmetry be proportional to the identity matrix. Letting τ_{rms} be the rms DGD of the emulated PMD vector and $\mathrm{E}\left\{\mathbf{R}_N^{\mathrm{T}}(\omega)\mathbf{R}_N(\omega+\Delta\omega)\right\} = g(\omega;\Delta\omega)\,\mathbf{I}_3$, I find

$$C_N(\omega;\Delta\omega) = \frac{\tau_{\mathrm{rms}}^2}{N} + g(\omega;\Delta\omega)C_{N-1}(\omega;\Delta\omega), \tag{1.41}$$

where $C_0(\omega;\Delta\omega) = 0$ and the mean-square DGD of the m:th emulator section is $\mathrm{E}\left\{|\vec{\tau}_m|^2\right\} = \tau_{\mathrm{rms}}^2/N$. Finally, solving for $C_N(\omega,\Delta\omega)$ yields

$$C_N(\omega;\Delta\omega) = \frac{\tau_{\mathrm{rms}}^2}{N}\left[\frac{1-g^N(\omega;\Delta\omega)}{1-g(\omega;\Delta\omega)}\right] \tag{1.42}$$

where the rotation autocorrelation function,

$$g(\omega;\Delta\omega) = \frac{1}{3}\,\mathrm{E}\left\{\mathrm{Trace}\left\{\mathbf{R}_N^{\mathrm{T}}(\omega)\mathbf{R}_N(\omega+\Delta\omega)\right\}\right\}, \tag{1.43}$$

may be evaluated in the above expression for any particular model of the local fiber birefringence.

For example, in the converged $N\to\infty$ limit, $g(\omega;\Delta\omega) \approx 1 - \tau_{\mathrm{rms}}^2\Delta\omega^2/3N$ while $g^N(\omega;\Delta\omega) \approx \exp(-\tau_{\mathrm{rms}}^2\Delta\omega^2/3)$ yielding

$$C_\infty(\omega;\Delta\omega) = \tau_{\mathrm{rms}}^2\left[\frac{1-\mathrm{e}^{-\Delta\omega^2/\omega_d^2}}{\Delta\omega^2/\omega_d^2}\right], \tag{1.44}$$

independent of the particular choice of statistical distribution chosen to model $\vec{\tau}_m$. Here, the decorrelation bandwidth $\omega_d = \sqrt{3}/\tau_{\rm rms}$ of the PMD vector represents the characteristic frequency interval over which the fiber emulator transitions into a largely uncorrelated state. The decorrelation frequency also provides an approximate, intuitive justification for the higher order PMD expansion of Eq. (1.32) [40, 64]. For modulated signals with bandwidths $\ll \omega_d$, the emulated PMD vector is approximately constant and a first order model is often sufficient for assessing PMD induced system penalty. For bandwidths approaching ω_d, however, higher order PMD can be significant and, as I will show in Chapter 2, ≥ 5 Taylor orders of the PMD vector are required to quantify the frequency variation of $\vec{\Omega}_N(\omega)$ with a high degree of confidence.

For finite N, however, the autocorrelation properties of the fiber emulator can be significantly affected by the particular stochastic model for each $\vec{\tau}_m$. To illustrate, I will first consider a constant modulus model of the segment PMD $\vec{\tau}_m = \tau\,\hat{p}_m$ where $\tau = \tau_{\rm rms}/\sqrt{N}$ is constant for all emulator sections and each $\hat{p}_m$ is selected from a uniform distribution over the unit sphere. In this case, I find

$$g(\omega;\Delta\omega) = \frac{1}{3} + \frac{2}{3}\cos\left(\frac{\tau_{\rm rms}}{\sqrt{N}}\Delta\omega\right), \tag{1.45}$$

and $C_N(\omega;\Delta\omega)$ is periodic in $\Delta\omega$ with free spectral range $\omega_{\rm FSR} = 2\pi\sqrt{N}/\tau_{\rm rms}$. The constant modulus model therefore approximates $C_\infty(\omega;\Delta\omega)$ of optical transmission fiber over bandwidths $< \omega_{\rm FSR}/2$, making it somewhat unsuitable for evaluating pulse distortion in high PMD or WDM systems employing several optical channels with small N. An alternative, improved model employs a zero-mean Gaussian distribution for the components of $\vec{\tau}_m$, with variance $\sigma_\tau^2 = \tau_{\rm rms}^2/3N$, so that the magnitude $|\vec{\tau}_m|$ is Maxwellian distributed with first and second moments $\mathrm{E}\{|\vec{\tau}_m|\} = \sigma_\tau\sqrt{8/\pi}$ and $\mathrm{E}\{|\vec{\tau}_m|^2\} = 3\,\sigma_\tau^2$, respectively. In this case,

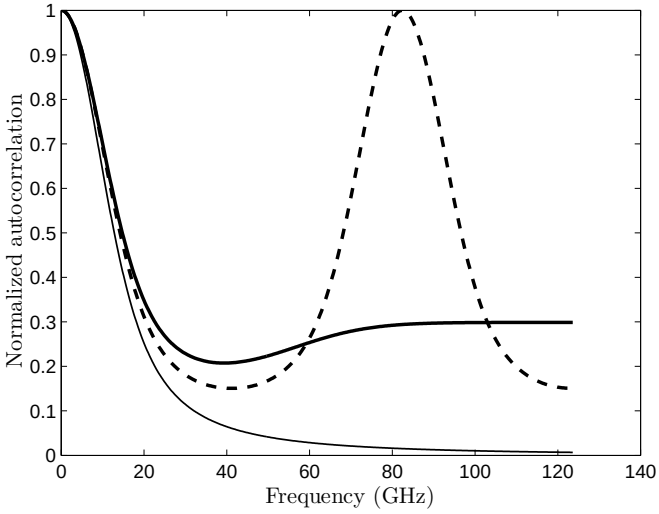
$$g(\omega;\Delta\omega) = \frac{1}{3} + \frac{2}{3}\left[1 - \sigma_\tau^2\Delta\omega^2\right]\exp\left(-\sigma_\tau^2\Delta\omega^2/2\right), \tag{1.46}$$

and the periodicities of $C_N(\omega;\Delta\omega)$ are suppressed for finite N, significantly increasing the free spectral range of the fiber emulator. Note that in deriving this last expression I have applied $\mathrm{E}\{\cos(\tau\Delta\omega)\} = \left[1-\sigma_\tau^2\Delta\omega^2\right]\exp\left(-\sigma_\tau^2\Delta\omega^2/2\right)$ whenever τ is Maxwellian distributed with $\mathrm{E}\{\tau\} = \sigma_\tau\sqrt{8/\pi}$. For sufficiently large $\Delta\omega$, $C_N(\omega;\Delta\omega)$ approaches a finite background autocorrelation of $3/2N$ that can be minimized only by further increasing the number of emulator sections. While previous authors have shown [10,70] that the Gaussian distribution of $\vec{\tau}_m$ accurately represents the statistics of higher order PMD with as few as $N = 5$ sections, I note that a significantly larger N is often required to accurately reproduce the frequency and temporal autocorrelation statistics of the converged result with reasonable accuracy.

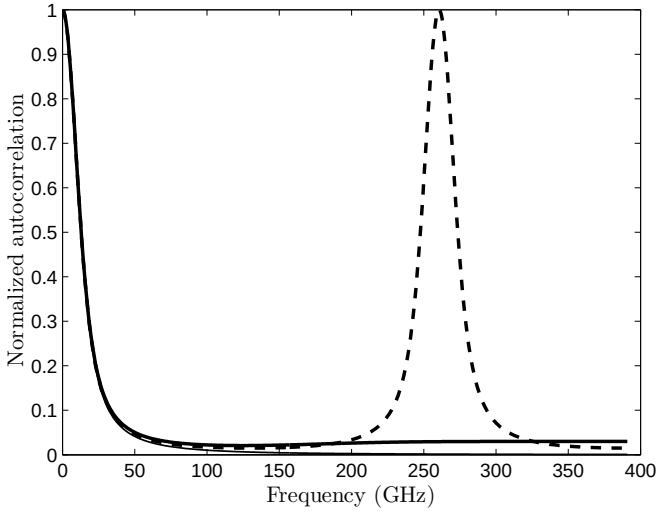
In Fig. 1.4, I illustrate the autocorrelation of the emulated PMD vector, normalized by τ_{rms}^2, for (a) $N = 5$ and (b) $N = 50$ emulator sections. The constant modulus and Gaussian models of $\vec{\tau}_m$ are shown in this figure as the dashed and thick solid lines, respectively, while the thin solid line displays the converged $N \to \infty$ result of Eq. (1.44). Deviations from the converged autocorrelation function are apparent with a reduced number of emulator sections. The difference between the emulated autocorrelation and the converged result is significantly reduced for $N = 50$, with the background autocorrelation in this case < 0.1.

1.4.2 Brownian bridge algorithm for PMD

A significant advantage of the Gaussian distributed model of segment PMD vectors $\vec{\tau}_m$ is the ability to precondition the fiber emulator in an arbitrary initial state at the central frequency ω_0 using the Brownian bridge algorithm [107]. This approach, while straightforward to implement, allows the precise numerical evaluation of conditional probability densities associated with the PMD vector and will be used extensively in Chapter 4 to precondition a temporal PMD emulator on a specified PMD vector at the initial time $t = 0$. To implement the Brownian bridge algorithm, I return to the PMD concatenation rule,



(a) N = 5



(b) N = 50

Figure 1.4: The autocorrelation of the emulated PMD vector, normalized by τ_{rms}^2, for (a) $N = 5$ and (b) $N = 50$ emulator sections. The thin solid line depicts Eq. (1.44) while the dashed and thick solid lines show the emulated autocorrelation function for the constant modulus and Gaussian distributed models of the segment PMD, respectively.

Eq. (1.36), and observe that at the optical carrier frequency ω_0 the PMD vector $\vec{\Omega}_N(\omega_0)$ is a sum of N independent Gaussian distributed vectors, $\vec{b}_m$, according to

$$\begin{aligned}\vec{\Omega}_N(\omega_0) &= \vec{\tau}_N + \mathbf{R}_N(\omega_0)\vec{\Omega}_{N-1}(\omega_0) \\ &= \vec{\tau}_N + \mathbf{R}_N(\omega_0)\vec{\tau}_{N-1} + \mathbf{R}_N(\omega_0)\mathbf{R}_{N-1}(\omega_0)\vec{\tau}_{N-2} + \ldots \\ &= \vec{b}_N + \vec{b}_{N-1} + \vec{b}_{N-2} + \ldots\end{aligned} \tag{1.47}$$

in which $\vec{b}_m = \mathbf{R}_N(\omega_0)\cdots\mathbf{R}_{m+1}(\omega_0)\,\vec{\tau}_m$. The algorithm is then as follows. Suppose that $\vec{\Omega}_c$ represents an arbitrary initial PMD vector at ω_0 (the constraint) while the components of the random vectors $\vec{r}_m$, for $m = 1, \ldots, N$, are selected from a zero-mean Gaussian distribution with $\mathrm{E}\{|\vec{r}_m|^2\} = \tau_{\mathrm{rms}}^2/N$. Next, I calculate [107]

$$\vec{b}_m = \vec{r}_m - \frac{1}{N}\sum_{k=1}^{N}\vec{r}_k + \frac{\vec{\Omega}_c}{N}, \tag{1.48}$$

while the polarization rotation introduced by each emulator segment is compensated by setting $\vec{\tau}_N = \vec{b}_N$, $\vec{\tau}_{N-1} = \mathbf{R}_N^{-1}(\omega_0)\vec{b}_{N-1}$, and so forth. The PMD vector $\vec{\Omega}_N(\omega)$ at all other frequencies is calculated from Eq. (1.36). In this manner, $\vec{\Omega}_N(\omega_0) = \vec{\Omega}_c$ since the sum over all $\vec{b}_m$ yields the desired constraint, while the jointly Gaussian $\vec{b}_m$ can be shown to satisfy the correct conditional densities [107]. It follows that the ensemble of emulator states produced by this algorithm matches that of a Monte-Carlo simulation in which only those $\vec{\Omega}_N(\omega_0)$ sufficiently close to $\vec{\Omega}_c$ are kept. Consequently, this algorithm significantly improves simulation efficiency in the typical case where the probability of randomly generating $\vec{\Omega}_c$ is otherwise extremely low.

An application of the Brownian bridge algorithm is shown in Fig. 1.5 in which an $N = 100$ section fiber emulator with mean DGD $\tau_{\mathrm{mean}} = 10$ ps is preconditioned to $\vec{\Omega}_c = 3\,\tau_{\mathrm{mean}}\,[1, 0, 0]^{\mathrm{T}}$. The black lines of Fig. 1.5 represent 10 different random realizations of the fiber emulator and clearly illustrate the constraint imposed at the central optical frequency. In typical network applications a DGD of $3\,\tau_{\mathrm{mean}}$ often defines the threshold of a

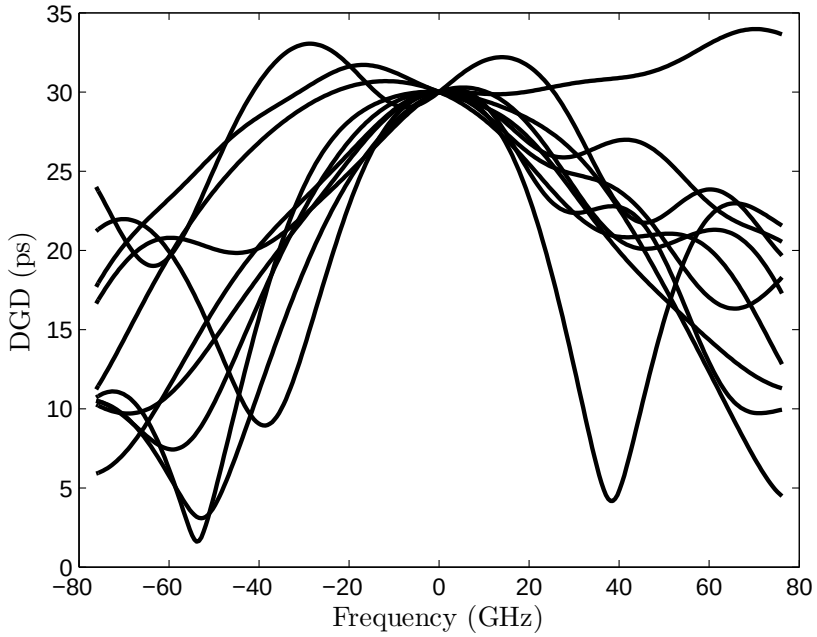


Figure 1.5: 10 independent stochastic realizations of an $N = 100$ section fiber emulator, with mean DGD $\tau_{\text{mean}} = 10$ ps, in which the DGD at the central frequency was constrained to $3\tau_{\text{mean}}$ using the Brownian bridge algorithm.

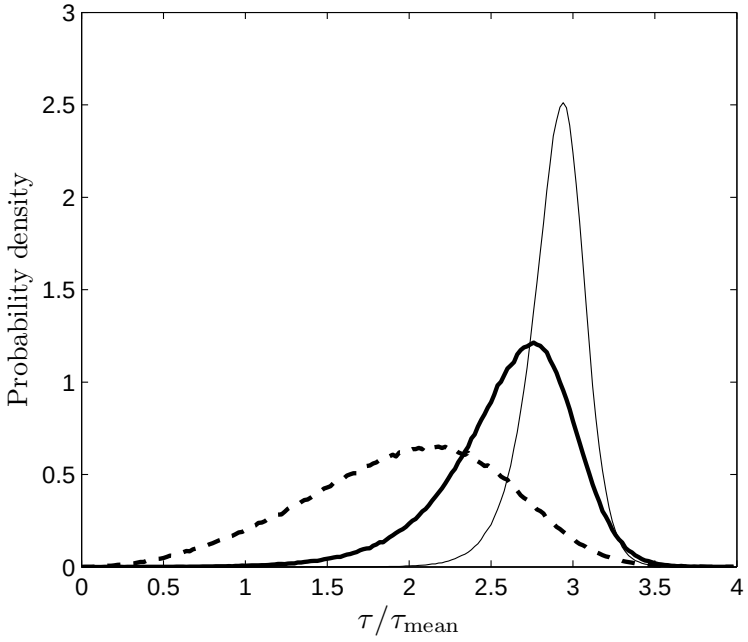


Figure 1.6: The probability density of the DGD at frequency offsets $\Delta\omega = 0.5\omega_d$ (thin solid line), ω_d (thick solid line) and $2\omega_d$ (dashed line) relative to the optical carrier for an $N = 100$ section emulator with mean DGD $\tau_{\text{mean}} = 10$ ps. Here, the emulated DGD was constrained to $3\tau_{\text{mean}}$ at $\Delta\omega = 0$ using the Brownian bridge algorithm.

PMD induced system outage event. A Maxwellian distributed DGD exceeds this threshold with a probability of $\approx 10^{-5}$, and several million Monte-Carlo iterations would be required on average to isolate 10 realizations of the emulator with DGD values sufficiently close to the constraint. Similarly, the probability density of the DGD is shown in Fig. 1.6 for frequency offsets $\Delta\omega = 0.5\omega_d$ (thin solid line), ω_d (thick solid line) and $2\omega_d$ (dashed line) relative to the optical carrier. The probability density converges to the Maxwellian distribution only for $\Delta\omega > 10\omega_d$, implying that a PMD outage event occurring in one channel of a multi-channel WDM system can degrade the performance of neighbouring channels within several PMD correlation bandwidths [107].

1.5 Polarization dependent loss

Optical communication systems typically include components such as amplifiers, couplers and isolators that unlike standard single mode fiber exhibit different attenuations or gains depending upon a signal's incident state of polarization, an effect termed polarization dependent loss/gain (PDL/PDG). These optical components are typically separated by long lengths of transmission fiber with significant PMD, leading to a complex interaction between PMD and PDL in which the total accumulated PDL/PDG of the system becomes strongly wavelength dependent [43, 55, 56]. In this section, I will summarize the important properties of PDL and will present an emulator model suitable for analyzing the combined effects of PMD and PDL in optical networks [94].

In an optical device with PDL or PDG, incident light generally experiences different frequency independent attenuations (or gains) α_1 and α_2 when polarized along one of two orthogonal principal polarization orientations [43]. The Jones matrix $\mathbf{H}_m$ describing the

polarization transformation through the m:th optical device is

$$\mathbf{H}_m = \mathbf{V}_m \begin{bmatrix} \mathrm{e}^{-\alpha_1/2} & 0 \\ 0 & \mathrm{e}^{-\alpha_2/2} \end{bmatrix} \mathbf{V}_m^\dagger \tag{1.49}$$

$$= \exp\left(-\frac{\bar{\alpha}_m}{2}\right) \mathbf{V}_m \begin{bmatrix} \mathrm{e}^{+\alpha_m/2} & 0 \\ 0 & \mathrm{e}^{-\alpha_m/2} \end{bmatrix} \mathbf{V}_m^\dagger, \tag{1.50}$$

where $\bar{\alpha}_m = (\alpha_2 + \alpha_1)/2$ and $\alpha_m = (\alpha_2 - \alpha_1)/2$ denote the mean and differential attenuation between polarization modes, respectively, and $\mathbf{V}_m$ is a unitary matrix specifying the arbitrary orientation of the principal axes relative to the input state of polarization. The PDL of the device is measured as the ratio of the maximum to minimum received optical power over all input polarization states and is related, in units of decibels, to the differential attenuation through

$$\alpha_m^{\mathrm{dB}} = |\alpha_m| \, \frac{20}{\ln 10}. \tag{1.51}$$

Finally, rewriting $\mathbf{H}_m$ in terms of a segment PDL vector, $\vec{\alpha}_m$,

$$\mathbf{H}_m = \exp\left(-\frac{\bar{\alpha}_m}{2}\right) \exp\left(\frac{1}{2}\, \vec{\alpha}_m \cdot \vec{\boldsymbol{\sigma}}\right) \tag{1.52}$$

can often provide more physical and geometric insight into the polarization dynamics of PDL [55, 94] and will be used in what follows to formulate a concatenation rule for the accumulated PDL in optical networks. Here, $\vec{\alpha}_m$ is a vector with magnitude equal to the differential attenuation, α_m, in the direction of the minimum attenuation axis on the Poincaré sphere. Note that in this thesis I will typically disregard the overall attenuation common to both polarization modes and will consider only the differential PDL affecting the polarization state of the optical field.

The Jones matrix $\mathbf{T}_N(\omega)$ of an optical system with both PMD and PDL is given by the product of N optical elements

$$\mathbf{T}_N(\omega) = \mathbf{H}_N \mathbf{U}_N(\omega) \mathbf{T}_{N-1}(\omega), \tag{1.53}$$

where $\mathbf{T}_0(\omega) = \mathbf{I}_2$ and, as in Eq. (1.29), $\mathbf{U}_m(\omega) = \exp\left(-i\omega\,\vec{\tau}_m \cdot \vec{\boldsymbol{\sigma}}/2\right)$ is the PMD Jones matrix of the m:th emulator section. To determine the total PDL associated with $\mathbf{T}_N(\omega)$, one widely-used measurement technique scans the polarization state of an incident constant-wave field $|A\rangle$ and records the ratio of the maximum and minimum received power over all input SOPs. Since $P = \langle A|\,\mathbf{T}_N^{\dagger}(\omega)\mathbf{T}_N(\omega)\,|A\rangle$ represents the power measured on a standard optical power meter, the ratio of the maximum to minimum powers is precisely equal to the ratio of the two eigenvalues $\lambda_{\max}(\omega)$ and $\lambda_{\min}(\omega)$ of $\mathbf{T}_N^{\dagger}(\omega)\mathbf{T}_N(\omega)$ and the total PDL of the emulated Jones matrix in units of decibels is given by

$$\alpha_{\mathrm{dB}}(\omega) = 10\log_{10}\left[\frac{\lambda_{\max}(\omega)}{\lambda_{\min}(\omega)}\right]. \tag{1.54}$$

Further, since $\mathbf{T}_N^{\dagger}(\omega)\mathbf{T}_N(\omega)$ is hermitian the two input polarization states that yield the maximum and minimum received power are orthogonal even if the eigenvectors of $\mathbf{T}_N(\omega)$ are not. Notice that unlike the PDL α_m of each segment, the total PDL $\alpha_{\mathrm{dB}}(\omega)$ of an optical system can exhibit strong frequency dependence due to the presence of PMD.

1.5.1 Concatenation rule for PDL in optical networks

For mean values of the PDL < 4 dB, it is possible to formulate an approximate concatenation rule for the total PDL vector, $\vec{\Gamma}_N(\omega)$, of N emulator sections in direct analogy to the PMD concatenation rule of Eq. (1.36). Here, the total PDL vector is defined as

$$\mathbf{T}_N^{\dagger}(\omega)\mathbf{T}_N(\omega) = \exp\left(\vec{\Gamma}_N(\omega)\cdot\vec{\boldsymbol{\sigma}}\right), \tag{1.55}$$

with the magnitude and direction of $\vec{\Gamma}_N(\omega)$ corresponding to the total PDL and minimum attenuation axis on the Poincaré sphere, respectively. While previous authors [81,131] have described the statistics of PDL using a stochastic differential equation for the longitudinal evolution of the PDL vector in a distributed system, here I will demonstrate that the PDL vectors of each segment, $\vec{\alpha}_m$, accumulate vectorially in a manner analogous to the

segment PMD vectors of Eq. (1.36). The resulting PDL concatenation rule provides a useful equation for analyzing the statistics of PDL in systems with only a finite number of PDL elements where the probability density of the total PDL can differ substantially from that of a distributed system [94]. Unlike Eq. (1.36), however, the PDL concatenation rule represents an approximation to the exact result that is most accurate for values of the mean PDL, α_{mean}, less than 4 dB. The concatenation rule is observed to provide a high degree of accuracy even for values of the instantaneous PDL that significantly exceed the mean.

To proceed, I will apply Eq. (1.53) to evaluate $\mathbf{T}_2^\dagger \mathbf{T}_2 = \exp\left(\vec{\Gamma}_2 \cdot \vec{\boldsymbol{\sigma}}\right)$ for $N = 2$ emulator sections and will subsequently generalize the result to arbitrary N. Observing that $\mathbf{H}_m = \exp\left(\vec{\alpha}_m \cdot \vec{\boldsymbol{\sigma}}/2\right)$ is hermitian and $\mathbf{U}_m = \exp\left(-i\omega\, \vec{\tau}_m \cdot \vec{\boldsymbol{\sigma}}/2\right)$ is unitary, I find

$$\begin{aligned} \mathbf{T}_2^\dagger \mathbf{T}_2 &= \left(\mathbf{U}_1^\dagger \mathbf{H}_1 \mathbf{U}_2^\dagger \mathbf{H}_2\right)\left(\mathbf{H}_2 \mathbf{U}_2 \mathbf{H}_1 \mathbf{U}_1\right) \\ &= \mathbf{U}_1^\dagger \mathbf{H}_1 \left(\mathbf{U}_2^\dagger \mathbf{H}_2^2 \mathbf{U}_2\right) \mathbf{H}_1 \mathbf{U}_1 \\ &= \mathbf{U}_1^\dagger \mathbf{H}_1 \, \exp\left(\, \mathbf{R}_2^{\mathrm{T}} \vec{\alpha}_2 \cdot \vec{\boldsymbol{\sigma}}\right) \, \mathbf{H}_1 \mathbf{U}_1 \end{aligned} \tag{1.56}$$

where $\mathbf{R}_2$ is the 3×3 Stokes matrix associated with $\mathbf{U}_2$, while applying the Baker-Campbell-Hausdorff (BCH) identity to $\mathbf{H}_1 \exp\left(\mathbf{R}_2^{\mathrm{T}} \vec{\alpha}_2 \cdot \vec{\boldsymbol{\sigma}}\right) \mathbf{H}_1$ yields

$$\begin{aligned} &\exp\left(\frac{1}{2}\vec{\alpha}_1 \cdot \vec{\boldsymbol{\sigma}}\right) \exp\left(\mathbf{R}_2^{\mathrm{T}} \vec{\alpha}_2 \cdot \vec{\boldsymbol{\sigma}}\right) \exp\left(\frac{1}{2}\vec{\alpha}_1 \cdot \vec{\boldsymbol{\sigma}}\right) \\ &\qquad = \exp\left(\left[\vec{\alpha}_1 + \mathbf{R}_2^{\mathrm{T}} \vec{\alpha}_2 + O(\alpha_{\text{mean}}^3)\right] \cdot \vec{\boldsymbol{\sigma}}\right). \end{aligned} \tag{1.57}$$

The higher order terms associated with the BCH expansion are random variables that depend upon the particular stochastic configuration of the emulator. In fact, since $\mathbf{H}_1$ appears symmetrically in Eq. (1.57), terms of order α_{mean}^2 cancel, leaving a lowest order contribution that scales as the cube of the mean PDL, α_{mean}^3 (numerical confirmation of

the α_{mean}^3 scaling will be presented in what follows). Next, I observe that

$$\begin{aligned}\mathbf{T}_2^\dagger \mathbf{T}_2 &= \mathbf{U}_1^\dagger \exp\left(\left[\vec{\alpha}_1 + \mathbf{R}_2^{\mathrm{T}}\vec{\alpha}_2 + O(\alpha_{\text{mean}}^3)\right]\cdot\vec{\boldsymbol{\sigma}}\right)\mathbf{U}_1 \\ &= \exp\left(\left[\mathbf{R}_1^{\mathrm{T}}\vec{\alpha}_1 + \mathbf{R}_1^{\mathrm{T}}\mathbf{R}_2^{\mathrm{T}}\vec{\alpha}_2 + O(\alpha_{\text{mean}}^3)\right]\cdot\vec{\boldsymbol{\sigma}}\right)\end{aligned} \tag{1.58}$$

and the total PDL vector is identified as

$$\vec{\Gamma}_2 = \mathbf{R}_1^{\mathrm{T}}\vec{\alpha}_1 + \mathbf{R}_1^{\mathrm{T}}\mathbf{R}_2^{\mathrm{T}}\vec{\alpha}_2 + O(\alpha_{\text{mean}}^3). \tag{1.59}$$

Letting $\boldsymbol{\mathcal{R}}_2 = \mathbf{R}_2\mathbf{R}_1$, this last result can be rewritten in a more convenient form by introducing the transformed PDL vector $\vec{\Gamma}'_2 = \boldsymbol{\mathcal{R}}_2\vec{\Gamma}_2$. In this manner, $\vec{\Gamma}'_2$ satisfies the concatenation rule

$$\vec{\Gamma}'_2 = \vec{\alpha}_2 + \mathbf{R}_2(\omega)\,\vec{\alpha}_1 + O(\alpha_{\text{mean}}^3), \tag{1.60}$$

while $\vec{\Gamma}_2$ can be recovered if necessary from $\vec{\Gamma}'_2$ by applying the inverse transformation. Finally, generalizing this result to arbitrary N, $\boldsymbol{\mathcal{R}}_N = \mathbf{R}_N\cdots\mathbf{R}_2\mathbf{R}_1$ and the transformed PDL vector $\vec{\Gamma}'_N = \boldsymbol{\mathcal{R}}_N\vec{\Gamma}_N$ satisfies the concatenation rule

$$\vec{\Gamma}'_N(\omega) = \vec{\alpha}_N + \mathbf{R}_N(\omega)\vec{\Gamma}'_{N-1}(\omega) + O(\alpha_{\text{mean}}^3), \tag{1.61}$$

where $\vec{\Gamma}'_0(\omega) = 0$ and $\mathbf{R}_m(\omega)$ is the 3×3 Stokes rotation matrix associated with $\mathbf{U}_m(\omega)$.

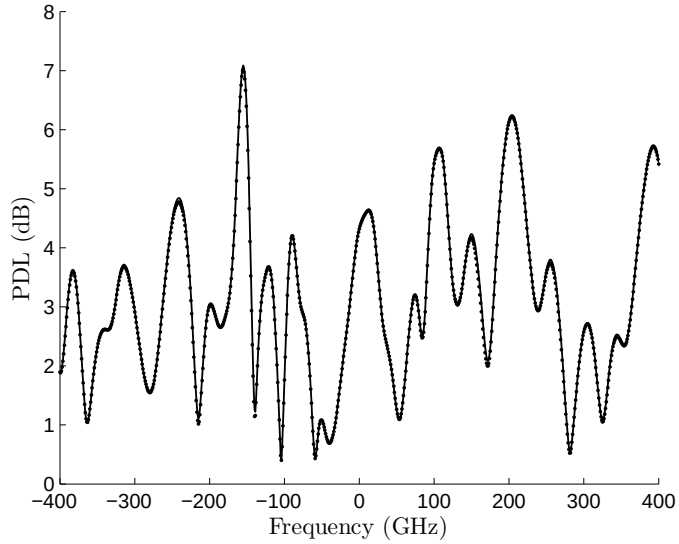
As with Eq. (1.36), a number of important statistical properties of the PDL follow directly from the PDL concatenation rule of Eq. (1.61). For example, if each $\vec{\alpha}_m$ is spherically symmetric, then for large N each component of $\vec{\Gamma}_N$ tends towards a zero-mean Gaussian distribution with the variance increasing as the square root of the number of PDL elements. Consequently, the instantaneous PDL $|\vec{\Gamma}_N|$ is Maxwellian distributed with the mean PDL, $\alpha_{\text{mean}} = \mathrm{E}\{|\vec{\Gamma}_N|\}$, increasing as the square root of the system length. Similarly, the frequency autocorrelation of $\vec{\Gamma}_N(\omega)$ can be analyzed using the procedure introduced previously in Section 1.4.1, with the final result possessing the same frequency dependence

as Eq. (1.42) for systems with only PMD. Although these statistical properties of PDL have been established previously in different contexts [55, 73], Eq. (1.61) often provides a more intuitive approach to their derivation and has the further advantage of allowing the emulator to be preconditioned to a particular PDL state at the optical carrier using the Brownian bridge algorithm.

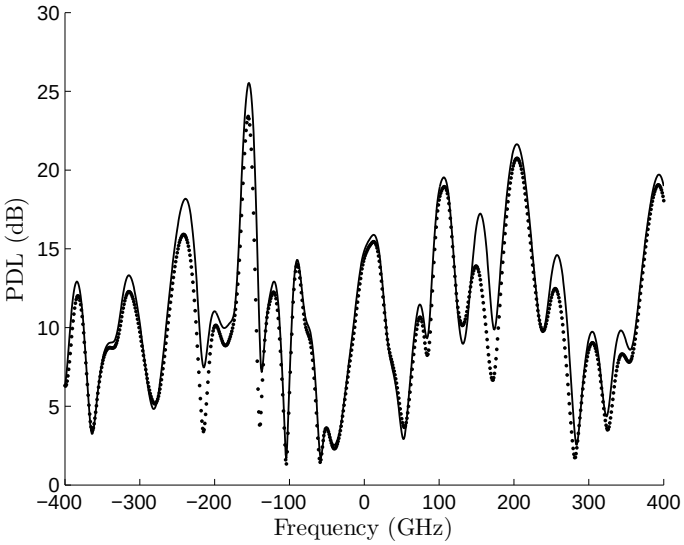
In Fig. 1.7a, I demonstrate the accuracy of the PDL concatenation rule by displaying the magnitude of the PDL for $\alpha_{\text{mean}} = 3$ dB and a mean DGD of 20 ps over a ± 400 GHz frequency interval for Eq. (1.53) (solid lines) and Eq. (1.61) ($\bullet$ markers) for an $N = 50$ section emulator. The PDL concatenation rule clearly coincides with the actual, emulated PDL even for instantaneous PDL values exceeding 2.5 times the average value. Similar results for $\alpha_{\text{mean}} = 10$ dB are shown in Fig. 1.7b, illustrating a regime in which the error terms associated with the PDL concatenation rule are no longer negligible. In Fig. 1.8, I show the variation of the $\log_{10}$ of the rms error $\epsilon_{\text{rms}} = \sqrt{\text{E}\left\{(\alpha_{est} - \alpha_{act})^2\right\}}$ between α_{act} from Eq. (1.54) and α_{est} from Eq. (1.61) with the $\log_{10}$ of the mean PDL. The rms error is well approximated by $\epsilon_{\text{rms}} \approx 10^{-2.92}\, \alpha_{\text{mean}}^3$ over $0 \leq \alpha_{\text{mean}} \leq 10$ dB, supporting the expected α_{mean}^3 scaling. Further, the rms error is < 0.04 dB for $\alpha_{\text{mean}} = 3$ dB with a maximum error of 0.35 dB recorded over 10^6 realizations of the emulator, making Eq. (1.61) suitable for the analysis of PDL over the range of mean PDL values typical in most telecom applications.

1.5.2 Brownian bridge algorithm for PDL

A potential significant benefit of Eq. (1.61) is the ability to apply the Brownian bridge algorithm to simulate the conditional densities associated with PDL using the procedure developed previously in Section 1.4.2 for systems with only PMD. Here, the PDL vector at the optical carrier frequency, ω_0, is prepared in a well defined initial state $\vec{\Gamma}_c$ by selecting each component of $\vec{\alpha}_m$ from a Gaussian distribution with variance $\text{E}\left\{|\vec{\alpha}_m|^2\right\} = \alpha_{\text{mean}}^2 3\pi/8N$ using a suitably modified version of Eq. (1.48). In Fig. 1.9 I demonstrate this method by



(a) $\alpha_{\text{mean}} = 3$ dB



(b) $\alpha_{\text{mean}} = 10$ dB

Figure 1.7: A comparison between the emulated PDL, Eq. (1.54) (solid line), and the PDL concatenation rule, Eq. (1.61) (• markers), for a single realization of an $N = 50$ section emulator with 20 ps mean DGD and (a) 3 dB or (b) 10 dB mean PDL.

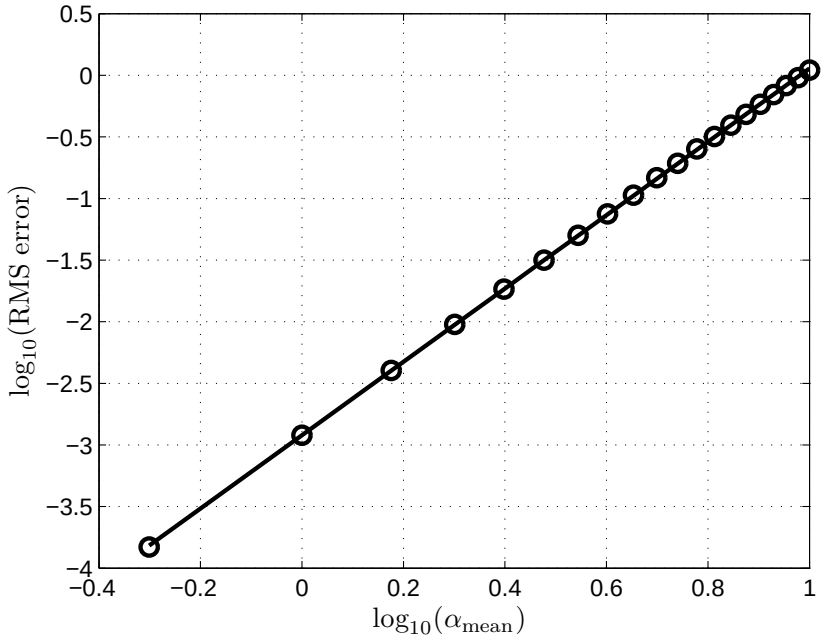


Figure 1.8: The $\log_{10}$ of the rms error $\epsilon_{\text{rms}} = \sqrt{\text{E}\left\{(\alpha_{est} - \alpha_{act})^2\right\}}$ between α_{act} from Eq. (1.54) and α_{est} from Eq. (1.61) as a function of the $\log_{10}$ of the mean PDL ($\circ$ markers). The solid line shows $\epsilon_{\text{rms}} = 10^{-2.92}\,\alpha_{\text{mean}}^3$.

conditioning a $N = 100$ section emulator with mean DGD of 20 ps and a mean PDL of 3 dB on the initial PDL vector, $\vec{\Gamma}_c = 3\,\alpha_{\text{mean}}\,[1,0,0]^{\text{T}}$. The PDL is calculated at all other optical frequencies using Eq. (1.54) and clearly reproduces the 9 dB PDL constraint at the optical carrier. Without the Brownian bridge algorithm, preconditioning the emulator in this manner would require on the order of several million Monte-Carlo emulator realizations to generate samples sufficiently close to the desired constraint.

1.5.3 Complex principal state vector for PMD and PDL

One of the main findings in the study of systems with PMD and PDL is the existence of two principal states of polarization at the output of the system that remain constant to first order with respect to small changes in the optical frequency. Unlike systems with only PMD, however, these two principal states are nonorthogonal and incident polarized light coupled to each interferes at the receiver, leading to an additional source of pulse broadening and fading in the received power that can be a significant penalty in direct detection systems [55, 116].

In the presence of PMD and PDL, the Jones matrix $\mathbf{T}_N(\omega)$ satisfies the differential equation [43, 55, 56]

$$\frac{\partial \mathbf{T}_N}{\partial \omega}\mathbf{T}_N^{-1} = -\frac{i}{2}\vec{W}_N \cdot \vec{\boldsymbol{\sigma}}, \tag{1.62}$$

where $\vec{W}_N(\omega) = \vec{\Omega}(\omega) + i\vec{\Lambda}(\omega)$ is often referred to as the *complex principal state vector* [20, 55, 98]. Analogous to the PMD vector of Eq. (1.31), the complex principal state vector characterizes the polarization frequency evolution of an optical system with PMD and PDL. I will present an algorithm for the exact evaluation of $\vec{W}_N(\omega)$ and its frequency derivatives in Chapter 2.

The two Jones space principal states, $|p_+\rangle$ and $|p_-\rangle$, are eigenvectors of the matrix $\vec{W}_N \cdot \vec{\boldsymbol{\sigma}}$ with complex eigenvalues $\pm\chi$ where $\chi = \sqrt{\vec{W}_N \cdot \vec{W}_N}$. That is, the two principal

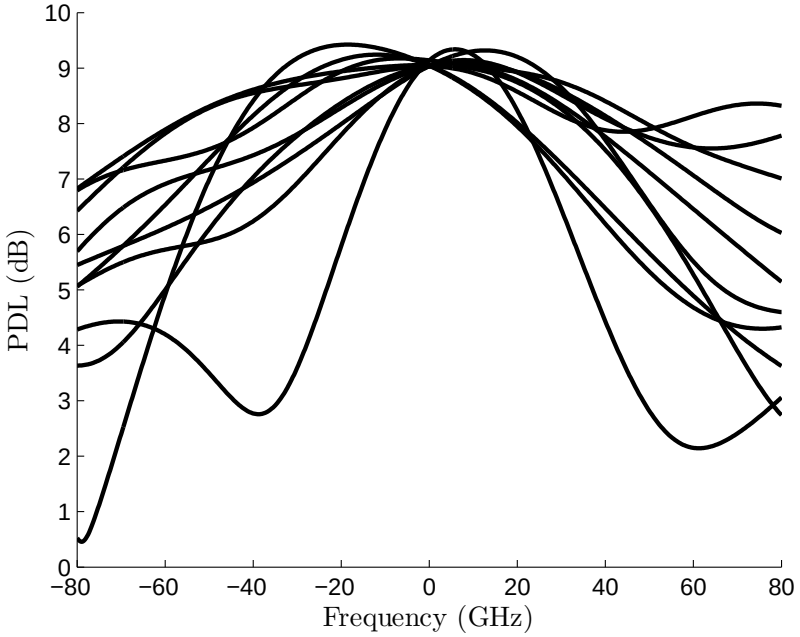


Figure 1.9: 10 independent stochastic realizations of an $N = 100$ section emulator with mean DGD $\tau_{\text{mean}} = 20$ ps and mean PDL $\alpha_{\text{mean}} = 3$ dB in which the PDL at the central frequency was constrained to $3\alpha_{\text{mean}}$ using the Brownian bridge algorithm.

states satisfy the eigenvalue equation

$$\left(\vec{W}_N \cdot \vec{\boldsymbol{\sigma}}\right) |p_\pm\rangle = \pm\chi\, |p_\pm\rangle, \tag{1.63}$$

where the real and imaginary components τ and η of $\chi = \tau + i\eta$ have the physical interpretation of the DGD and the frequency derivative of the differential attenuation between the two fundamental polarization modes. The angle between the two principal states can be evaluated directly from Eq. (1.63), and I find [94]

$$\begin{aligned} \hat{p}_\pm &= \langle p_\pm | \vec{\boldsymbol{\sigma}} | p_\pm \rangle \\ &= \pm\, 2 \frac{\tau\vec{\Omega} + \eta\vec{\Lambda} \pm \vec{\Omega} \times \vec{\Lambda}}{\tau^2 + \eta^2 + |\vec{\Omega}|^2 + |\vec{\Lambda}|^2}, \end{aligned} \tag{1.64}$$

while the relative angular separation is

$$\begin{aligned} \cos\theta &= \hat{p}_+ \cdot \hat{p}_- \\ &= 4 \frac{|\vec{\Omega} \times \vec{\Lambda}|^2 - |\tau\vec{\Omega} + \eta\vec{\Lambda}|^2}{\left(\tau^2 + \eta^2 + |\vec{\Omega}|^2 + |\vec{\Lambda}|^2\right)^2}. \end{aligned} \tag{1.65}$$

Notice that τ, η, $\vec{\Omega}$ and $\vec{\Lambda}$ are stochastic variables that may be evaluated from Eq. (1.53) for each particular realization of the emulator. Interference between the two principal states at the optical receiver can often lead to somewhat counter-intuitive propagation results. For example, an optical pulse can experience anomalous dispersion in which the temporal broadening can exceed the linear sum of all DGD elements in the system [43,56]. Several other examples illustrating the interactions between PMD and PDL, such as pulse spreading that can occur even with zero DGD, are discussed in Ref. [55].

1.6 Summary and outline

In this chapter, I have introduced the phenomenon of PMD and PDL in optical communication systems, with a particular emphasis on the general stochastic properties of fiber

"emulator" models that are often widely used for the theoretical and numerical evaluation of system performance. To summarize the key points from this chapter:

1. To first order, PMD is a differential group delay (DGD) introduced between two orthogonal polarization modes. Equivalently, in the frequency domain, PMD acts to rotate an incident SOP with increasing frequency by an angle proportional to the DGD about the fast polarization axis on the Poincaré sphere.

2. PDL represents a differential attenuation between two orthogonal polarization modes that renders the two principal polarization axes non-orthogonal.

3. In long lengths of transmission fiber, both the magnitude and direction of the PMD vector vary stochastically with optical frequency. In particular, the PMD vector magnitude follows a Maxwellian distribution with mean and rms values τ_{mean} and $\tau_{\text{rms}} = \tau_{\text{mean}}\sqrt{3\pi/8}$ respectively. The decorrelation bandwidth of the PMD vector is related to the rms DGD through $\omega_d = \sqrt{3}/\tau_{\text{rms}}$.

4. Concatenation rules for the accumulation of PMD and PDL with an arbitrary number of optical elements provide a useful analytic and numerical tool for evaluating polarization evolution in optical systems. For finite N, the statistical properties of the concatenation can differ substantially from the $N \to \infty$ limit.

5. Finally, PMD and PDL emulators may be preconditioned to an arbitrary initial state with Brownian bridge algorithm. This numerical technique can reduce simulation run times by orders of magnitude as compared to standard Monte Carlo calculations in which the probability of occurrence of the initial emulator state is otherwise extremely low.

The remaining chapters of this thesis are organized as follows. In Chapter 2, I will discuss the Magnus expansion of PMD and PDL that was introduced for the first time in

Ref. [94] and will establish its accuracy to fifth order in frequency. I will also present for the first time a concatenation rule for the complex principal state vector, $\vec{W}$, applicable to systems with both PMD and PDL and will develop a novel method for the exact numerical evaluation of arbitrary derivatives of $\vec{W}$.

Chapter 3 presents an experimental measurement of high-frequency polarization transients resulting from low amplitude mechanical impacts of a commercially available DCM that are unexpectedly reproducible between successive measurements. This allows measurements of the transient evolution of PMD on time scales that are not otherwise possible on most commercially available PMD measurement systems. Mechanically induced polarization transients have since been studied experimentally in Ref. [28] and using the physical models developed in Ref. [111].

In Chapter 4, I will introduce a model of the temporal evolution of PMD based upon the formalism of a uniform random walk over the Poincaré sphere that, while being straightforward to implement numerically, provides a simple relation between the model parameters (angular step size and the number of emulator segments N) and the decorrelation time of the fiber emulator. I will then demonstrate a novel modification of the PMD emulator that allows the temporal dynamics of the PMD emulator to be biased towards low-probability regions of polarization activity using multicanonical sampling algorithms.

Next, in Chapter 5, I will adapt the basic transition matrix procedure to the temporal dynamics of PMD by associating each transition in the Markov chain with a simulated temporal step, leading to the accurate reproduction of the temporal dynamics of PMD induced system outages using far fewer samples. In Chapter 6, I examine the impact of the acceptance rule enforced between adjacent states of the Markov chain on the accuracy of the resulting pdf estimate and will show that for a suitable range of input parameters simple modifications of the decision rule can significantly enhance algorithmic efficiency. Chapter 7 illustrates the application of the transition matrix procedure to the analysis of

the outage dynamics associated with the hinge model of polarization dynamics. Finally, I establish the more general applicability of the transition matrix method by calculating the pdf of the duration of fading events in a wireless communication system in Chapter 8.

Chapter 2

The Accuracy of the Magnus Expansion for Polarization Mode Dispersion and Polarization Dependent Loss

We employ the Magnus expansion to solve the differential equation for the frequency dependence of the Mueller matrix in the presence of polarization mode dispersion and polarization dependent loss. We then compare this solution with the results of previous authors and establish its accuracy to fifth order in frequency.

2.1 Introduction

The input and output Stokes vectors [47] in a linear optical system such as a fiber or waveguide are related by a rotation matrix in the absence of polarization dependent loss

(PDL). This matrix is often described by an expansion in polarization mode dispersion (PMD) vectors of increasing order [64]. Here, we generalize this procedure to the 4×4 Mueller matrix transformation for systems with both PMD and PDL. This matrix obeys a simple differential equation involving only the real and imaginary components of the complex principal state vector [25, 55, 96, 98] that we solve with the Magnus expansion. Unlike other procedures, truncating this solution yields approximations that preserve the group properties of the exact expression [78, 87, 102, 103]. In the absence of PDL, we will further demonstrate that the Magnus expansion [98] agrees with previous power series formulations [47, 51, 64] for the frequency variation of the Stokes vector, while a detailed numerical analysis indicates that for a specified model order the Magnus expansion maximizes the bandwidth of high Jones matrix estimation accuracy and therefore can efficiently invert the transformation between the frequency dependent PMD vector and the Jones/Mueller matrix. We also discuss the relationship between higher-order PMD and PDL effects, which are described with both 4×4 Mueller and 2×2 Jones matrices.

2.2 Mueller and Jones matrix formalisms

2.2.1 The Magnus expansion

In a linear optical system characterized by two orthogonally polarized guided modes with different group velocities and losses (PMD and PDL), the Jones vectors of the output and the frequency-independent input electric field polarizations, $|A_{out}(\omega)\rangle$ and $|A_{in}\rangle$, at an optical frequency ω are related through a 2×2 complex transfer matrix $|A_{out}(\omega)\rangle = \mathbf{T}(\omega)|A_{in}\rangle$ (as this chapter only considers attenuation and phase differences between the two polarizations, we further set $\det(\mathbf{T}) = 1$ [55]). The matrix $\mathbf{T}(\omega)$ satisfies [56]

$$\frac{\partial \mathbf{T}}{\partial \omega}\mathbf{T}^{-1} = -\frac{i}{2}\vec{W}(\omega) \cdot \vec{\boldsymbol{\sigma}} \tag{2.1}$$

in which $\vec{W}(\omega) = \vec{\Omega}(\omega) + i\vec{\Lambda}(\omega)$ is termed the complex principal state vector (PSV). The corresponding input and output 4×1 Stokes vectors for a completely polarized incident electric field

$$\begin{aligned} \tilde{A}_{in} &= \begin{bmatrix} \langle A_{in}|A_{in}\rangle \\ \langle A_{in}|\vec{\boldsymbol{\sigma}}|A_{in}\rangle \end{bmatrix} \\ \tilde{A}_{out} &= \begin{bmatrix} \langle A_{out}|A_{out}\rangle \\ \langle A_{out}|\vec{\boldsymbol{\sigma}}|A_{out}\rangle \end{bmatrix} \end{aligned} \tag{2.2}$$

are similarly related through

$$\tilde{A}_{out}(\omega) = \mathbf{M}(\omega)\, \tilde{A}_{in} \tag{2.3}$$

in which

$$M_{ab}(\omega) = \frac{1}{2}\ \text{Trace}\left\{\, \mathbf{T}^{\dagger}(\omega)\, \boldsymbol{\sigma}_a\, \mathbf{T}(\omega)\, \boldsymbol{\sigma}_b \,\right\} \tag{2.4}$$

designate the components of the 4×4 Mueller matrix $\mathbf{M}(\omega)$, $a, b = 0, 1, 2, 3$, and $\vec{\boldsymbol{\sigma}} = [\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, \boldsymbol{\sigma}_3]^{\mathrm{T}}$ is the vector of Pauli spin matrices while $\boldsymbol{\sigma}_0 = \mathbf{I}_2$ is the 2×2 identity matrix [39, 47].

Equivalently, the Mueller matrix can be expressed as a Kronecker matrix product [2, 6, 21, 98]

$$\mathbf{M}(\omega) = \mathbf{A}\, [\mathbf{T}(\omega) \otimes \mathbf{T}^{*}(\omega)]\, \mathbf{A}^{\dagger} \tag{2.5}$$

in which

$$\mathbf{A} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & i & -i & 0 \end{bmatrix} \tag{2.6}$$

is unitary, while $*$ and $\dagger$ represent complex and hermitian conjugation, respectively. Defining

$$\begin{bmatrix} 0 & \vec{a}^{\mathrm{T}} \\ \vec{a} & \vec{b}\times \end{bmatrix} = \begin{bmatrix} 0 & a_1 & a_2 & a_3 \\ a_1 & 0 & -b_3 & b_2 \\ a_2 & b_3 & 0 & -b_1 \\ a_3 & -b_2 & b_1 & 0 \end{bmatrix} \tag{2.7}$$

for any two vectors $\vec{a}$ and $\vec{b}$ [48], it follows

$$\mathbf{A}\,[\,(\vec{a}\cdot\vec{\boldsymbol{\sigma}})\otimes\mathbf{I}_2\,]\,\mathbf{A}^{\dagger} = \begin{bmatrix} 0 & \vec{a}^{\mathrm{T}} \\ \vec{a} & i\vec{a}\times \end{bmatrix} \tag{2.8}$$

while

$$\mathbf{A}\,[\,\mathbf{I}_2\otimes(\vec{a}\cdot\vec{\boldsymbol{\sigma}})^{*}\,]\,\mathbf{A}^{\dagger} = \begin{bmatrix} 0 & \vec{a}^{\mathrm{T}} \\ \vec{a} & -i\vec{a}\times \end{bmatrix}. \tag{2.9}$$

Accordingly, in terms of $\vec{\Omega}$ and $\vec{\Lambda}$ differentiation of Eq. (2.5) yields [25, 96, 98]

$$\frac{\partial\mathbf{M}}{\partial\omega}\mathbf{M}^{-1} = \mathbf{H}(\omega) \tag{2.10}$$

$$\mathbf{H}(\omega) = \begin{bmatrix} 0 & \vec{\Lambda}^{\mathrm{T}} \\ \vec{\Lambda} & \vec{\Omega}\times \end{bmatrix}. \tag{2.11}$$

The matrix $\mathbf{H}(\omega)$ can then be considered a generator of an infinitesimal Lorentz transformation in Stokes space [6, 48, 95].

The variation of $\mathbf{M}(\omega)$ with frequency can now be determined iteratively in terms of $\mathbf{M}(\omega_0)$ at the optical carrier frequency, ω_0, according to [103, 126]

$$\mathbf{M}(\omega) = \mathbf{M}(\omega_0) + \int_{\omega_0}^{\omega} \mathrm{d}\omega_1 \mathbf{H}(\omega_1)\mathbf{M}(\omega_1). \tag{2.12}$$

This yields

$$\mathbf{M}(\omega) = [\mathbf{I}_4 + \mathbf{J}_1 + \mathbf{J}_2 + \mathbf{J}_3 + \ldots]\,\mathbf{M}(\omega_0) \tag{2.13}$$

where $\mathbf{I}_N$ is the $N \times N$ identity matrix and

$$\mathbf{J}_n(\omega) = \int_{\omega_0}^{\omega} \mathrm{d}\omega_1 \cdots \int_{\omega_0}^{\omega_{n-1}} \mathrm{d}\omega_n \; \mathbf{H}(\omega_1) \cdots \mathbf{H}(\omega_n) \tag{2.14}$$

are the series solution coefficients of Eq. (2.10). On the other hand, as demonstrated in subsequent sections, the frequency interval over which Eq. (2.13) approximates $\mathbf{M}$ to within a given level of accuracy is improved if we instead truncate the Magnus expansion

$$\mathbf{M}(\omega) = \exp\left[\mathbf{B}_1 + \mathbf{B}_2 + \mathbf{B}_3 + \ldots\right] \mathbf{M}(\omega_0) \tag{2.15}$$

to a given number of terms, which preserves relevant group properties of the exact solution [78,87,98,103]. To fourth order in $\mathbf{H}$ the expansion coefficients $\mathbf{J}_n$ and $\mathbf{B}_n$ of Eq. (2.13) and Eq. (2.15) are related by

$$\begin{aligned}
\mathbf{B}_1 &= \mathbf{J}_1 \\
\mathbf{B}_2 &= \mathbf{J}_2 - \frac{\mathbf{J}_1^2}{2} \\
\mathbf{B}_3 &= \mathbf{J}_3 - \frac{1}{2}(\mathbf{J}_1\mathbf{J}_2 + \mathbf{J}_2\mathbf{J}_1) + \frac{\mathbf{J}_1^3}{3} \\
\mathbf{B}_4 &= \mathbf{J}_4 - \frac{1}{2}(\mathbf{J}_1\mathbf{J}_3 + \mathbf{J}_3\mathbf{J}_1 + \mathbf{J}_2\mathbf{J}_2) \\
&\quad + \frac{1}{3}(\mathbf{J}_1\mathbf{J}_1\mathbf{J}_2 + \mathbf{J}_1\mathbf{J}_2\mathbf{J}_1 + \mathbf{J}_2\mathbf{J}_1\mathbf{J}_1) - \frac{\mathbf{J}_1^4}{4}
\end{aligned} \tag{2.16}$$

with analogous formulas for higher orders [102]. Rewriting $\mathbf{B}_n$ in terms of the commutators $[\mathbf{H}(\omega_i), \mathbf{H}(\omega_j)]$ of $\mathbf{H}(\omega)$ at two optical frequencies ω_i and ω_j yields an alternative,

computationally convenient formula for the Magnus expansion coefficients [87, 102]

$$
\begin{aligned}
\mathbf{B}_1 &= \int_{\omega_0}^{\omega} \mathrm{d}\omega_1 \mathbf{H}(\omega_1) \\
\mathbf{B}_2 &= \frac{1}{2} \int_{\omega_0}^{\omega} \mathrm{d}\omega_1 \int_{\omega_0}^{\omega_1} \mathrm{d}\omega_2 \left[\mathbf{H}(\omega_1), \mathbf{H}(\omega_2)\right] \\
\mathbf{B}_3 &= \frac{1}{6} \int_{\omega_0}^{\omega} \mathrm{d}\omega_1 \int_{\omega_0}^{\omega_1} \mathrm{d}\omega_2 \int_{\omega_0}^{\omega_2} \mathrm{d}\omega_3 \\
&\qquad \Big(\left[\mathbf{H}(\omega_1), \left[\mathbf{H}(\omega_2), \mathbf{H}(\omega_3)\right]\right] + \left[\left[\mathbf{H}(\omega_1), \mathbf{H}(\omega_2)\right], \mathbf{H}(\omega_3)\right] \Big) \\
\mathbf{B}_4 &= \frac{1}{12} \int_{\omega_0}^{\omega} \mathrm{d}\omega_1 \int_{\omega_0}^{\omega_1} \mathrm{d}\omega_2 \int_{\omega_0}^{\omega_2} \mathrm{d}\omega_3 \int_{\omega_0}^{\omega_3} \mathrm{d}\omega_4 \\
&\qquad \Big(\left[\mathbf{H}(\omega_1), \left[\left[\mathbf{H}(\omega_2), \mathbf{H}(\omega_3)\right], \mathbf{H}(\omega_4)\right]\right] + \left[\left[\mathbf{H}(\omega_1), \left[\mathbf{H}(\omega_2), \mathbf{H}(\omega_3)\right]\right], \mathbf{H}(\omega_4)\right] \\
&\qquad\quad + \left[\left[\mathbf{H}(\omega_1), \mathbf{H}(\omega_2)\right], \left[\mathbf{H}(\omega_3), \mathbf{H}(\omega_4)\right]\right] \Big).
\end{aligned}
\tag{2.17}
$$

If the matrix $\mathbf{H}(\omega)$ is now expanded in a Taylor series

$$
\mathbf{H}(\omega) = \sum_{n=0}^{\infty} \mathbf{H}^{(n)} \frac{\Delta\omega^n}{n!}, \tag{2.18}
$$

where the coefficient $\mathbf{H}^{(n)} = [\partial^n \mathbf{H}/\partial\omega^n]|_{\omega_0}$ is the n:th order PMD/PDL matrix, we obtain

to fifth order in $\Delta\omega = \omega - \omega_0$ from the Magnus coefficients, Eq. (2.17),

$$\begin{aligned}
\mathbf{M}(\omega) &= \exp\left[\mathbf{N}(\Delta\omega)\right]\, \mathbf{M}(\omega_0) \\
\\
\mathbf{N}(\Delta\omega) &= \mathbf{H}^{(0)}\Delta\omega + \mathbf{H}^{(1)}\frac{\Delta\omega^2}{2!} \\
&\quad + (\mathbf{H}^{(2)} - \frac{1}{2}\,[\mathbf{H}^{(0)}, \mathbf{H}^{(1)}])\frac{\Delta\omega^3}{3!} \\
&\quad + (\mathbf{H}^{(3)} - [\mathbf{H}^{(0)}, \mathbf{H}^{(2)}])\frac{\Delta\omega^4}{4!} \\
&\quad + \Big(\mathbf{H}^{(4)} - [\mathbf{H}^{(1)}, \mathbf{H}^{(2)}] - \frac{3}{2}[\mathbf{H}^{(0)}, \mathbf{H}^{(3)}] \\
&\quad + \frac{1}{6}[\mathbf{H}^{(0)}, [\mathbf{H}^{(0)}, \mathbf{H}^{(2)}]] + \frac{1}{2}[\mathbf{H}^{(1)}, [\mathbf{H}^{(1)}, \mathbf{H}^{(0)}]] \\
&\quad + \frac{1}{6}[\mathbf{H}^{(0)}, [\mathbf{H}^{(0)}, [\mathbf{H}^{(0)}, \mathbf{H}^{(1)}]]]\Big)\frac{\Delta\omega^5}{5!} \\
&\quad + O(\Delta\omega^6).
\end{aligned} \tag{2.19}$$

Alternatively, Eq. (2.11) expresses the Mueller matrix directly in terms of the vectors $\vec{\Omega}$ and $\vec{\Lambda}$. To fourth order,

$$\mathbf{M}(\omega) = \exp\begin{bmatrix} 0 & \vec{a}^{\,\mathrm{T}} \\ \vec{a} & \vec{b}\times \end{bmatrix} \mathbf{M}(\omega_0) \tag{2.20}$$

in which $\vec{a}$ and $\vec{b}$ designate

$$\begin{aligned}
\vec{a} &= \vec{\Lambda}^{(0)}\Delta\omega + \vec{\Lambda}^{(1)}\frac{\Delta\omega^2}{2!} \\
&\quad + (\vec{\Lambda}^{(2)} - \frac{1}{2}\,\vec{\Omega}^{(0)} \times \vec{\Lambda}^{(1)} - \frac{1}{2}\,\vec{\Lambda}^{(0)} \times \vec{\Omega}^{(1)})\frac{\Delta\omega^3}{3!} \\
&\quad + (\vec{\Lambda}^{(3)} - \vec{\Omega}^{(0)} \times \vec{\Lambda}^{(2)} - \vec{\Lambda}^{(0)} \times \vec{\Omega}^{(2)})\frac{\Delta\omega^4}{4!} + O(\Delta\omega^5)
\end{aligned} \tag{2.21}$$

and

$$\begin{aligned}\vec{b} &= \vec{\Omega}^{(0)}\Delta\omega + \vec{\Omega}^{(1)}\frac{\Delta\omega^2}{2!} \\ &+ (\vec{\Omega}^{(2)} - \frac{1}{2}\,\vec{\Omega}^{(0)} \times \vec{\Omega}^{(1)} - \frac{1}{2}\,\vec{\Lambda}^{(1)} \times \vec{\Lambda}^{(0)})\frac{\Delta\omega^3}{3!} \\ &+ (\vec{\Omega}^{(3)} - \vec{\Omega}^{(0)} \times \vec{\Omega}^{(2)} - \vec{\Lambda}^{(2)} \times \vec{\Lambda}^{(0)})\frac{\Delta\omega^4}{4!} + O(\Delta\omega^5).\end{aligned} \tag{2.22}$$

We observe that $\mathbf{M}(\omega)$ as given by Eqs. (2.20)-(2.22) includes previously unnoticed correlations between the Taylor orders of $\vec{\Lambda}$ and $\vec{\Omega}$.

Additionally, since any Jones matrix of the form $\mathbf{T}(\omega) = \exp[\mathbf{K}]\,\mathbf{T}(\omega_0)$, for some 2×2 complex matrix $\mathbf{K}$, can be transformed according to Eq. (2.5) into the Mueller matrix [6,98]

$$\mathbf{M}(\omega) = \exp\left[\mathbf{A}\ (\mathbf{K}\otimes\mathbf{I}_2 + \mathbf{I}_2\otimes\mathbf{K}^*)\ \mathbf{A}^\dagger\right]\mathbf{M}(\omega_0) \tag{2.23}$$

we conclude from Eq. (2.19) that

$$\mathbf{T}(\omega) = \exp\left[-\frac{i}{2}\,\vec{n}(\Delta\omega)\cdot\vec{\boldsymbol{\sigma}}\right]\ \mathbf{T}(\omega_0), \tag{2.24}$$

where

$$\begin{aligned}\vec{n}(\Delta\omega) = \vec{W}^{(0)}\Delta\omega &+ \vec{W}^{(1)}\frac{\Delta\omega^2}{2!} \\ &+\ (\vec{W}^{(2)} - \frac{1}{2}\,\vec{W}^{(0)} \times \vec{W}^{(1)})\frac{\Delta\omega^3}{3!} \\ &+\ (\vec{W}^{(3)} - \vec{W}^{(0)} \times \vec{W}^{(2)})\frac{\Delta\omega^4}{4!} \\ &+\ \Big(\vec{W}^{(4)} - \vec{W}^{(1)} \times \vec{W}^{(2)} - \frac{3}{2}\vec{W}^{(0)} \times \vec{W}^{(3)} \\ &+\ \frac{1}{6}\vec{W}^{(0)} \times (\vec{W}^{(0)} \times \vec{W}^{(2)}) \\ &+\ \frac{1}{2}\vec{W}^{(1)} \times (\vec{W}^{(1)} \times \vec{W}^{(0)}) \\ &+\ \frac{1}{6}\vec{W}^{(0)} \times (\vec{W}^{(0)} \times (\vec{W}^{(0)} \times \vec{W}^{(1)}))\Big)\frac{\Delta\omega^5}{5!} \\ &+\ O(\Delta\omega^6)\end{aligned} \tag{2.25}$$

in which $\vec{W}_n = [\partial^n \vec{W}/\partial\omega^n]|_{\omega_0}$, or, in terms of $\vec{\Omega}$ and $\vec{\Lambda}$, cf. Eqs. (2.21)-(2.22),

$$\mathbf{T}(\omega) = \exp\left[-\frac{i}{2}(\vec{b} + i\vec{a}) \cdot \vec{\boldsymbol{\sigma}}\right] \mathbf{T}(\omega_0). \tag{2.26}$$

Finally, we observe that the matrix exponential of Eq. (2.20) can be computed efficiently for known values of $\vec{a}$ and $\vec{b}$ by defining the 2×2 matrix

$$\mathbf{Q} = \left(\vec{b} + i\vec{a}\right) \cdot \vec{\boldsymbol{\sigma}} \tag{2.27}$$

and complex variable

$$\lambda = \sqrt{(\vec{b} + i\vec{a}) \cdot (\vec{b} + i\vec{a})}. \tag{2.28}$$

The Jones matrix of Eq. (2.26) can then be expanded as

$$\begin{aligned} \mathbf{F} &= \exp\left[-\frac{i}{2}(\vec{b} + i\vec{a}) \cdot \vec{\boldsymbol{\sigma}}\right] \\ &= \cos\left(\frac{\lambda}{2}\right)\mathbf{I}_2 - \frac{i}{\lambda}\sin\left(\frac{\lambda}{2}\right)\mathbf{Q} \end{aligned} \tag{2.29}$$

with the final result

$$\exp\begin{bmatrix} 0 & \vec{a}^{\,\mathrm{T}} \\ \vec{a} & \vec{b}\times \end{bmatrix} = \mathbf{A}\,(\mathbf{F} \otimes \mathbf{F}^*)\,\mathbf{A}^\dagger, \tag{2.30}$$

providing a useful formula for the numerical evaluation of the 4×4 exponential Mueller matrix. In particular, applying Eq. (2.30) with $\vec{a} = 0$ and $\vec{b} = \psi\,\hat{p}$ yields

$$\exp\begin{bmatrix} 0 & 0 \\ 0 & \psi\,\hat{p}\times \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \cos\psi\,\mathbf{I}_3 + (1 - \cos\psi)\,\hat{p}\hat{p}^{\mathrm{T}} + \sin\psi\,\hat{p}\times \end{bmatrix} \tag{2.31}$$

which clearly reproduces the known result for the 3×3 Stokes rotation matrix $\exp\bigl(\psi\,\hat{p}\times\bigr)$. Similarly, for $\vec{a} = \alpha\,\hat{r}$ and $\vec{b} = 0$,

$$\exp\begin{bmatrix} 0 & \alpha\,\hat{r}^{\mathrm{T}} \\ \alpha\,\hat{r} & 0 \end{bmatrix} = \begin{bmatrix} \cosh\alpha & \sinh\alpha\,\hat{r}^{\mathrm{T}} \\ \sinh\alpha\,\hat{r} & \mathbf{I}_3 + (\cosh\alpha - 1)\,\hat{r}\hat{r}^{\mathrm{T}} \end{bmatrix} \tag{2.32}$$

is the 4×4 Mueller matrix representation of the PDL Jones matrix $\exp\bigl(\alpha\,\hat{r} \cdot \vec{\boldsymbol{\sigma}}/2\bigr)$.

2.2.2 Exponential series

In Ref. [51], the Jones matrix associated with PMD was evaluated to fourth-order in the frequency offset $\Delta\omega$. While, as we will show below, this procedure is equivalent to the series expansion of Eq. (2.26) the series when truncated to finite order is non-unitary in the absence of PDL and therefore is accurate within a somewhat reduced frequency interval.

Specializing to the case of zero PDL, and setting $\vec{b} = \psi\,\hat{r}$ with $\vec{\Lambda} = 0$, Eq. (2.26) can be written in terms of the unitary Jones matrix

$$\begin{aligned}\mathbf{U}(\Delta\omega) &= \mathbf{T}(\omega)\mathbf{T}(\omega_0)^{-1} \\ &= \cos\left(\frac{\psi}{2}\right)\mathbf{I}_2 - i\sin\left(\frac{\psi}{2}\right)(\hat{r}\cdot\vec{\boldsymbol{\sigma}}) \\ &= u_0\,\mathbf{I}_2 - i\vec{u}\cdot\vec{\boldsymbol{\sigma}},\end{aligned} \tag{2.33}$$

in which we have introduced the vector $\vec{u}$ and scalar u_0, components of the Stokes vector formed from the elements of the Jones matrix $\mathbf{U}$ according to $u_0 = \text{Trace}\,\{\mathbf{U}\}/2$ and $\vec{u} = i\,\text{Trace}\,\{\mathbf{U}\,\vec{\boldsymbol{\sigma}}\}/2$.

In terms of the unit vector $\hat{p} = \vec{\Omega}^{(0)}/|\vec{\Omega}^{(0)}|$, we find after expanding the quantities u_0 and $\vec{u}$ in a power series, cf. Eq. (2.22), and identifying terms of equal orders in $\Delta\omega$,

$$\begin{aligned}u_0 &= 1 - |\vec{\Omega}^{(0)}|^2\frac{\Delta\omega^2}{8} - \vec{\Omega}^{(0)}\cdot\vec{\Omega}^{(1)}\frac{\Delta\omega^3}{8} \\ &\quad + \left(\frac{1}{16}|\vec{\Omega}^{(0)}|^4 - \frac{3}{4}|\vec{\Omega}^{(1)}|^2 - \vec{\Omega}^{(0)}\cdot\vec{\Omega}^{(2)}\right)\frac{\Delta\omega^4}{24} + \ldots \\ \hat{p}\cdot\vec{u} &= |\vec{\Omega}^{(0)}|\frac{\Delta\omega}{2} + \hat{p}\cdot\vec{\Omega}^{(1)}\frac{\Delta\omega^2}{4} \\ &\quad + \left(\hat{p}\cdot\vec{\Omega}^{(2)} - \frac{1}{4}\,|\vec{\Omega}^{(0)}|^3\right)\frac{\Delta\omega^3}{12} \\ &\quad + \left(\hat{p}\cdot\vec{\Omega}^{(3)} - \frac{3}{2}\,|\vec{\Omega}^{(0)}|\,\vec{\Omega}^{(0)}\cdot\vec{\Omega}^{(1)}\right)\frac{\Delta\omega^4}{48} + \ldots\end{aligned} \tag{2.34}$$

$$
\begin{aligned}
\vec{u}_{\perp} &= \vec{u} - \left[\hat{p}\cdot\vec{u}\right]\hat{p} \\
&= \vec{\Omega}^{(1\perp)}\,\frac{\Delta\omega^2}{4} + \left(\vec{\Omega}^{(2\perp)} - \frac{1}{2}\,\vec{\Omega}^{(0)}\times\vec{\Omega}^{(1\perp)}\right)\frac{\Delta\omega^3}{12} \\
&+ \left(\vec{\Omega}^{(3\perp)} - \vec{\Omega}^{(0)}\times\vec{\Omega}^{(2)} - \frac{1}{2}|\vec{\Omega}^{(0)}|^2\,\vec{\Omega}^{(1\perp)}\right)\frac{\Delta\omega^4}{48} + \ldots
\end{aligned}
$$

with $\vec{\Omega}^{(n\perp)} = \vec{\Omega}^{(n)} - \left[\hat{p}\cdot\vec{\Omega}^{(n)}\right]\hat{p}$. The results of Ref. [51] are reproduced by substituting $\vec{\tau} = \vec{\Omega}/2$ indicating that the Magnus formula, Eq. (2.24), accurately represents the higher-order frequency variation of the Jones (Mueller) matrix.

2.2.3 Operator symmetrization

To recast the Magnus solution of Eq. (2.19) into a product of exponential operators each of which corresponds to a pure n:th-order PMD/PDL component, the various non-commuting operators must be correctly symmetrized to preserve the commutation relations between the different matrix orders of $\mathbf{H}^{(n)}$ in Eq. (2.19) [98,124,126]. The expansions that contain a minimum number of operators up to and including fourth order in $\Delta\omega$ can be derived from the recursive form of the Magnus coefficients, Eq. (2.16), through repeated application

of the Baker-Campbell-Hausdorff (BCH) identity [44, 98] and are given by [23, 45, 46, 118]

$$\mathbf{M}(\omega)\mathbf{M}(\omega_0)^{-1} = \exp\left[\mathbf{H}^{(0)}\Delta\omega\right] + O(\Delta\omega^2) \tag{2.35}$$

$$= \exp\left[\frac{1}{2}\mathbf{H}^{(0)}\Delta\omega\right]\exp\left[\frac{1}{2}\mathbf{H}^{(1)}\Delta\omega^2\right]\times \exp\left[\frac{1}{2}\mathbf{H}^{(0)}\Delta\omega\right] + O(\Delta\omega^3) \tag{2.36}$$

$$= \exp\left[\frac{1}{3}\mathbf{H}^{(0)}\Delta\omega\right]\exp\left[\frac{1}{2}\,\mathbf{H}^{(1)}\Delta\omega^2\right]\times \exp\left[\frac{2}{3}\mathbf{H}^{(0)}\Delta\omega\right]\exp\left[\frac{1}{6}\mathbf{H}^{(2)}\Delta\omega^3\right] + O(\Delta\omega^4) \tag{2.37}$$

$$= \exp\left[\frac{1}{6}\mathbf{H}^{(1)}\Delta\omega^2\right]\exp\left[\frac{1}{4}\mathbf{H}^{(0)}\Delta\omega\right]\times \exp\left[\frac{1}{6}\mathbf{H}^{(2)}\Delta\omega^3\right]\exp\left[\frac{1}{4}\mathbf{H}^{(0)}\Delta\omega\right]\times \exp\left[\frac{1}{3}\mathbf{H}^{(1)}\Delta\omega^2\right]\exp\left[\frac{1}{2}\mathbf{H}^{(0)}\Delta\omega\right]\times \exp\left[\frac{1}{24}\mathbf{H}^{(3)}\Delta\omega^4\right] + O(\Delta\omega^5). \tag{2.38}$$

2.2.4 Non-symmetric model

In Section 2.4 we compare numerically the accuracy of the symmetrized to the non-symmetrized Jones matrix operator expansion [30, 64, 130]

$$\mathbf{T}(\omega)\mathbf{T}(\omega_0)^{-1} \approx \mathbf{U}_n\mathbf{U}_{n-1}\cdots\mathbf{U}_2\mathbf{U}_1 \tag{2.39}$$

in which each matrix

$$\mathbf{U}_n(\Delta\omega) = \exp\left[-\frac{i}{2}\frac{\Delta\omega^n}{n!}\vec{p}^{(n)}\cdot\vec{\boldsymbol{\sigma}}\right] \tag{2.40}$$

is associated as in Section 2.2.3 with a pure n:th order PMD component with corresponding PMD vector [64]

$$\frac{\partial \mathbf{U}_n}{\partial \omega}\mathbf{U}_n^{-1} = -\frac{i}{2}\left[\frac{\Delta\omega^{n-1}}{(n-1)!}\,\vec{p}^{(n)}\right]\cdot\vec{\boldsymbol{\sigma}}. \tag{2.41}$$

To ensure $\vec{\Omega}(\omega)\cdot\vec{\boldsymbol{\sigma}} = 2i\left(\partial\mathbf{T}/\partial\omega\right)\mathbf{T}^{-1}$ to order $\Delta\omega^4$ therefore requires [64]

$$\begin{aligned}
\vec{p}^{(1)} &= \vec{\Omega}^{(0)} \\
\vec{p}^{(2)} &= \vec{\Omega}^{(1)} \\
\vec{p}^{(3)} &= \vec{\Omega}^{(2)} + \vec{p}^{(1)}\times\vec{p}^{(2)} \\
\vec{p}^{(4)} &= \vec{\Omega}^{(3)} + \vec{p}^{(1)}\times\vec{p}^{(3)} \\
\vec{p}^{(5)} &= \vec{\Omega}^{(4)} + \vec{p}^{(1)}\times\vec{p}^{(4)} - 3\,\vec{p}^{(2)}\times\left(\vec{p}^{(2)}\times\vec{p}^{(1)}\right) + 4\,\vec{p}^{(2)}\times\vec{p}^{(3)}.
\end{aligned} \tag{2.42}$$

2.3 Emulator model for PMD and PDL

The frequency variation of the polarization vector of an optical system with PMD and PDL can be modeled by dividing the system into a large number of constant birefringent and lossy segments. In each segment, the light polarization precesses around the axis of the section's polarization dispersion vector at a rate determined by the magnitude of the birefringence while the attenuation is a maximum and minimum in two generally different orthogonal directions.

We consider an optical system with mean PDL $0 \le \alpha_{\text{mean}} \le 10$ dB and differential group delay (DGD) $\tau_{\text{mean}} = 20$ ps (as calculated for 0 dB mean PDL). To represent this by an $N = 100$ section emulator, we set the frequency-independent PMD vector magnitude of the m:th section and the differential attenuation between the most and least attenuated polarization states to $|\vec{\tau}_m| = \tau_{\text{mean}}\sqrt{3\pi/8N} = 2.17$ ps and $0 \le |\vec{\alpha}_m| \le 1$ dB, respectively. The orientation of the birefringent and attenuation axes of the m:th section, specified by

the unit vectors $\vec{\tau}_m/|\vec{\tau}_m|$ and $\vec{\alpha}_m/|\vec{\alpha}_m|$, respectively, are selected from a uniform statistical distribution over the Poincaré sphere.

Denoting the 2×2 Jones and a complex 3×3 Stokes matrices of the m:th emulator section, $m = 1, 2, \ldots, N$, by

$$\mathbf{V}_m(\omega) = \exp\left[\frac{1}{2}\,\vec{\alpha}_m \cdot \vec{\boldsymbol{\sigma}}\right]\,\exp\left[-\frac{i}{2}\,\omega\,\vec{\tau}_m \cdot \vec{\boldsymbol{\sigma}}\right]$$

$$\mathbf{R}_m(\omega) = \exp\left[i\,\vec{\alpha}_m\times\right]\,\exp\left[\omega\,\vec{\tau}_m\times\right] \tag{2.43}$$

the corresponding complex PSV,

$$\frac{\partial \mathbf{V}_m}{\partial \omega}\mathbf{V}_m^{-1} = -\frac{i}{2}\vec{v}_m \cdot \vec{\boldsymbol{\sigma}} \tag{2.44}$$

of the m:th section can be written as

$$\vec{v}_m = \exp\left[i\,\vec{\alpha}_m\times\right]\vec{\tau}_m. \tag{2.45}$$

Here, for any real valued α and unit vector $\hat{r}$

$$\exp\left[i\,\alpha\,\hat{r}\times\right] = \cos\alpha\,\mathbf{I}_3 + (1-\cos\alpha)\,\hat{r}\hat{r}^{\mathrm{T}} + i\sin\alpha\,\hat{r}\,\times\,. \tag{2.46}$$

Finally, the Jones matrix, $\mathbf{T}_N(\omega)$, after N emulator sections is calculated recursively according to

$$\mathbf{T}_N(\omega) = \mathbf{V}_N(\omega)\mathbf{T}_{N-1}(\omega) \tag{2.47}$$

with $\mathbf{T}_0(\omega) = \mathbf{I}_2$, while a concatenation rule for the total complex PSV, $\vec{W}_N(\omega)$, after N

emulator segments can be derived by considering

$$
\begin{aligned}
-\frac{i}{2}\vec{W}_N(\omega)\cdot\vec{\boldsymbol{\sigma}} &= \frac{\partial \mathbf{T}_N}{\partial\omega}\,\mathbf{T}_N^{-1} \\
&= \frac{\partial \mathbf{V}_N}{\partial\omega}\,\mathbf{V}_N^{-1} + \mathbf{V}_N\left(\frac{\partial \mathbf{T}_{N-1}}{\partial\omega}\,\mathbf{T}_{N-1}^{-1}\right)\mathbf{V}_N^{-1} \\
&= -\frac{i}{2}\vec{v}_N\cdot\vec{\boldsymbol{\sigma}} + \mathbf{V}_N\left(-\frac{i}{2}\vec{W}_{N-1}\cdot\vec{\boldsymbol{\sigma}}\right)\mathbf{V}_N^{-1} \\
&= -\frac{i}{2}\vec{v}_N\cdot\vec{\boldsymbol{\sigma}} - \frac{i}{2}\left(\mathbf{R}_N\vec{W}_{N-1}\right)\cdot\vec{\boldsymbol{\sigma}} \\
&= -\frac{i}{2}\left[\vec{v}_N + \mathbf{R}_N\vec{W}_{N-1}\right]\cdot\vec{\boldsymbol{\sigma}}.
\end{aligned}
\tag{2.48}
$$

In analogy with Eq. (1.36),

$$
\vec{W}_N(\omega) = \vec{v}_N(\omega) + \mathbf{R}_N(\omega)\,\vec{W}_{N-1}(\omega) \tag{2.49}
$$

with $\vec{W}_0(\omega) = 0$, provides a general rule for the accumulation of the complex principal state vector in optical systems containing both PMD and PDL that is applicable to the analysis of the distribution and frequency autocorrelation statistics of $\vec{W}_N(\omega)$. Notice that unlike systems with PMD, the quantities $\vec{v}_m$ and $\mathbf{R}_m(\omega)$ appearing in Eq. (2.49) are complex valued.

By induction we find that the n:th derivative, $n \geq 1$, of the total complex PSV after m emulator sections, $\vec{W}_m = \vec{v}_m + \mathbf{R}_m\vec{W}_{m-1}$, is given by

$$
\begin{aligned}
\frac{\partial^n \vec{W}_m}{\partial\omega^n} &= \mathbf{R}_m\frac{\partial^n \vec{W}_{m-1}}{\partial\omega^n} \\
&\quad - \sum_{k=1}^{n}\binom{n}{k}\left[-\vec{v}_m\times\right]^k\,\frac{\partial^{n-k}\vec{W}_m}{\partial\omega^{n-k}}
\end{aligned}
\tag{2.50}
$$

with $\vec{W}_0 = 0$. Notice that this last expression provides an exact recursive formula for

evaluating $\partial^n \vec{W}_m/\partial\omega^n$. Analogously, setting $\mathbf{T}_0 = \mathbf{I}_2$,

$$\begin{aligned} \frac{\partial^n \mathbf{T}_m}{\partial\omega^n} &= \mathbf{V}_m \frac{\partial^n \mathbf{T}_{m-1}}{\partial\omega^n} \\ &- \sum_{k=1}^{n} \binom{n}{k} \left[\frac{i}{2}\vec{v}_m \cdot \vec{\boldsymbol{\sigma}}\right]^k \frac{\partial^{n-k}\mathbf{T}_m}{\partial\omega^{n-k}} \end{aligned} \tag{2.51}$$

yields an exact recursive formula for the n:th derivative of the Jones matrix $\mathbf{T}_m(\omega) = \mathbf{V}_m(\omega)\mathbf{T}_{m-1}(\omega)$ after m segments [99]. While seemingly complex, both Eqs. (2.50)-(2.51) can be easily implemented numerically [94]. The algorithm is as follows: with $m = 1$, $\vec{W}_1 = \vec{v}_1$, and $\partial^n \vec{W}_1/\partial\omega^n = 0$ for all $n \geq 1$. For $m = 2$, we calculate sequentially each $\vec{W}_2$, $\partial\vec{W}_2/\partial\omega$, $\partial^2\vec{W}_2/\partial\omega^2$, $\ldots$, $\partial^n\vec{W}_2/\partial\omega^n$ from repeated application of Eq. (2.50). Once all the $\partial^n\vec{W}_2/\partial\omega^n$ are determined, we calculate in order $\vec{W}_3$, $\partial\vec{W}_3/\partial\omega$, $\partial^2\vec{W}_3/\partial\omega^2$, $\ldots$, $\partial^n\vec{W}_3/\partial\omega^n$. These steps are then performed N times. Pseudocode implementing this algorithm is presented in Listing 2.1.

In view of the Taylor expansion of the Jones matrix in orders of $\vec{\Omega}(\omega)$ presented in Sec. 2.2.2, Eq. (2.51) will be subsequently employed in the numerical evaluation of Eqs. (2.34).

Listing 2.1: Pseudocode for the exact evaluation of $\vec{W}^{(n)} = \partial^n \vec{W}_N / \partial \omega^n$ for all derivative orders $n = 0, \ldots, N_d$ following an N section PMD/PDL emulator.

$\vec{W}^{(n)} = 0$; % *Initialize for all* $n = 0, \ldots, N_d$
for $m = 1$ to N % *Loop over emulator segments*
  $\mathbf{U}_m = \exp(\omega \vec{\tau}_m \times)$;
  $\mathbf{H}_m = \exp(i \vec{\alpha}_m \times)$;
  $\vec{v}_m = \mathbf{H}_m \vec{\tau}_m$;

  $\vec{W}^{(0)} = \vec{v}_m + \mathbf{H}_m \mathbf{U}_m \vec{W}^{(0)}$; % *First order complex PSV*
  for $n = 1$ to N_d % *Loop over order of derivative*
    $\vec{S} = 0$;
    for $k = 1$ to n
      $\vec{S} = \vec{S} - \binom{n}{k} [-\vec{v}_m \times]^k \vec{W}^{(n-k)}$;
    end
    $\vec{W}^{(n)} = \vec{S} + \mathbf{H}_m \mathbf{U}_m \vec{W}^{(n)}$; % *Complex PSV of order* $n+1$
  end
end

2.4 Numerical results

2.4.1 PMD model comparison

We now compare the procedures of Section 2.2 for zero PDL through an analysis of the maximum error, $\epsilon_\kappa(\Delta\omega)$, between the estimated $\mathbf{T}_{\text{est}}(\omega)$ and the emulated Jones matrices $\mathbf{T}(\omega)$ over the frequency interval $|\omega - \omega_0| \leq \Delta\omega$. This error is given for each realization $\kappa = 1, 2, \ldots, (N_\kappa = 10^6)$ of the emulator by

$$\epsilon_\kappa(\Delta\omega) = \max_{|\omega-\omega_0|\leq\Delta\omega} \left[\frac{\|\mathbf{T}_{\text{est}}(\omega) - \mathbf{T}(\omega)\|}{\|\mathbf{T}(\omega)\|} \right] \cdot 100\% \tag{2.52}$$

where $\|\mathbf{T}\|^2 = \text{Trace}\left\{\mathbf{T}^\dagger\mathbf{T}\right\}$ is the Frobenius matrix norm. For each emulator realization κ the bandwidth $\Delta\omega_\kappa$ is found with a bisection method applied to the condition that $\epsilon_\kappa(\Delta\omega_\kappa) = \epsilon_{\max}$ for a maximum specified error $\epsilon_{\max}$. The probability density function (pdf) of $\Delta\omega_\kappa$ is then estimated for each of the numerical techniques of Section 2.2 from the N_κ emulator realizations. Note that Eq. (2.52), unlike the Q value or bit-error rate, is independent of a specific system implementation and therefore provides an unambiguous measure of the algorithmic precision [99].

The pdf of $\Delta\omega_\kappa$ for $\epsilon_{\max} = 2\%$ as a function of the normalized frequency $\Delta\nu = \tau_{\text{mean}}\Delta\omega_\kappa/2\pi$ is displayed in Fig. 2.1, in which $\mathbf{T}_{\text{est}}(\omega)$ is approximated by the fourth-order versions of the Magnus expansion Eq. (2.24) ($\circ$), Eqs. (2.34) ($\bullet$), the symmetric expansion of Eq. (2.38) ($\times$) and the non-symmetric model, Eq. (2.39) ($+$). Here we have displayed every third data point to improve legibility. Clearly, the frequency range over which the Jones matrix can be accurately estimated is greatest for the Magnus expansion. However, Eq. (2.38) yields more precise results than Eq. (2.39) though within a more restricted frequency range than that of Eq. (2.24).

To determine the variation of the frequency range over which each model accurately represents the Jones matrix with the expansion order M, we determine as in Ref. [99] the normalized frequency $\Delta\nu_{\min}$ such that the probability $P(\Delta\nu \geq \Delta\nu_{\min})$ of $\Delta\nu$ exceeding the minimum frequency, $\Delta\nu_{\min}$, is 99.99% for values of M between 1 and 5. Fig. 2.2 then displays $\Delta\nu_{\min}$ as a function of M in which the $\circ$, $\times$, $+$ and $\bullet$ markers indicate the Magnus expansion of Eq. (2.24), the symmetric model of Eqs. (2.35)-(2.38), the non-symmetric model Eq. (2.39) and Eqs. (2.34), respectively. Of course, as a result of statistical fluctuations, the modeling error exceeds $\epsilon_{\max}$ with a probability $1 - P(\Delta\nu \geq \Delta\nu_{\min})$. However, we have previously established that the variation of $\Delta\nu_{\min}$ with expansion order M is qualitatively independent of $P(\Delta\nu \geq \Delta\nu_{\min})$ and $\epsilon_{\max}$ [99]. Therefore, while a $\Delta\nu_{\min}$ cannot be specified such that $\Delta\nu \geq \Delta\nu_{\min}$ for all possible emulator realizations, for values of $\Delta\nu$ less than ≈ 0.1 the pdf decreases rapidly in Fig. 2.1 indicating that slightly lowering

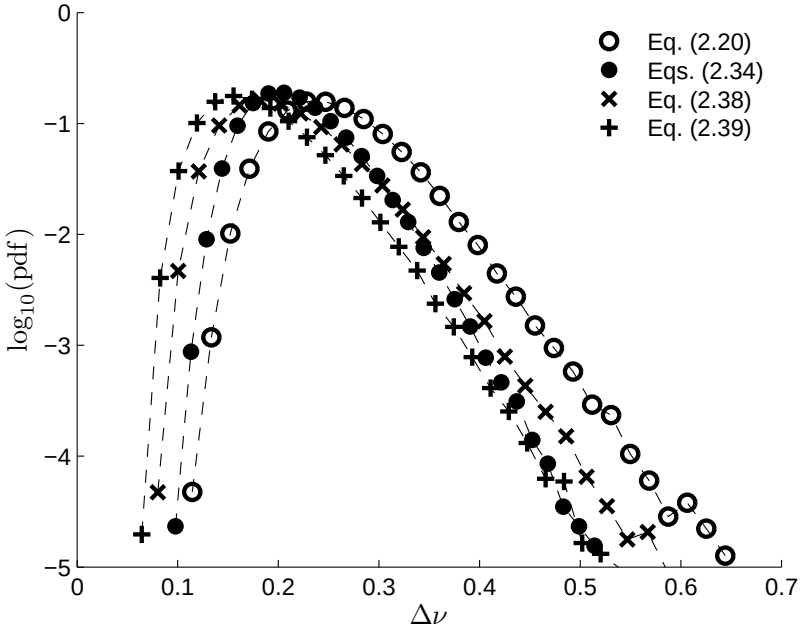


Figure 2.1: The pdf of the normalized bandwidth, $\Delta\nu = \tau_{\text{mean}}\Delta\omega_\kappa/2\pi$, for $\epsilon_\kappa(\Delta\omega_\kappa) = \epsilon_{\max} = 2\%$ evaluated from $N_\kappa = 10^6$ realizations of a $N = 100$ section fiber emulator with 0 dB PDL as determined by the fourth order Magnus expansion Eq. (2.24) ($\circ$), Eqs. (2.34) ($\bullet$), the symmetric product of Eq. (2.38) ($\times$) and the non-symmetric expansion of Eq. (2.39) ($+$).

$\Delta\nu_{\min}$ substantially reduces the probability that the estimation error exceeds $\epsilon_{\max}$.

In Fig. 2.3 we display as a function of τ_{mean} the number of expansion orders required for each of the models of Section 2.2 to reproduce the frequency evolution of the Jones matrix to within $\epsilon_{\max} = 2\%$ for $P(\Delta\nu \geq \Delta\nu_{\min}) = 99.99\%$ over a $|\omega - \omega_0|/2\pi \leq 10$ GHz frequency interval. While the range of τ_{mean} for the desired level of accuracy is greatest for the Magnus expansion, Eq. (2.24), even for relatively low values of τ_{mean} many terms in the Taylor expansion of the PMD vector are required to ensure sufficient precision over the entire 20 GHz frequency interval although reducing the confidence level $P(\Delta\nu \geq \Delta\nu_{\min})$ or increasing the acceptable modeling error $\epsilon_{\max}$ decreases the expansion order required to attain a given value of τ_{mean}.

2.4.2 Magnus expansion with PMD and PDL

Next, in Fig. 2.4 we examine the relative accuracy of the Magnus expansion in the presence of PDL by displaying the variation of $\Delta\nu_{\min}$ with the mean PDL, α_{mean}, for a fiber emulator model after truncating Eq. (2.24) to first ($\circ$), second ($\times$), third ($+$), fourth ($\bullet$) and fifth ($\triangle$) order in $\Delta\omega$. Here we have employed a constant segment DGD of $|\vec{\tau}_{(m)}| = 2.17$ ps for $m = 1, \ldots, N$. The normalized bandwidth $\Delta\nu = \tau_{\text{mean}}\Delta\omega_\kappa/2\pi$ in Fig. 2.4 accounts for the slight dependence of τ_{mean} on α_{mean}, arising from the relationship $\tau_{\text{mean}} = \mathrm{E}\{\mathrm{Re}\{\sqrt{\vec{W}\cdot\vec{W}}\}\}$ for non-zero PDL, where $\mathrm{E}\{\cdots\}$ represents an ensemble average over emulator realizations [55]. Although the bandwidth reduction at large α_{mean} is most evident at high expansion orders, for $\alpha_{\text{mean}} \leq 4$ dB the worst case reduction is $\approx 7\times10^{-3}$ in $\Delta\nu_{\min}$ up to fifth order, implying that a nominal 7.6 GHz bandwidth would be reduced by 0.35 GHz for $\tau_{\text{mean}} = 20$ ps and by ≈ 0.042 or 2.1 GHz for an α_{mean} of 10 dB.

Finally, in Fig. 2.5 we display as in Fig. 2.3 the variation of the number of Magnus expansion orders required to reproduce the frequency evolution of the Jones matrix to within $\epsilon_{\max} = 2\%$ for $P(\Delta\nu \geq \Delta\nu_{\min}) = 99.99\%$ over a $|\omega - \omega_0|/2\pi \leq 10$ GHz frequency

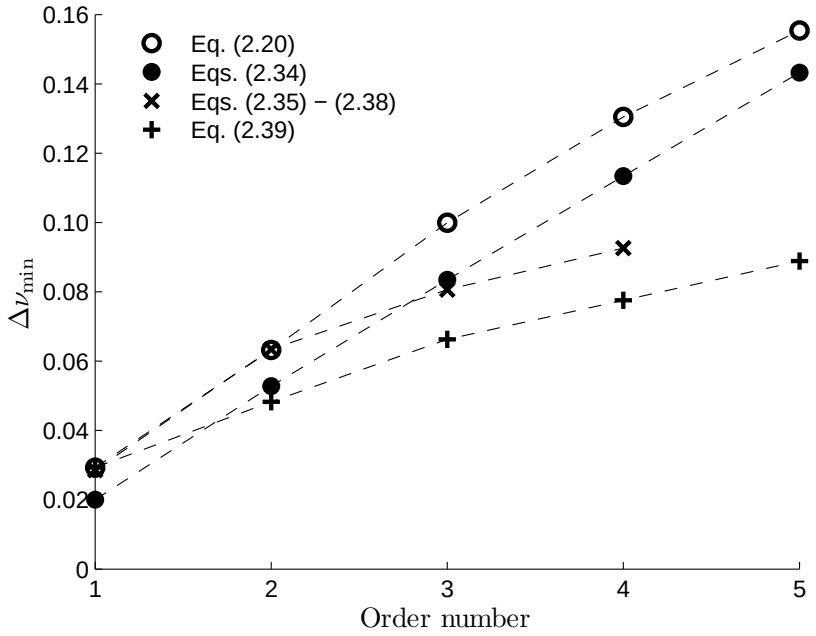


Figure 2.2: The smallest normalized bandwidth, $\Delta\nu_{\min}$, for which the maximum observed error of Eq. (2.52) is $< 2\%$ for 99.99% of the randomly generated emulator realizations versus the expansion order, M, for the Magnus expansion of Eq. (2.24) ($\circ$), the symmetric operator expansion of Eqs. (2.35)-(2.38) ($\times$), the non-symmetric expansion of Eq. (2.39) ($+$), and Eqs. (2.34) ($\bullet$).

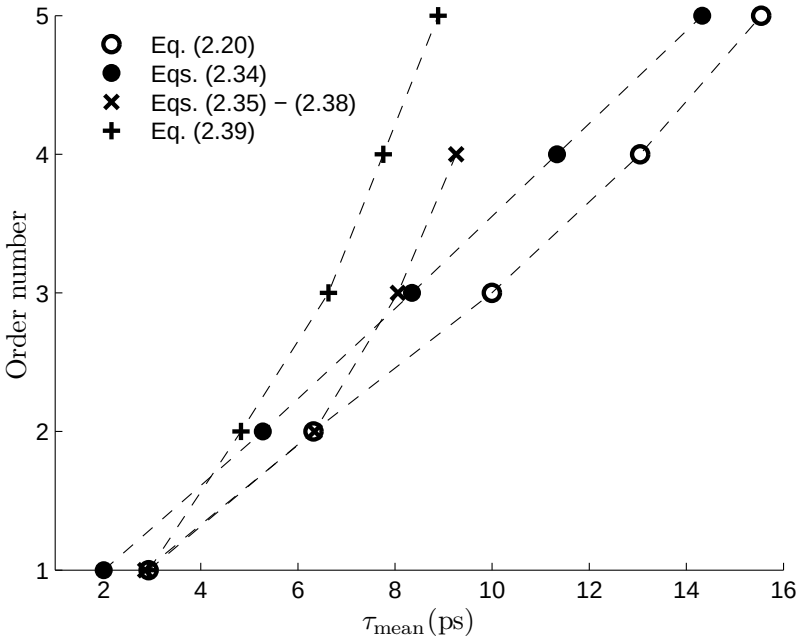


Figure 2.3: Same as Fig. 2.2, but for the number of expansion orders required to determine the frequency evolution of the Jones matrix with a maximum error of $\epsilon_{\max} = 2\%$ with 99.99% confidence over a $|\omega - \omega_0|/2\pi \leq 10$ GHz frequency interval for mean DGD, τ_{mean}, and 0 dB mean PDL.

interval with $\tau_{\rm mean}$, for values of mean PDL in the range $0 \leq \alpha_{\rm dB} \leq 10$ dB. Evidently, relative to $\alpha_{\rm mean} = 0$ dB, the mean DGD must be reduced by approximately 4.2 ps to limit sufficiently the modeling error for $\Delta\omega/(2\pi) = 10$ GHz for the worst case of 9.9 dB mean PDL and $M = 5$ expansion orders, while for $\alpha_{\rm mean} \leq 4$ dB $\tau_{\rm mean}$ should be reduced by ≈ 0.7 ps.

2.4.3 Taylor expansion estimation bandwidth

We next examine the width of the frequency region over which the truncated Taylor expansion

$$\vec{W}_{\rm est}(\omega) = \sum_{n=0}^{N} \vec{W}^{(n)} \frac{\Delta\omega^n}{n!} \tag{2.53}$$

accurately reproduces the frequency evolution of $\vec{W}(\omega)$. Here we first construct the 3×3 correlation matrix $\rho_{ab} = \mathrm{E}\{W_a\, W^*_{{\rm est},b}\}/(\sigma_a\, \sigma_{{\rm est},b})$, for $a, b = 1, 2, 3$, formed from the components $W_a(\omega)$ and $W_{{\rm est},b}(\omega)$ of the $\vec{W}(\omega)$ and $\vec{W}_{\rm est}(\omega)$ vectors where $\mathrm{E}\{\cdots\}$ represents an ensemble average over 10^5 emulator realizations. Since $\mathrm{E}\{W_a\} = \{W_{{\rm est},b}\} = 0$ and W_a is statistically uncorrelated with $W_{{\rm est},b}$ for $a \neq b$, $\sigma_a^2 = \mathrm{E}\{|W_a|^2\}$ and $\sigma^2_{{\rm est},b} = \mathrm{E}\{|W_{{\rm est},b}|^2\}$ while ρ_{ab} is proportional to the identity matrix. Accordingly, we consider only the frequency variation of $\rho_{11} = \mathrm{E}\{W_1\, W^*_{{\rm est},1}\}/(\sigma_1\, \sigma_{{\rm est},1})$ of the first component of each vector. In Fig. 2.6 we plot the correlation $|\rho_{11}|$ as a function of normalized frequency $\Delta\nu = \tau_{\rm mean}\Delta\omega/2\pi$ for expansion orders $M = 0, 2, 4, \ldots 20$ where $\vec{W}_n$ is calculated from Eq. (2.50) with 0 dB mean PDL. We then employ $|\rho_{11}| = 0.97$, which results from a $M = 0$ Taylor expansion calculation performed according to Section 2.4.1 at the half-bandwidth, $\Delta\nu_{\rm psp}/2 = 1/16$, of the principal state, to display as + markers in Fig. 2.7 the frequency interval $\Delta\nu$ for which $|\rho_{11}| \geq 0.97$ as a function of the expansion order. The inclusion of PDL only moderately reduces the bandwidth for which $|\rho_{11}| \geq 0.97$ as indicated by the $\bullet$ markers in Fig. 2.7 which instead use $\alpha_{\rm mean} = 9.9$ dB.

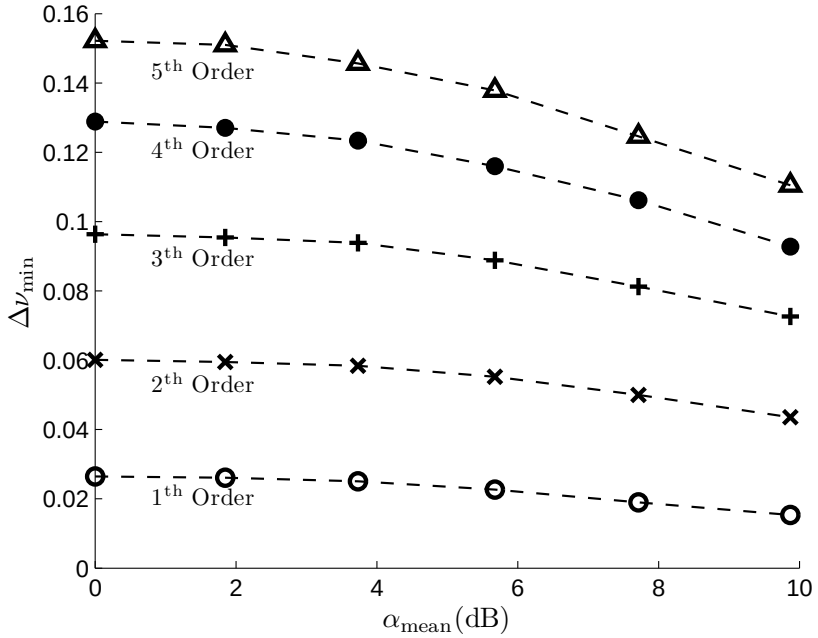


Figure 2.4: The variation of $\Delta\nu_{\min}$ with the mean PDL, α_{mean}, of a $N = 100$ section emulator for the Magnus expansion of Eq. (2.24) truncated to first ($\circ$), second ($\times$), third ($+$), fourth ($\bullet$) and fifth ($\triangle$) order in $\Delta\omega$.

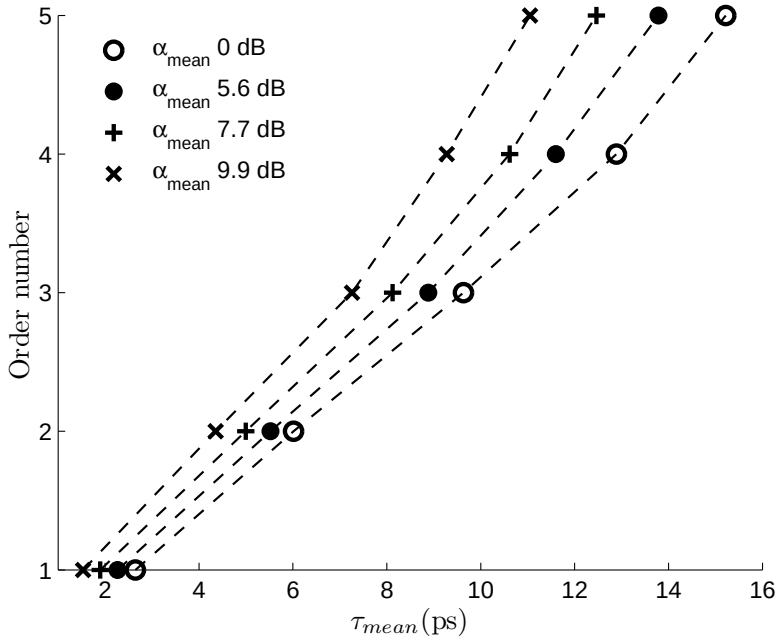


Figure 2.5: Same as Fig. 2.3, but for the number of Magnus expansion orders required to limit the error in the computed frequency evolution of the Jones matrix to $\epsilon_{\max} = 2\%$ with 99.99% confidence over a $|\omega - \omega_0|/2\pi \leq 10$ GHz frequency interval for a mean DGD τ_{mean} and values of the mean PDL in the range $0 \leq \alpha_{\text{mean}} \leq 10$ dB.

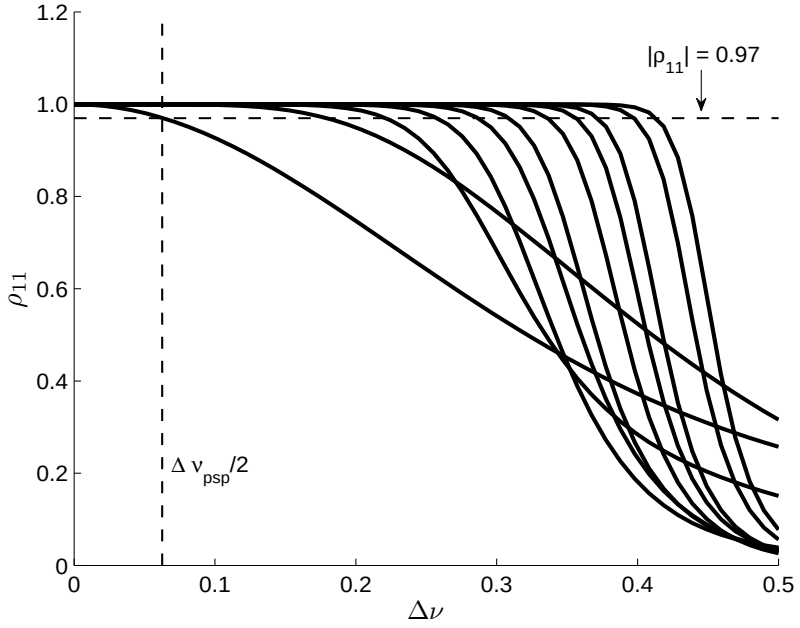


Figure 2.6: The variation with normalized frequency, $\Delta\nu = \tau_{\text{mean}}\Delta\omega/2\pi$, of the correlation $|\rho_{11}|$ between the x-components of $\vec{W}$ and $\vec{W}_{\text{est}}$ for Taylor orders $N = 0, 2, 4, \ldots 20$ and 0 dB mean PDL. A vertical dotted line indicates the half bandwidth of the principal state, $\Delta\nu_{\text{psp}}/2$, with $\Delta\nu_{\text{psp}} = 1/8$.

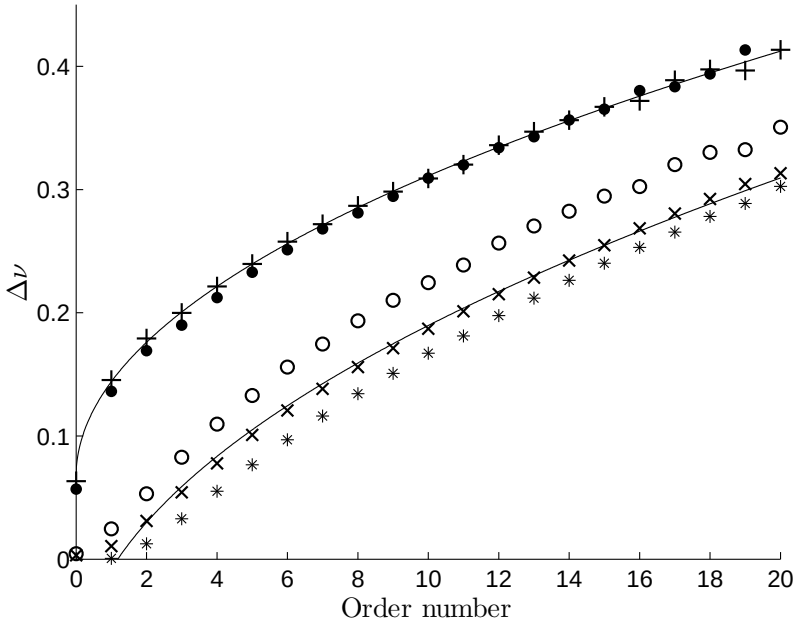


Figure 2.7: The variation of the frequency interval $\Delta\nu$ for which $|\rho_{11}| \geq 0.97$ (+ markers) and $|\rho_{11}| \geq 0.9999$ ($\circ$ markers) with the Taylor expansion order, N and 0 dB mean PDL. Analogous results for $|\rho_{11}| \geq 0.97$ and 9.9 dB mean PDL are indicated by the $\bullet$ markers. Also shown is $\Delta\nu_{\min}$, as calculated with the numerical methods of Section 2.4.1 for $\epsilon_{\max} = 2\%$ and $P(\Delta\nu \geq \Delta\nu_{\min}) = 99.99\%$, with 0 dB ($\times$ markers) and 9.9 dB ($*$ markers) mean PDL. The solid lines are optimal fits to square-root functions.

In a second calculation, shown as $\times$ markers in Fig. 2.7, we instead apply the numerical techniques of Section 2.4.1 to the maximum error

$$\epsilon_W(\Delta\omega) = \max_{|\omega-\omega_0|\leq\Delta\omega} \left[\frac{\|\vec{W}_{\text{est}}(\omega) - \vec{W}(\omega)\|}{\|\vec{W}(\omega)\|} \right] \cdot 100\% \tag{2.54}$$

to identify the frequency interval $\Delta\nu_{\min}$, with 0 dB mean PDL, for which $\epsilon_{\max} = 2\%$ and $P(\Delta\nu \geq \Delta\nu_{\min}) = 99.99\%$ (the solid lines of Fig. 2.7 are least-square fits of the results of both calculations to square-root functions [19]). Analogous results for $\alpha_{\text{mean}} = 9.9$ dB, marked $*$ in Fig. 2.7, are only weakly dependent on the mean emulator PDL. Evidently, limiting the maximum modeling error significantly reduces the bandwidth compared to enforcing a minimum statistical correlation between the estimated and exact complex principal state vectors. This is clearly evident from the $|\rho_{11}| = 0.9999$ results for $\alpha_{\text{mean}} = 0$ dB in Fig. 2.7 ($\circ$ markers), which more accurately reproduce the $\epsilon_{\max} = 2\%$ curve.

2.5 Conclusions

While the Magnus expansion model of the frequency dependence of the Mueller matrix in linear optical systems affected by both PMD and PDL reproduces previous power-series expansion and differential equation solution techniques, it additionally preserves the symmetry of the Mueller matrix in every expansion order. This yields physically realizable operator expansions that facilitate the design of joint PMD and PDL compensators [98, 124, 126]. Other practically relevant models of PMD frequency variation [30, 126] can be obtained by applying the BCH identity [44, 98] to the novel low order expansions of Eq. (2.19) while improved accuracy within a broader frequency range could potentially be achieved by extending the formalism according to Refs. [62, 103].

Although the frequency behaviour of the Jones or Mueller matrix is generally characterized through Taylor expansion coefficients, a fifth order Magnus expansion is, for example,

required to represent the Jones matrix with 99.99% accuracy for a mean DGD of 15 ps. Thus other parameterizations may be preferable for the numerical simulation of large mean DGD fibers. Numerical models such as the Jones matrix expansions and the interpolation procedures of Ref. [99] that describe a large span of frequencies can as well lead to enhanced numerical accuracy and programming simplicity for high bit-rate single or multiple channel optical system simulations. Of course, in this case the physical connection between specific Taylor orders of the complex principal state vector and the resulting PMD/PDL induced pulse distortion is lost.

Finally, although PMD model performance is quantified in this chapter by the maximum error incurred within a given frequency range, the system penalty due to PMD and PDL in an optical network may be a highly-nonlinear function of this error, dependent upon implementation details [121, 129], or may possibly be sensitive only to the coefficients of the low Magnus expansion orders.

Chapter 3

Transient Evolution of Polarization in Dispersion Compensation Modules

In this chapter, we employ a falling weight to generate high-frequency fluctuations in the polarization and PMD of commercial dispersion compensation modules and observe that for low impact forces the polarization variation is nearly identical for successive measurements.

3.1 Introduction

High frequency SOP variations can significantly affect the design and performance of both PMD compensation devices and advanced coherent lightwave systems based on polarization sensitive detection. In aerial fiber for example the SOP has been observed to rotate on the Poincaré sphere at frequencies up to several kilohertz [18, 113], while the results of several field trials [16,67] have demonstrated high speed polarization transient events occurring on timescales < 1 ms even in buried fiber links. As well, the mechanical vibrations generated by the impact of metal tools on dispersion compensation modules (DCMs) containing long

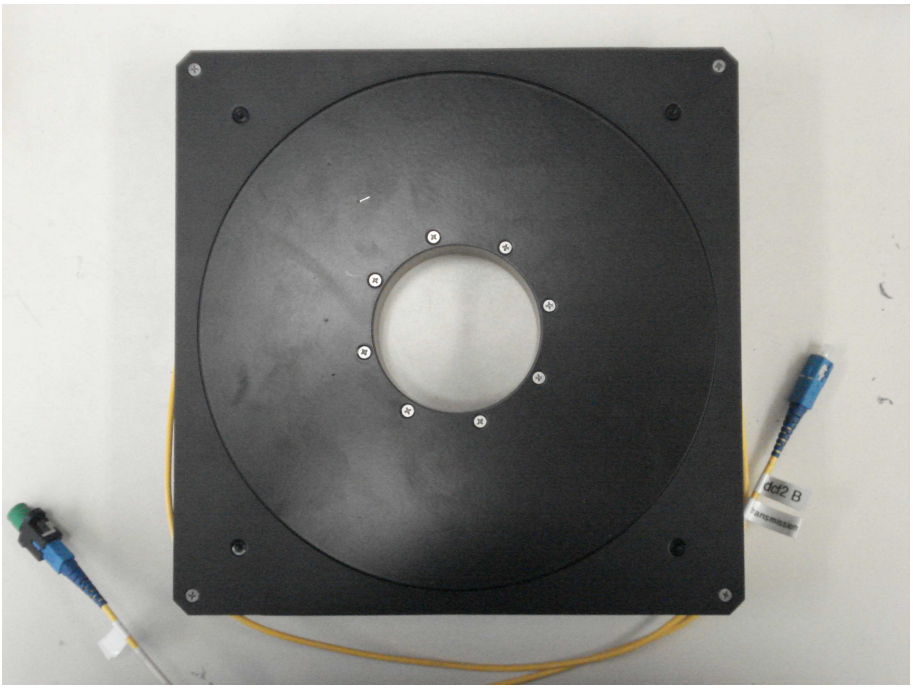


Figure 3.1: The Avanex 20 km dispersion compensation module. Impacts were targeted 1 cm radially outward from the inner cylindrical core of the DCM housing.

lengths of dispersion compensating fiber have been associated with random trajectories of the SOP on the Poincaré sphere varying at instantaneous rates equivalent to $> 45 \times 10^3$ complete rotations of the sphere per second (Rot/s) [66].

While the remainder of this thesis examines the temporal evolution of polarization characterized by a stochastic evolution of the PMD vector over the Poincaré sphere, in this chapter we will briefly consider the polarization transients generated during mechanical impacts of a commercially available DCM (a typical dispersion compensation module is illustrated in Fig. 3.1). Here, we will demonstrate transient events with instantaneous rotation rates exceeding > 75 kRot/s by releasing a 2.6 g steel ball from heights of several centimeters onto a DCM and, more significantly, will show that after a suitable restoring time the observed SOP trajectory is approximately retraversed in a subsequent impact.

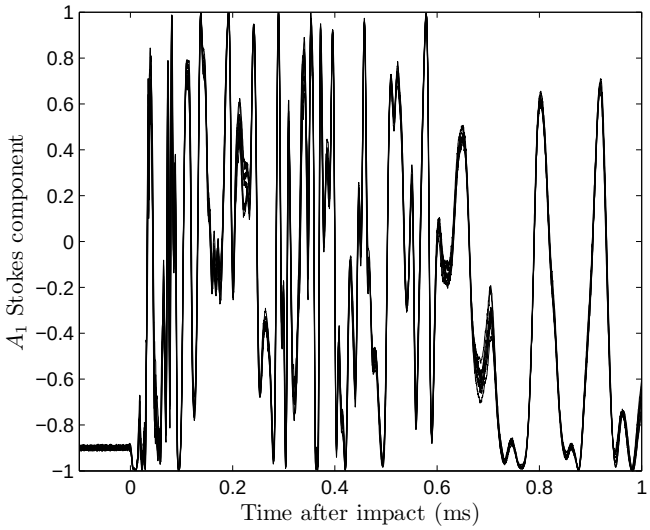
These "elastic" polarization transient events were studied in greater experimental detail in Ref. [28] and modeled numerically in Ref. [111]. More recently, elastic transients were reported in large-scale field trials involving 1800 km terrestrial fiber links [82]. As a result, an optical network can experience transient degradations of the bit error rate in which the instantaneous rate of polarization rotation or the PMD and PDL briefly exceed acceptable performance thresholds.

3.2 Experimental setup

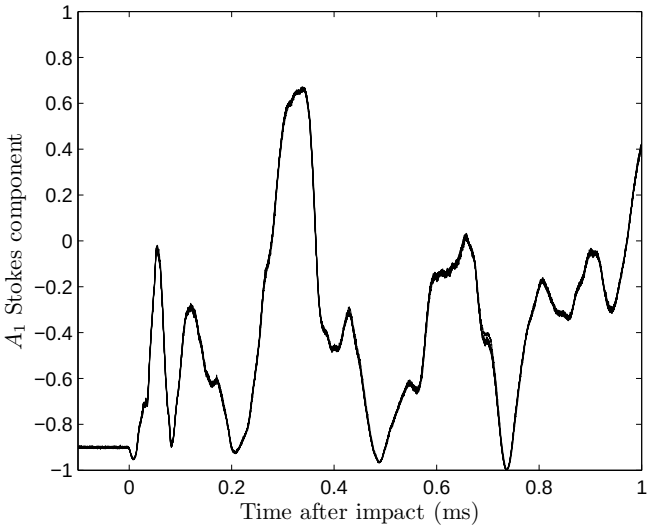
We characterize high-speed polarization transients by inserting polarized light from a Photonetics Tunics-PRI tunable laser into a Nortel DCM60 or an Avanex 20 km dispersion compensation module. The SOP at the input of the DCM was controlled with an HP 11896A while the outgoing polarization state was sampled at a 2.5×10^6 Hz rate with a Tektronix digitizing oscilloscope attached to the detectors of a custom built, high speed polarimeter based upon an integrated Lightwaves2020 device. The polarimeter's analog electrical bandwidth was approximately 40 MHz.

The polarimeter was itself calibrated against an HP 8509B polarization analyzer. Successive measurements of the SOP Stokes vector on the Poincaré sphere, $\hat{A} = [A_1, A_2, A_3]^{\mathrm{T}}$, coincided to within a mean angular error of 0.07 deg with the HP 8509B. The magnitude of the error further remained < 0.25 deg over 200 independent measurements.

In Fig. 3.2, we display the results of 10 measurements of the time variation of the A_1 Stokes component of the SOP at the output of the DCM. Similar results for the temporal variation of $\hat{A}$ on the Poincaré sphere are displayed in Fig. 3.2 for the Avanex DCM. Here, we dropped a 2.6 g and 1.3 cm diameter steel ball 9 cm above a location 1 cm from the inner cylindrical core of a Nortel DCM. Although the polarization response of the Nortel and Avanex DCMs in Figs. 3.2a and 3.2b are qualitatively similar, we have observed that the



(a) Nortel DCM60



(b) Avanex 20 km DCM

Figure 3.2: The time variation of the Stokes component, A_1, at the output of (a) a Nortel DCM60 and (b) an Avanex 20 km DCM during the first 1 ms after ten collisions separated by $\approx$ 20 sec intervals with a 2.6 g steel ball released from 9 cm above a point 1 cm from the DCM inner core.

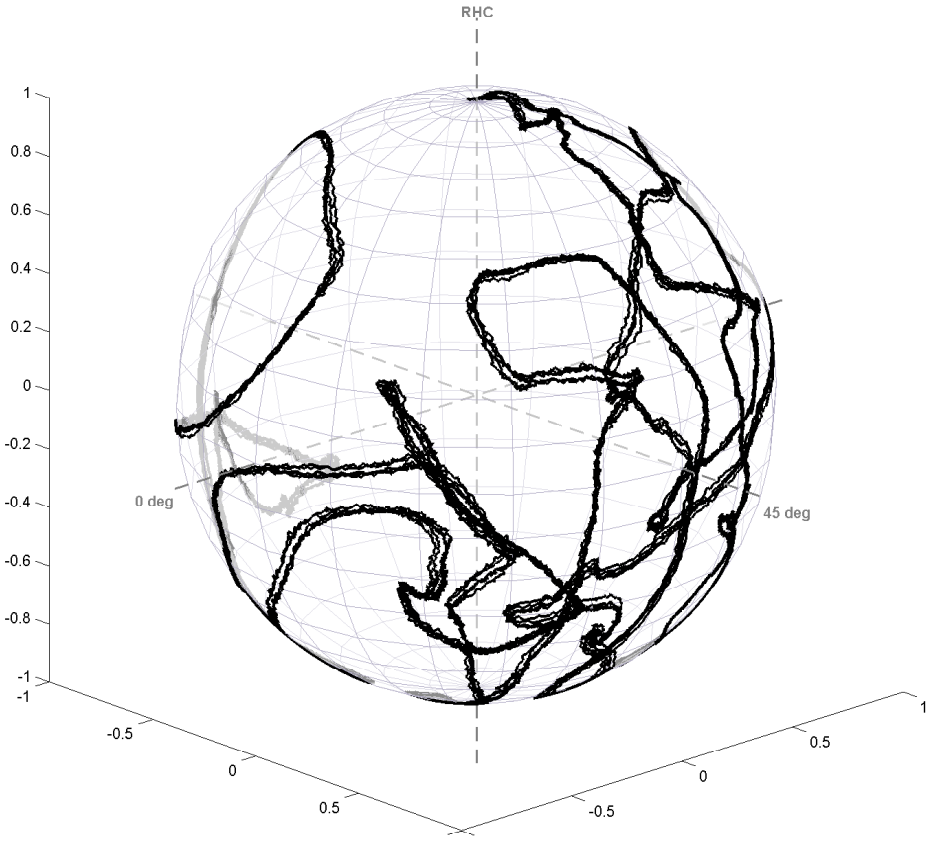


Figure 3.3: The time variation of $\hat{A}$ on the Poincaré sphere at the output of the Avanex 20 km DCM during the first 0.75 ms after five collisions separated by $\approx$ 20 sec intervals with a 2.6 g steel ball released from 9 cm above a point 1 cm from the DCM inner core.

minimum required relaxation time can vary between 1 sec and > 40 sec between different DCM models. In particular, a ~ 20 sec time between measurements is typically required to restore the initial state of the Nortel DCM60.

Fig. 3.2 indicates that for times < 0.2 ms after the collision the output SOP rotates at approximate rates > 75 kRot/s and > 10 kRot/s for the Nortel and Avanex DCM, respectively, comparable to the previously reported value of 45 kRot/s [66]. Low frequency oscillations then persist for an additional ~ 20 ms. However, Fig. 3.2 further demonstrates that for an identical impact location and release height, the individual components of the output Stokes vectors coincide to within 4% when averaged over the time series, although sufficiently large impact forces alter the internal state of the DCM and therefore reduce the measurement reproducibility. The increased polarization activity of the Nortel DCM60 in Fig. 3.2 was related in Ref. [111] to the increased fiber length and reduced modal size relative to the Avanex unit.

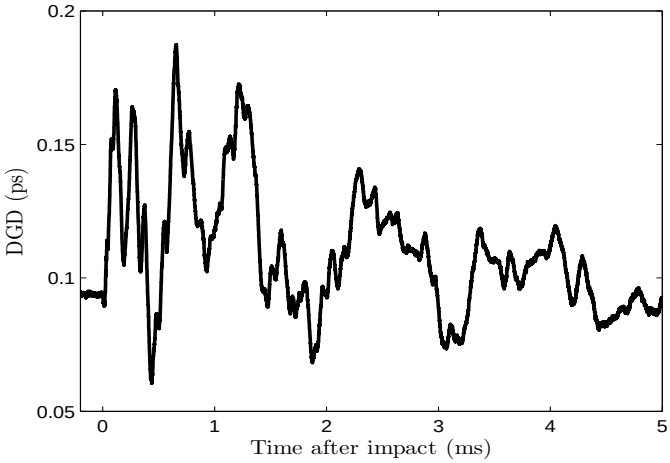
Since the temporal variation of the SOP is nearly invariant for successive collisions, the time-variation of the PMD of the DCM can be determined by repeating the impact experiment with different input polarization states and optical frequencies. To first order in the PMD, the output SOP of the DCM $\hat{A}_{out}(\omega_1; t)$ and $\hat{A}_{out}(\omega_0; t)$ at time t and optical frequencies ω_0 and $\omega_1 = \omega_0 + \Delta\omega$ are related through

$$\hat{A}_{out}(\omega_0 + \Delta\omega; t) = \mathbf{U}(\Delta\omega; t)\, \hat{A}_{out}(\omega_0; t), \tag{3.1}$$

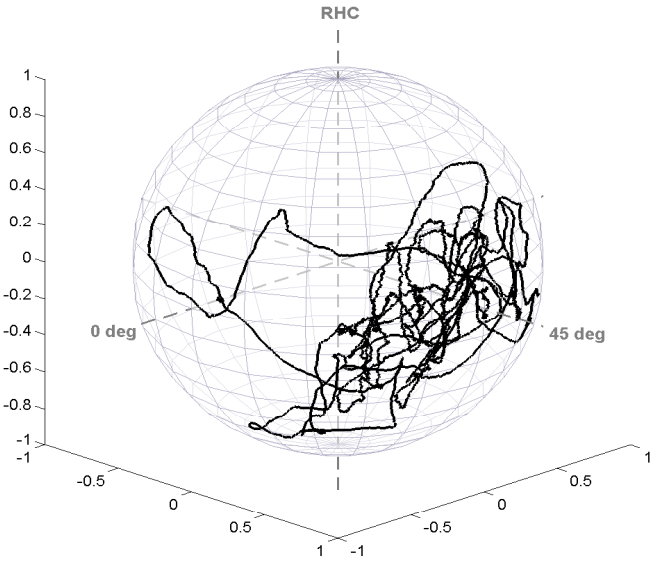
where the 3×3 rotation matrix $\mathbf{U}(\Delta\omega; t)$ is specified in terms of a time-dependent PMD vector $\vec{\tau}(t)$ according to

$$\mathbf{U}(\Delta\omega; t) = \exp\left[\Delta\omega\, \vec{\tau}(t)\times\right]. \tag{3.2}$$

We have employed the least-square method of Refs. [97, 104] to extract $\mathbf{U}(\Delta\omega; t)$ from measurements of $\hat{A}_{out}(\omega_0 + \Delta\omega; t)$ and $\hat{A}_{out}(\omega_0; t)$ for 6 input polarization states maximally

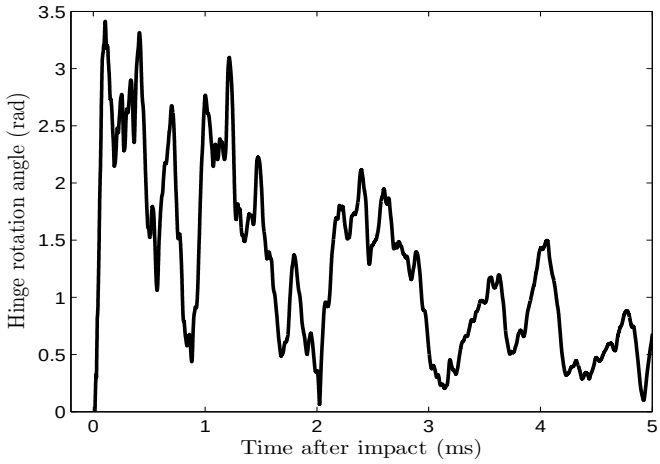


(a) Differential group delay

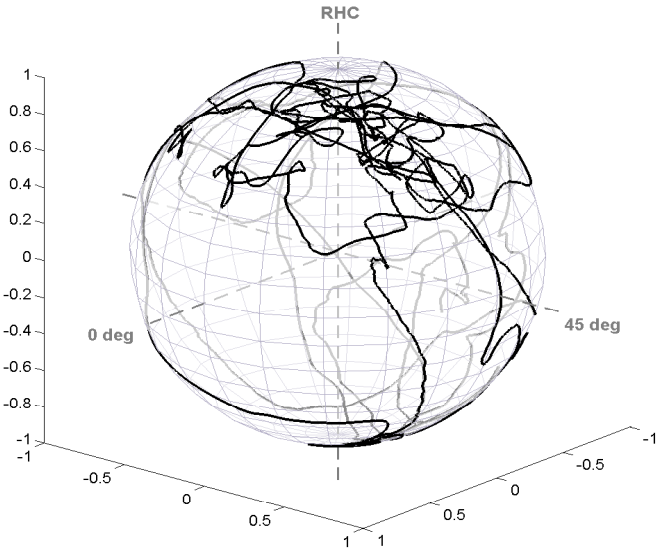


(b) Principal state of polarization

Figure 3.4: The variation with time of (a) the DGD and (b) the PSP of an Avanex 20 km DCM during the first 5 ms after collision with a 2.6 g steel ball, again for a 11 cm release height and a 1.0 cm radial displacement from the inner core.



(a) Rotation angle



(b) Rotation axis

Figure 3.5: The time variation of the DCMs hinge (a) rotation angle and (b) rotation axis after the mechanical excitation of Fig. 3.4.

separated on the Poincaré sphere and a frequency offset $\Delta\omega$ corresponding to a 5 nm wavelength interval. To reduce experimental error, we further averaged each $\hat{A}_{out}(\omega;t)$ over 5 successive impacts. Fig. 3.4 displays the DGD, $\tau(t) = |\vec{\tau}(t)|$, and principal state, $\hat{p}(t) = \vec{\tau}(t)/|\vec{\tau}(t)|$, as a function of time over a 5 ms interval after a collision with the ball released from 11 cm above a point 1 cm from the cylindrical core of the Avanex DCM. Here the DGD fluctuates between ≈ 0.06 ps and 0.19 ps at frequencies exceeding 10 kHz, while the PSP covers approximately one quarter of the Poincaré sphere.

As the measured DGD of the DCM is low (< 0.2 ps for the Avanex model), the mechanical excitations can be modeled as purely rotational hinge sites in the context of the hinge-model of PMD in optical networks [65]. That is, the relation between the output SOPs at an initial t_0 and a later time $t_1 = t_0 + \Delta t$ at a constant optical frequency ω is approximately described by

$$\hat{A}_{out}(\omega; t_0 + \Delta t) = \mathbf{R}(\Delta t)\, \hat{A}_{out}(\omega; t_0), \tag{3.3}$$

with

$$\mathbf{R}(\Delta t) = \exp\left[\psi(\Delta t)\ \hat{r}(\Delta t)\times\right] \tag{3.4}$$

and a time-dependent rotation angle $\psi(\Delta t)$ and rotation axis $\hat{r}(\Delta t)$. As in the preceding paragraph, we determined $\mathbf{R}(\Delta t)$ from measurements of $\hat{A}_{out}(\omega;t)$, averaged over 5 successive collision events and 6 random input polarizations. With the identical mechanical excitation as in Fig. 3.4, we then obtain Fig. 3.5 for the variation of $\psi(\Delta t)$ and $\hat{r}(\Delta t)$ during the first 5 ms after collision. Evidently, even relatively small mechanical excitations generate large variations in both the rotation angle $\psi(\Delta t)$ and the rotation axis $\hat{r}(\Delta t)$, with the rotation axis traversing a large region of the Poincaré sphere.

3.3 Conclusions

We have demonstrated that low-amplitude mechanical excitations of standard DCMs can yield high frequency, > 75 kRot/s, polarization transients that are nearly invariant between successive measurements. This reproducibility enables the exhaustive characterization of the PMD of such components as well as their precise hinge model parametrization. Further, by placing the DCM in series with one or more sections of PMD fiber, inexpensive, high-frequency and highly reproducible PMD scramblers can be constructed for system testing purposes. These hinge model results could further be incorporated into system simulators, which would, for example, enable the performance of polarization sensitive detectors to be analyzed with increased accuracy.

Chapter 4

A Time-Dependent Polarization Mode Dispersion Emulator

We implement a fiber emulator based on a random walk over the unit sphere that accurately simulates the temporal evolution of polarization-mode dispersion (PMD) in fiber optic communication systems with relatively few emulator sections. We then derive a simple expression relating the characteristic decorrelation time of the emulated PMD vector to the properties of the underlying random walk. Finally, the model is adapted in a multicanonical calculation of low-probability PMD induced system outage events.

4.1 Introduction

Fiber emulators predict the frequency variation of the state of polarization (SOP) and the PMD vector at the output of standard, single mode optical fiber [10, 24, 59, 64, 70, 79]. Consequently, they have been successfully employed to simulate the effects of polarization mode dispersion (PMD) on system performance at transmission rates exceeding 10 Gb/s.

In an emulator, N randomly oriented, linearly birefringent sections are concatenated in such a manner that the emulator yields the correct fiber statistics for distances sufficiently large compared to the characteristic birefringence correlation length [114].

The temporal variations of the fiber birefringence have previously been incorporated into the emulator model by, for example, either varying the differential group delay (DGD) with time while preserving the birefringence orientation [33, 61], or considering temporal variations as a small perturbation on an otherwise constant background birefringence [3, 80]. Both of these procedures approximate buried or aerial optical fiber links in which the thermal and mechanical fluctuations responsible for the PMD temporal dynamics are distributed along the fiber length in such a manner that the polarization evolution can be represented as a diffusion process. As $N \to \infty$, however, each method predicts equivalent temporal autocorrelation statistics [80].

In this chapter, we extend this prior work by changing the birefringence orientation of each emulator section by a constant but randomly oriented angle on the Poincaré sphere for each simulated time step. This procedure, while simple to implement, correctly models the temporal statistics expected in the diffusion limit for a relatively small value of N, as verified here by ensuring that the resulting probability of a PMD induced system outage as a function of time coincides with the predictions of a previous stochastic differential equation analysis [3]. Additionally, by adapting the emulator into standard multicanonical procedures [7, 53, 122, 123, 125], we accurately model the temporal dynamics of low probability regions with orders of magnitude less computation time than required by standard Monte-Carlo techniques [127, 128].

In standard emulator models [60, 70, 79, 80, 108] if $\vec{\tau}_{m,k}$ denotes the PMD vector of the m:th element of a series of birefringent segments at the k:th simulated temporal step, the total PMD vector $\vec{\Omega}_N(\omega, k)$ at optical angular frequency ω after the final, N:th element is

given by [24, 47, 59]

$$\vec{\Omega}_N(\omega, k) = \vec{\tau}_{N,k} + \mathbf{R}_N(\omega, k)\,\vec{\Omega}_{N-1}(\omega, k) \tag{4.1}$$

with $\vec{\Omega}_0(\omega, k) = 0$. Each rotation matrix

$$\mathbf{R}_m(\omega, k) = \exp\left(\omega\,\vec{\tau}_{m,k}\,\times\right) \tag{4.2}$$

in Eq. (4.1) precesses the direction of the incident polarization vector around the linear birefringence of the m:th optical element according to Rodrigues' formula for rotating a vector by an angle ψ about a second unit vector $\hat{r}$ [47]

$$\exp\left(\psi\,\hat{r}\,\times\right) = \cos\psi\,\mathbf{I}_3 + (1 - \cos\psi)\,\hat{r}\hat{r}^{\mathrm{T}} + \sin\psi\,\hat{r}\,\times\,. \tag{4.3}$$

Here $\mathbf{I}_3$ is the 3×3 identity matrix and

$$\hat{r}\,\times = \begin{bmatrix} 0 & -r_3 & r_2 \\ r_3 & 0 & -r_1 \\ -r_2 & r_1 & 0 \end{bmatrix} \tag{4.4}$$

where $\hat{r} = [r_1, r_2, r_3]^{\mathrm{T}}$ and T indicates the transpose.

In this chapter, we describe a technique that varies $\vec{\tau}_{m,k}$ at the k:th simulation step by a constant angle α [101] on the Poincaré sphere such that if $\vec{\tau}_{m,k} = \tau_m\,\hat{p}_{m,k}$, where τ_m is time-independent, then $\hat{p}_{m,(k+1)} \cdot \hat{p}_{m,k} = \cos\alpha$ for $k \geq 0$, although other implementations have employed e.g. randomly distributed values of α [101]. The resulting random walk approaches a diffusion process on the unit sphere for $\alpha \ll 1$. The components of $\vec{\tau}_{m,0}$ at the initial $k = 0$ time index are typically independent, zero mean Gaussian distributed random variables with variance σ_τ^2 [70], although in later sections $\vec{\tau}_{m,0}$ will be selected using the Brownian bridge algorithm [107]. The initial unit vector $\hat{p}_{m,0} = \vec{\tau}_{m,0}/|\vec{\tau}_{m,0}|$ in the former case is uniformly distributed over the sphere, while the DGD of each emulator section, $\tau_m = |\vec{\tau}_{m,k}|$, possesses a Maxwellian probability density function (pdf)

$$f_\tau(x) = \sqrt{\frac{2}{\pi}}\frac{x^2}{\sigma_\tau^3}\exp\left(-\frac{x^2}{2\sigma_\tau^2}\right) \tag{4.5}$$

with first and second moments $\mathrm{E}\{\tau_m\} = \sigma_\tau\sqrt{8/\pi}$ and $\mathrm{E}\{\tau_m^2\} = 3\,\sigma_\tau^2$, respectively, yielding an rms DGD, $\tau_{\mathrm{rms}}^2 = \mathrm{E}\left\{|\vec{\Omega}_N(\omega,k)|^2\right\}$, of $\tau_{\mathrm{rms}} = \sigma_\tau\sqrt{3N}$ for the emulated PMD vector. Unlike models involving a constant DGD for each section, this approach typically reproduces higher order PMD statistics [64] with fewer emulator sections while suppressing periodicities in the frequency autocorrelation function for small N [79], increasing the free spectral range of the emulator.

To model the temporal properties of $\vec{\Omega}_N(\omega,k)$ for an arbitrary number of emulator sections N, we consider the PMD autocorrelation function [60, 79]

$$C_N(\omega,k;\Delta\omega,n) = \mathrm{E}\left\{\vec{\Omega}_N(\omega,k)\cdot\vec{\Omega}_N(\omega+\Delta\omega,k+n)\right\} \tag{4.6}$$

in which $\mathrm{E}\{\cdots\}$ denotes an expectation over realizations of the stochastic process. If the $\vec{\tau}_{m,k}$ average to zero and are further independent and identically distributed between emulator sections [59, 70],

$$\begin{aligned} C_N(\omega,k;\Delta\omega,n) = {} & C_\tau(k;n) \\ & + \mathrm{E}\Big\{\vec{\Omega}_{N-1}(\omega,k)^{\mathrm{T}} \\ & \times \mathbf{R}_N(\omega,k)^{\mathrm{T}}\mathbf{R}_N(\omega+\Delta\omega,k+n) \\ & \times \vec{\Omega}_{N-1}(\omega+\Delta\omega,k+n)\Big\}. \end{aligned} \tag{4.7}$$

Here, we have defined $C_\tau(k;n) = \mathrm{E}\{\vec{\tau}_{m,k}\cdot\vec{\tau}_{m,k+n}\}$ as the temporal autocorrelation function of the m:th emulator section. Further, if the distribution of the orientation of $\vec{\tau}_{m,k}$ is spherically symmetric, by symmetry the autocorrelation of $\mathbf{R}_m(\omega,k)$ is proportional to the identity matrix so that $\mathrm{E}\left\{\mathbf{R}_m(\omega,k)^{\mathrm{T}}\mathbf{R}_m(\omega+\Delta\omega,k+n)\right\} = g(\omega,k;\Delta\omega,n)\,\mathbf{I}_3$ and

$$\begin{aligned} C_N(\omega,k;\Delta\omega,n) = {} & C_\tau(k;n) \\ & + g(\omega,k;\Delta\omega,n)\,C_{N-1}(\omega,k;\Delta\omega,n). \end{aligned} \tag{4.8}$$

Repeatedly applying Eq. (4.8) yields the autocorrelation of the output PMD vector [61, 79],

$$C_N(\omega,k;\Delta\omega,n) = C_\tau(k;n)\left[\frac{1-g^N(\omega,k;\Delta\omega,n)}{1-g(\omega,k;\Delta\omega,n)}\right], \tag{4.9}$$

where the autocorrelations associated with each emulator section,

$$C_\tau(k;n) = \mathrm{E}\left\{\vec{\tau}_{m,k} \cdot \vec{\tau}_{m,k+n}\right\} \tag{4.10}$$

and

$$g(\omega,k;\Delta\omega,n) = \frac{1}{3}\,\mathrm{E}\left\{\mathrm{Trace}\left\{\mathbf{R}_m(\omega,k)^{\mathrm{T}}\mathbf{R}_m(\omega+\Delta\omega,k+n)\right\}\right\}, \tag{4.11}$$

can be evaluated for any time dependent local birefringence model. We demonstrate below that the autocorrelation functions $C_\tau(k;n)$ and $g(\omega,k;\Delta\omega,n)$ of the random walk method accurately represent the temporal dynamics of optical fiber links even for a small, $N \approx 20$, number of emulator sections.

4.2 Uniform random walk on a sphere

Omitting the segment index m, to ensure that for each time index k the unit vector $\hat{p}_k$ is rotated by a constant angle α in a random direction we calculate

$$\hat{p}_{k+1} = \cos\alpha\,\hat{p}_k + \sin\alpha\,\hat{q}_k. \tag{4.12}$$

The random unit vector $\hat{q}_k$ is selected from a uniform distribution over the great circle in the plane orthogonal to $\hat{p}_k$, ensuring that $\hat{p}_{k+1}\cdot\hat{p}_k = \cos\alpha$. To generate $\hat{q}_k$ we observe that, by spherical symmetry, projecting $\vec{r}_k$ onto the plane orthogonal to $\hat{p}_k$ yields a vector $\vec{q}_k = \vec{r}_k - (\vec{r}_k\cdot\hat{p}_k)\,\hat{p}_k$, and an associated unit vector $\hat{q}_k = \vec{q}_k/|\vec{q}_k|$ that are uniformly distributed in angle within this plane if the components of $\vec{r}_k$ are obtained from a zero mean unit variance Gaussian random number generator. The random walk can however be preferentially directed along any desired spatial directions by appropriately biasing the selection of $\hat{q}_k$.

Denoting the probability density of the initial unit vector $\hat{p}_0$ over the unit sphere by $f_0(\theta,\phi)$, with polar and azimuthal coordinates θ and ϕ, so that

$$\int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi} f_0(\theta,\phi)\,\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi = 1, \tag{4.13}$$

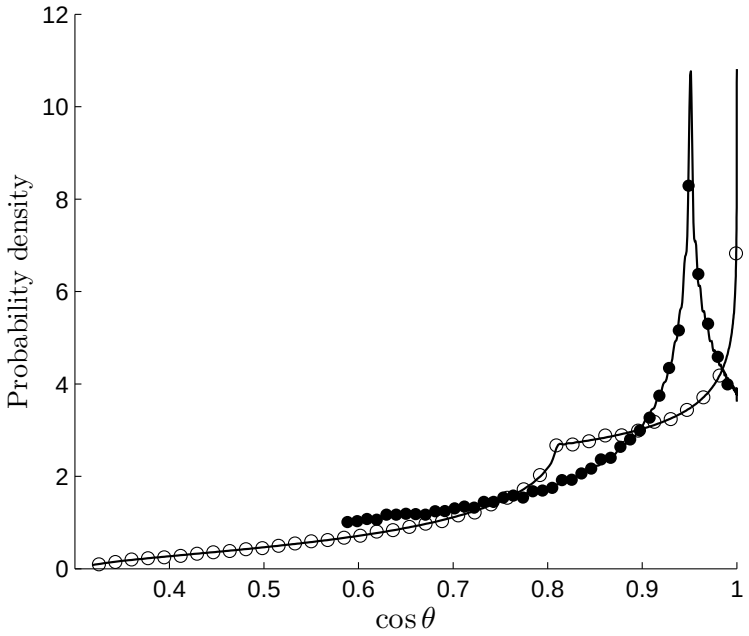


Figure 4.1: The pdf, $f_z(\cos\theta)$, of the z-component of $\hat{p}_k$ after k random steps for a vertically directed initial vector $\hat{p}_0$ as estimated from a 10^6 sample Monte-Carlo simulation after $k = 3$ ($\bullet$ markers) and 4 steps ($\circ$ markers) of a random walk with an angular step size of $\alpha = \pi/10$, respectively. The solid lines display the analytic result of Eq. (4.17).

and expanding in spherical harmonics $Y_l^m(\theta,\phi)$ according to

$$f_0(\theta,\phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} F_{lm} Y_l^m(\theta,\phi), \tag{4.14}$$

after k steps of the unbiased random walk the probability density of $\hat{p}_k$ is given by [101]

$$f_k(\theta,\phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \left[P_l\left(\cos\alpha\right)\right]^k F_{lm} Y_l^m(\theta,\phi), \tag{4.15}$$

where $P_l\left(\cos\alpha\right)$ represents the ordinary Legendre polynomial of degree l and the $Y_l^m(\theta,\phi)$ are orthonormalized with respect to integration over all solid angles. In the two cases of interest here, the initial vector $\hat{p}_0$ is (1) either oriented along the z-axis with arbitrary azimuthal angle ϕ so that

$$\begin{aligned} f_0(\theta,\phi) &= \frac{1}{2\pi\sin\theta}\,\delta(\theta) \\ &= \sum_{l=0}^{\infty} \sqrt{\frac{2l+1}{4\pi}}\, Y_l^0\left(\theta,\phi\right), \end{aligned} \tag{4.16}$$

or (2) uniformly distributed over the sphere with $f_0(\theta,\phi) = Y_0^0(\theta,\phi)/\sqrt{4\pi}$. Applying Eq. (4.15) for k random steps yields

$$f_k(\theta,\phi) = \sum_{l=0}^{\infty} \left[P_l\left(\cos\alpha\right)\right]^k \sqrt{\frac{2l+1}{4\pi}}\, Y_l^0\left(\theta,\phi\right) \tag{4.17}$$

and $f_k(\theta,\phi) = Y_0^0(\theta,\phi)/\sqrt{4\pi}$ for case (1) and (2) respectively.

In Fig. 4.1 the solid ($\bullet$) and open ($\circ$) circles display the pdf, $f_z(\cos\theta)$, obtained from a 10^6 sample Monte-Carlo simulation of the z-component, $\cos\theta$, of $\hat{p}_k$ after $k = 3$ and 4 random steps for case (1) and an angular step size of $\alpha = \pi/10$ while the solid lines in the figure correspond to the analytic result of Eq. (4.17) with $f_z(\cos\theta) = 2\pi f_k(\theta,0)$. Note the agreement between the simulated and analytic results which yield a minimum value of $\cos\theta$ after k steps of $\cos(k\alpha)$ for $k\alpha < \pi$ and -1 otherwise.

Expectation values of the polarization vector or its components can be obtained either from the recursion of Eq. (4.12), or by expanding each component of $\hat{p}_k$ in orders of the spherical harmonics and averaging with respect to $f_k(\theta, \phi)$ according to

$$\mathrm{E}\{G\} = \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} G(\theta, \phi)\, f_k(\theta, \phi)\, \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\phi, \tag{4.18}$$

where $G(\theta, \phi)$ represents an arbitrary function defined over the unit sphere. Since integrals involving $f_k(\theta, \phi)$ often reduce to simple algebraic expressions as a result of the orthogonality of the spherical harmonics we follow the latter approach, which further does not require correlation functions of the random vector $\hat{q}_k$.

Eq. (4.18) for $\mathrm{E}\{\hat{p}_k\}$ is evaluated in case (1) by expanding the components $p_{x,k} = \cos\phi \sin\theta$, $p_{y,k} = \sin\phi \sin\theta$ and $p_{z,k} = \cos\theta$ of $\hat{p}_k$ in orders of $Y_l^m(\theta, \phi)$ according to

$$\begin{aligned} \cos\phi \sin\theta &= \sqrt{\frac{2\pi}{3}} \left[Y_1^{-1*}(\theta, \phi) - Y_1^{1*}(\theta, \phi) \right] \\ \sin\phi \sin\theta &= -i\sqrt{\frac{2\pi}{3}} \left[Y_1^{-1*}(\theta, \phi) + Y_1^{1*}(\theta, \phi) \right] \\ \cos\theta &= \sqrt{\frac{4\pi}{3}}\, Y_1^0(\theta, \phi). \end{aligned} \tag{4.19}$$

This yields

$$\begin{aligned} \mathrm{E}\{\cos\phi \sin\theta\} = \mathrm{E}\{\sin\phi \sin\theta\} &= 0 \\ \mathrm{E}\{\cos\theta\} &= [P_1(\cos\alpha)]^k, \end{aligned} \tag{4.20}$$

and consequently

$$\mathrm{E}\{\hat{p}_k\} = [P_1(\cos\alpha)]^k \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \tag{4.21}$$

Therefore, for small angular step sizes $\alpha \ll 1$, $[P_1(\cos\alpha)]^k \approx \exp(-\alpha^2 k/2)$ and the asymmetry induced by the initial condition $\hat{p}_0 = [0, 0, 1]^{\mathrm{T}}$ becomes negligible after approximately $2/\alpha^2$ steps.

Considering next the expectation of the outer product $\mathrm{E}\left\{\hat{p}_k\hat{p}_k^{\mathrm{T}}\right\}$, the average of the off-diagonal matrix elements vanishes by symmetry in either case (1) or (2) while the diagonal elements satisfy $\mathrm{E}\left\{p_{x,k}^2\right\} = \mathrm{E}\left\{p_{y,k}^2\right\}$ in case (1). Further, $p_{x,k}^2$, $p_{y,k}^2$ and $p_{z,k}^2$ are given in terms of the spherical harmonics by

$$\begin{aligned}
\cos^2\phi\sin^2\theta &= \frac{1}{3}\sqrt{4\pi}\,Y_0^0(\theta,\phi) - \frac{1}{3}\sqrt{\frac{4\pi}{5}}\,Y_2^0(\theta,\phi) \\
&\quad + \sqrt{\frac{2\pi}{15}}\,\left[Y_2^{-2*}(\theta,\phi) + Y_2^{2*}(\theta,\phi)\right] \\
\sin^2\phi\sin^2\theta &= \frac{1}{3}\sqrt{4\pi}\,Y_0^0(\theta,\phi) - \frac{1}{3}\sqrt{\frac{4\pi}{5}}\,Y_2^0(\theta,\phi) \\
&\quad - \sqrt{\frac{2\pi}{15}}\,\left[Y_2^{-2*}(\theta,\phi) + Y_2^{2*}(\theta,\phi)\right] \\
\cos^2\theta &= \frac{1}{3}\sqrt{4\pi}\,Y_0^0(\theta,\phi) + \frac{2}{3}\sqrt{\frac{4\pi}{5}}\,Y_2^0(\theta,\phi).
\end{aligned} \tag{4.22}$$

Accordingly, from Eq. (4.18)

$$\begin{aligned}
\mathrm{E}\left\{\cos^2\phi\sin^2\theta\right\} &= \mathrm{E}\left\{\sin^2\phi\sin^2\theta\right\} \\
&= \frac{1}{3} - \frac{1}{3}\left[P_2(\cos\alpha)\right]^k \\
\mathrm{E}\left\{\cos^2\theta\right\} &= \frac{1}{3} + \frac{2}{3}\left[P_2(\cos\alpha)\right]^k
\end{aligned} \tag{4.23}$$

and

$$\mathrm{E}\left\{\hat{p}_k\hat{p}_k^{\mathrm{T}}\right\} = \frac{1}{3}\mathbf{I}_3 + \frac{1}{3}\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix}\left[P_2(\cos\alpha)\right]^k. \tag{4.24}$$

If $\alpha \ll 1$, $\left[P_2(\cos\alpha)\right]^k \approx \exp\left(-3\alpha^2 k/2\right)$, and $\mathrm{E}\left\{\hat{p}_k\hat{p}_k^{\mathrm{T}}\right\}$ tends towards $\mathbf{I}_3/3$ after approximately $2/3\alpha^2$ iterations of the random walk. Since this is a factor of 3 less than the analogous time scale associated with $\mathrm{E}\left\{\hat{p}_k\right\}$, the random walk approaches a uniform distribution over the sphere with a characteristic time constant of $2/\alpha^2$.

In the case that the initial vector $\hat{p}_0$ instead possesses a uniform density, $f_0(\theta,\phi) = Y_0^0(\theta,\phi)/\sqrt{4\pi}$, Eq. (4.19) and Eq. (4.22) imply $\mathrm{E}\{\hat{p}_k\} = 0$, and $\mathrm{E}\left\{p_{x,k}^2\right\} = \mathrm{E}\left\{p_{y,k}^2\right\} = \mathrm{E}\left\{p_{z,k}^2\right\} = 1/3$ for all steps k. Further, by symmetry the off-diagonal elements of $\mathrm{E}\left\{\hat{p}_k\hat{p}_k^{\mathrm{T}}\right\}$ average to zero, resulting in $\mathrm{E}\left\{\hat{p}_k\hat{p}_k^{\mathrm{T}}\right\} = \mathbf{I}_3/3$, consistent with the initial uniform distribution.

4.2.1 Expectation of the emulator rotation matrix

We will next demonstrate that the autocorrelation matrix
$\mathbf{C}_{\mathbf{R}}(\omega,k;\Delta\omega,n) = \mathrm{E}\left\{\mathbf{R}(\omega,k)^{\mathrm{T}}\mathbf{R}(\omega+\Delta\omega,k+n)\right\}$ associated with the emulator rotation $\mathbf{R}(\omega,k) = \exp(\omega\tau\,\hat{p}_k\times)$ is proportional to the identity matrix,
$\mathbf{C}_{\mathbf{R}}(\omega,k;\Delta\omega,n) = g(\omega,k;\Delta\omega,n)\,\mathbf{I}_3$ and that $\mathbf{C}_{\mathbf{R}}(\omega,k;\Delta\omega,n) \approx \mathbf{C}_{\mathbf{R}}(\Delta\omega,n)$ implying that $\mathbf{R}(\omega,k)$ is approximately wide sense stationary with respect to both frequency and time.

Specializing for simplicity to case (2), all terms in the expansion of $\mathbf{C}_{\mathbf{R}}(\omega,k;\Delta\omega,n)$ involving an odd number of occurrences of either $\hat{p}_k$ or $\hat{p}_{k+n}$ vanish after averaging, while from the preceding analysis $\mathrm{E}\left\{\hat{p}_k\hat{p}_k^{\mathrm{T}}\right\} = \mathbf{I}_3/3$ for all k. Therefore only the two-time correlation functions $\mathrm{E}\{(\hat{p}_k\times)(\hat{p}_{k+n}\times)\}$ and $\mathrm{E}\left\{(\hat{p}_k\hat{p}_k^{\mathrm{T}})(\hat{p}_{k+n}\hat{p}_{k+n}^{\mathrm{T}})\right\}$ require evaluation. From the spherical symmetry of the distribution of $\hat{p}_0$, $\mathrm{E}\left\{\hat{p}_{k+n}\hat{p}_k^{\mathrm{T}}\right\} = \mathrm{E}\{\hat{p}_{k+n}\cdot\hat{p}_k\}\,\mathbf{I}_3/3$ which can be demonstrated by aligning $\hat{p}_k$ with the polar axis, $\hat{z} = [0,0,1]^{\mathrm{T}}$. By symmetry all elements of $\mathrm{E}\left\{\hat{p}_{k+n}\hat{z}^{\mathrm{T}}\right\}$ then average to zero at step $k+n$ except for the lower right diagonal $p_{z,k+n} = \hat{p}_{k+n}\cdot\hat{z}$ corresponding to the angle subtended by $\hat{p}_{k+n}$ and $\hat{p}_k = \hat{z}$. The factor $1/3$ arises from the uniform density $f_0(\theta,\phi)$ so that

$$\begin{aligned}\mathrm{E}\{(\hat{p}_k\times)(\hat{p}_{k+n}\times)\} &= \mathrm{E}\left\{\left[\hat{p}_{k+n}\hat{p}_k^{\mathrm{T}} - \hat{p}_{k+n}\cdot\hat{p}_k\mathbf{I}_3\right]\right\}\\ &= -\frac{2}{3}\mathrm{E}\{\hat{p}_{k+n}\cdot\hat{p}_k\}\,\mathbf{I}_3.\end{aligned} \tag{4.25}$$

Since $\mathrm{E}\{\hat{p}_{k+n}\cdot\hat{p}_k\}$ only depends on the relative orientations of $\hat{p}_{k+n}$ and $\hat{p}_k$, from Eq. (4.20)

$$\mathrm{E}\{\hat{p}_{k+n}\cdot\hat{p}_k\} = [P_1(\cos\alpha)]^n. \tag{4.26}$$

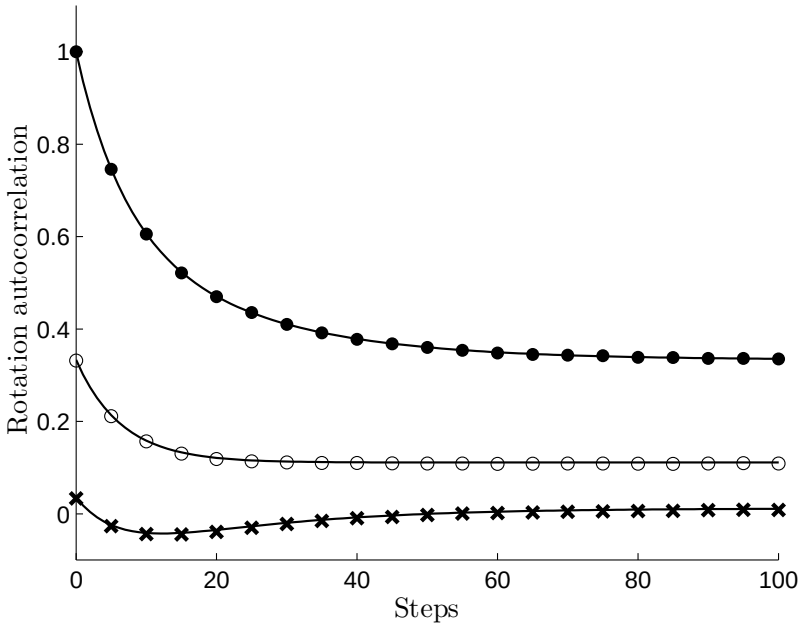


Figure 4.2: The exact (solid line) autocorrelation of the rotation matrix, Eq. (4.28), and the result of 5×10^4 repeated random walk simulations with an angular step size $\alpha = \pi/10$, a Maxwellian distributed DGD with mean $\mathrm{E}\{\tau\} = \sigma_\tau \sqrt{8/\pi} = 5$ ps and a uniform initial distribution of polarizations $\hat{p}_0$ for frequency offsets $\Delta\omega = 0$ ($\bullet$ markers), $1/\sigma_\tau$ ($\circ$ markers) and $\sqrt{3}/\sigma_\tau$ ($\times$ markers).

Similarly, after orienting the $\hat{z}$ axis to coincide with the direction of $\hat{p}_k$ at the k:th temporal step,

$$\mathrm{E}\left\{\left(\hat{z}\hat{z}^{\mathrm{T}}\right)\left(\hat{p}_{k+n}\hat{p}_{k+n}^{\mathrm{T}}\right)\right\}$$

$$= \mathrm{E}\left\{\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ p_{x,k+n}\,p_{z,k+n} & p_{y,k+n}\,p_{z,k+n} & p_{z,k+n}^2 \end{bmatrix}\right\},$$

where by symmetry only $\mathrm{E}\left\{p_{z,k+n}^2\right\} = \mathrm{E}\left\{\cos^2\theta\right\}$ is nonzero after averaging. Including an additional factor of $1/3$ to incorporate the random initial orientation of $\hat{p}_k$ we find from Eq. (4.23)

$$\mathrm{E}\left\{(\hat{p}_k\hat{p}_k^{\mathrm{T}})(\hat{p}_{k+n}\hat{p}_{k+n}^{\mathrm{T}})\right\} = \left(\frac{1}{9} + \frac{2}{9}\left[P_2(\cos\alpha)\right]^n\right)\mathbf{I}_3. \tag{4.27}$$

Finally, combining the above results with Rodrigues' formula, Eq. (4.3), for $\mathbf{R}(\omega,k)$ yields the autocorrelation of the emulator rotation matrix $\mathrm{E}\left\{\mathbf{R}(\omega,k)^{\mathrm{T}}\mathbf{R}(\omega+\Delta\omega,k+n)\right\} = g(\omega,k;\Delta\omega,n)\,\mathbf{I}_3$ in which

$$\begin{aligned} g(\omega,k;\Delta\omega,n) = \; & \frac{1}{9}\Big(1 + 2\left[P_2(\cos\alpha)\right]^n\Big) \\ & + \frac{1}{9}\Big(2 + 3\left[P_1(\cos\alpha)\right]^n + \left[P_2(\cos\alpha)\right]^n\Big) \\ & \times \left[1 - \sigma_\tau^2\Delta\omega^2\right]\exp(-\sigma_\tau^2\Delta\omega^2/2), \end{aligned} \tag{4.28}$$

τ is a Maxwellian distributed random variable with $\mathrm{E}\left\{\tau\right\} = \sigma_\tau\sqrt{8/\pi}$,

$$\mathrm{E}\left\{\cos(\tau\Delta\omega)\right\} = \left[1 - \sigma_\tau^2\Delta\omega^2\right]\exp(-\sigma_\tau^2\Delta\omega^2/2), \tag{4.29}$$

and $\mathrm{E}\left\{\cos(\tau\omega)\right\} \approx 0$ for $\sigma_\tau\omega \gg 1$. Similarly, $\mathrm{E}\left\{\mathbf{R}(\omega,k)\right\} = \mathbf{I}_3/3$, while from Eq. (4.28) $g(\omega,k;\Delta\omega,n) = g(\Delta\omega,n)$. Accordingly, both the rotation matrix $\mathbf{R}(\omega,k)$ and the emulated PMD vector $\vec{\Omega}_N(\omega,k)$ are wide sense stationary with respect to both optical frequency and time if $\sigma_\tau\omega \gg 1$ and $f_0(\theta,\phi)$ is uniform. Otherwise, the evolution of the random walk is

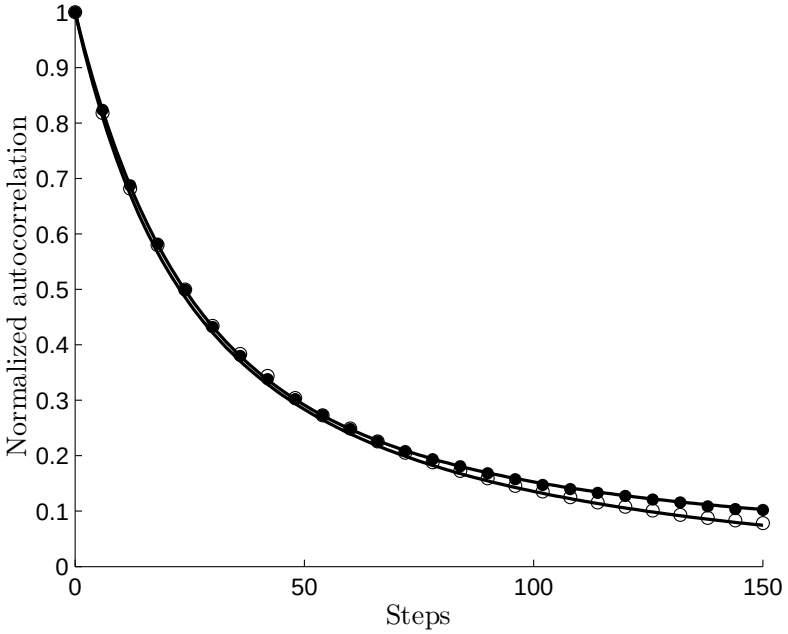


Figure 4.3: The temporal autocorrelation function of PMD emulators with $N = 5$ ($\circ$) and $N = 25$ ($\bullet$ markers) sections normalized by τ^2_{rms}, as evaluated from the results of 5×10^4 $t_d = 15$ step random walks. The analytic result is displayed as the solid line.

influenced by the starting condition so that the correlation function $g(\omega, k; \Delta\omega, n)$ becomes independent of k only after a number of the order of $\sim 2/\alpha^2$ steps.

In Fig. 4.2 the exact autocorrelation of the rotation matrix, Eq. (4.28) (solid line), is compared to the results of 5×10^4 random walk simulations for an angular step size $\alpha = \pi/10$, a Maxwellian distributed DGD with mean $\mathrm{E}\{\tau\} = 5$ ps and a frequency ω corresponding to a $\lambda = 1550$ nm optical wavelength. Results are shown for frequency offsets $\Delta\omega = 0$ ($\bullet$ markers), $1/\sigma_\tau$ ($\circ$ markers), and $\sqrt{3}/\sigma_\tau$ ($\times$ markers) the latter two of which correspond to the zero and minimum values of $\mathrm{E}\{\cos(\tau\Delta\omega)\}$, respectively. Note that as n increases, $g(\Delta\omega, n)$ approaches a limiting value $1/9+2/9\left[1-\sigma_\tau^2\Delta\omega^2\right]\exp(-\sigma_\tau^2\Delta\omega^2/2)$ that varies with frequency from $1/9-4/9\exp(-3/2) \approx 0.012$ at $\Delta\omega = \sqrt{3}/\sigma_\tau$ to $1/3$. Again the Monte-Carlo and analytic results nearly coincide.

4.2.2 PMD emulation

The temporal statistics of the PMD vector $\vec{\Omega}_N(\omega, k)$ are now obtained by allowing each $\hat{p}_{m,k}$ in Eq. (4.1) to execute an independent random walk on the unit sphere as discussed in Section 4.2; that is, in terms of the PMD vector of the m:th emulator section at time index k, $\vec{\tau}_{m,k} = \tau_m \hat{p}_{m,k}$, and $C_\tau(k; n) = \mathrm{E}\{\vec{\tau}_{m,k} \cdot \vec{\tau}_{m,k+n}\} = 3\sigma_\tau^2\,[P_1(\cos\alpha)]^n$. The PMD vector autocorrelation after N sections is therefore given by

$$C_N(\Delta\omega, n) = \frac{\tau_{\mathrm{rms}}^2}{N}\,[P_1(\cos\alpha)]^n \left[\frac{1-g^N(\Delta\omega, n)}{1-g(\Delta\omega, n)}\right], \tag{4.30}$$

where $g(\Delta\omega, n)$ is defined according to Eq. (4.28). Since the rms DGD of the N section emulator, τ_{rms}, exceeds that of each section by $\sqrt{N}$, $\mathrm{E}\{\tau_m^2\} = 3\sigma_\tau^2 = \tau_{\mathrm{rms}}^2/N$.

To relate the characteristic decorrelation time, t_d, of the PMD vector to the number of emulator sections N and the angular step size α, consider $g^N(\Delta\omega, n)$ in the $n \ll 1/\alpha^2$ and $\Delta\omega \ll 1/\sigma_\tau$ limit. Approximating $[P_1(\cos\alpha)]^n \approx 1-\alpha^2 n/2$, $[P_2(\cos\alpha)]^n \approx 1-3\alpha^2 n/2$

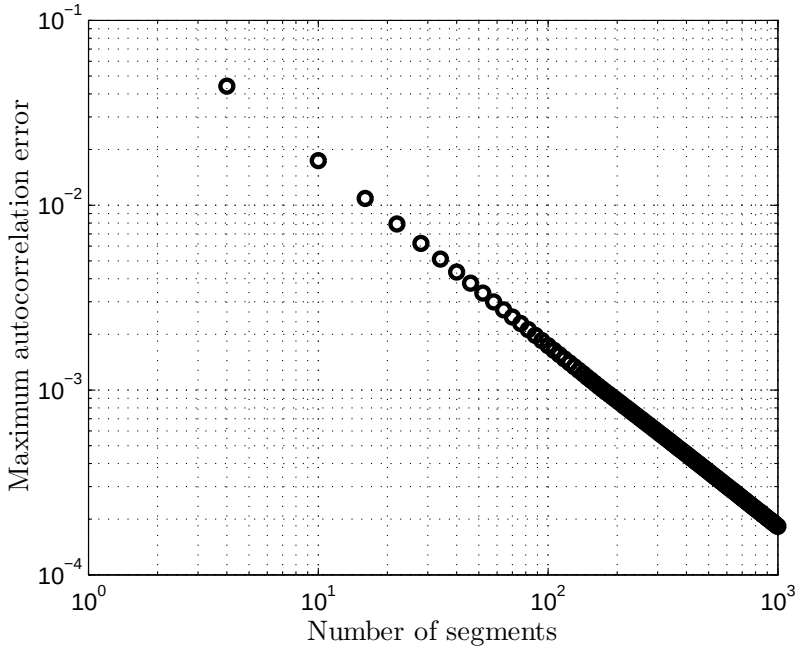


Figure 4.4: The maximum difference, $\epsilon_{\max}$, between the temporal autocorrelation of an N section PMD emulator, Eq. (4.30), and the $N \to \infty$ result normalized by τ_{rms}^2.

gives $g(\Delta\omega, n) \approx 1 - 2\alpha^2 n/3 - \sigma_\tau^2 \Delta\omega^2$ and

$$g^N(\Delta\omega, n) \approx \mathrm{e}^{-2N\alpha^2 n/3} \mathrm{e}^{-N\sigma_\tau^2 \Delta\omega^2}, \tag{4.31}$$

yielding the PMD decorrelation frequency $\omega_d = 1/\sqrt{N\sigma_\tau^2} = \sqrt{3}/\tau_{\mathrm{rms}}$ [61, 108], and the decorrelation time

$$t_d = \frac{3}{2} \frac{1}{N\alpha^2} \tag{4.32}$$

over which the PMD vector $\vec{\Omega}_N(\omega, k)$ evolves into a statistically independent state. Specializing Eq. (4.30) further to $\Delta\omega = 0$ and $n \ll N t_d$, leads to $P_1(\cos\alpha) \approx 1$ and

$$C_N(0, n) \approx \tau_{\mathrm{rms}}^2 \frac{\left(1 - \mathrm{e}^{-n/t_d}\right)}{n/t_d} \tag{4.33}$$

which reproduces the known temporal autocorrelation function of the PMD vector, $C_\infty(n)$, in the $N \to \infty$ limit [61, 80]. Consequently, selecting an angular step size $\alpha = \sqrt{3/(2Nt_d)}$ reproduces the known asymptotic result for the temporal autocorrelation statistics with decorrelation time t_d.

Next, in Fig. 4.3 we display the PMD temporal autocorrelation function normalized by τ_{rms}^2 for PMD emulators with $N = 5$ ($\circ$ markers) and $N = 25$ ($\bullet$ markers) sections as evaluated from 5×10^4 random walks, while the solid line corresponds to the analytic result of Eq. (4.30). In this calculation, the rms DGD of the emulator and the decorrelation time were set to $\tau_{\mathrm{rms}} = 21.7$ ps and 15 steps, respectively. Here $C_N(0, n)$ varies perceptibly with N only for $N < 25$, while the convergence of Eq. (4.30) to the infinite-section limit is further illustrated in Fig. 4.4 in which the circles indicate the variation of the maximum difference $\epsilon_{\max} = \max_n |C_N(0, n) - C_\infty(n)|/\tau_{\mathrm{rms}}^2$ with the number of emulator sections N. From the figure, $\epsilon_{\max} \approx 0.175/N$ implying that for $N > 18$ sections the maximum error in the emulated PMD temporal autocorrelation is $< 1\%$.

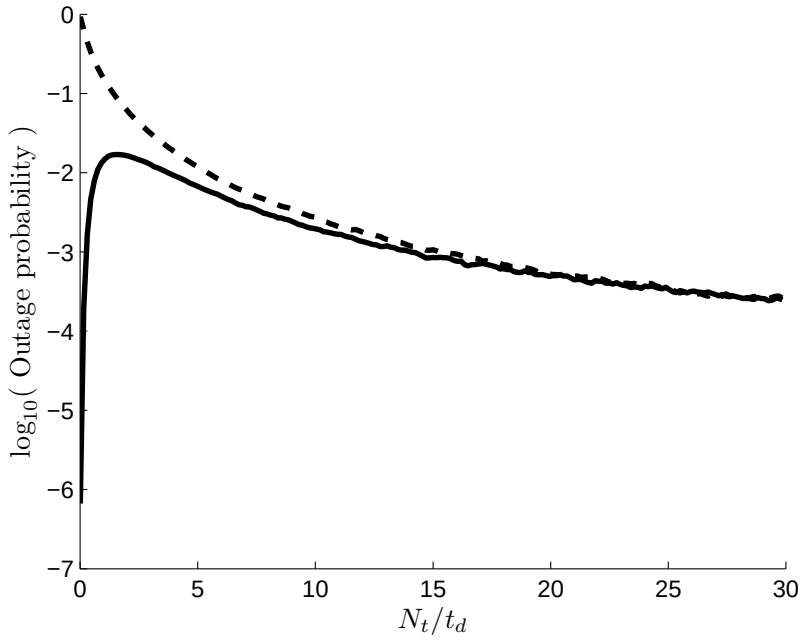


Figure 4.5: The emulated PMD-induced system outage probability after N_t temporal steps of an $N = 50$ section emulator with an rms DGD of $\tau_{\text{rms}} = 21.7$ ps and a decorrelation time of $t_d = 20$ steps calculated with 10^6 realizations of Eq. (4.1) and an initial PMD vector $\vec{\Omega}_N(\omega, 0)$ parallel (solid line) and perpendicular (dashed line) to the signal SOP.

4.3 System applications

4.3.1 Outage probability evaluation

We now employ the above formalism to determine the evolution of the PMD-induced outage probability of a system with time. An outage is assumed to occur if the magnitude of the component of the PMD vector

$$\tau_{\perp}(k) = |\vec{\Omega}_N(\omega, k) \times \hat{s}| \tag{4.34}$$

orthogonal to a signal's incident state of polarization (SOP), $\hat{s}$, exceeds a specified threshold, here $\tau_{\text{outage}} = 2.77\,\tau_{\text{rms}}$ [80, 93]. We first apply the Brownian bridge technique [107] to generate an initial set of Gaussian distributed PMD vectors $\vec{\tau}_{m,0}$ for each of the $m = 1, \ldots, N$ emulator sections such that the total PMD vector $\vec{\Omega}_N(\omega, 0)$ of the initial state is either parallel or perpendicular to $\hat{s}$ with magnitude $|\vec{\Omega}_N(\omega, 0)| = 3\tau_{\text{rms}}$. In the first case $\tau_{\perp}(0) = 0$ so that an outage only occurs at a later time while in the second case only initial vectors with $\tau_{\perp}(0) > \tau_{\text{outage}}$ are selected. Subsequently, we set $\tau_m = |\vec{\tau}_{m,0}|$ and $\hat{p}_{m,0} = \vec{\tau}_{m,0}/|\vec{\tau}_{m,0}|$, since for Gaussian distributed $\vec{\tau}_{m,0}$, τ_m is Maxwellian while $\hat{p}_{m,0}$ is uniformly distributed over solid angle. The unit vector $\hat{p}_{m,k}$ of each segment then performs a random walk according to Eq. (4.12).

Fig. 4.5 illustrates the result of this calculation after N_t temporal steps of an $N = 50$ section emulator with an rms DGD of $\tau_{\text{rms}} = 21.7$ ps and a decorrelation time of $t_d = 20$ steps involving 10^6 realizations of Eq. (4.1) for an initial PMD vector $\vec{\Omega}_N(\omega, 0)$ parallel (solid line) and perpendicular (dashed line) to the signal SOP. Evidently, for $N_t < 3t_d$ the outage probability increases rapidly when the PMD vector is initially parallel to the signal SOP, while for $N_t > 10\,t_d$ the PMD vector becomes effectively uncorrelated from its initial state and the outage probabilities for either initial condition coincide, as previously predicted through a stochastic differential equation approach [3].

4.3.2 Multicanonical evaluation of PMD temporal dynamics

The multicanonical procedure was first introduced in the context of statistical mechanics [7] and later adapted to communication systems in [122, 123, 125]. Subsequently it was realized that multicanonical Markov chain Monte Carlo methods could be applied to the time evolution of a system and in particular to adapt such calculations towards physically unlikely but practically important configurations [127–129]. This section accordingly discusses the manner in which the multicanonical framework can be employed to bias a random walk toward configurations with large orthogonal PMD vector components after a given number of temporal steps.

In particular, consider a system described by N_E observables, here written as a vector $\vec{E}(\vec{a})$, that depend on N_a stochastic parameters $\vec{a}$ (in this paper $N_E = 1$ and E is identified with the magnitude of the orthogonal component of the PMD vector $E = \tau_\perp(N_t)$ of Eq. (4.34) after N_t time steps). To compute a stochastic function $f(\vec{E})$, here the pdf, an appropriate region of $\vec{E}$ is partitioned into N_b histogram bins centered at $\vec{E}_l$ with $l = 1, 2, \ldots, N_b$. Two histograms are allocated and initialized to unity to store the current estimate of the unnormalized function (pdf), $f^0(\vec{E}_l)$, and the intermediate result of each iteration while the system variables are assigned values $\vec{a}^{\text{cur}}$ selected from a stochastic distribution. Subsequently, small variations $\vec{a}^{\text{cur}} \rightarrow \vec{a}^{\text{new}}$ are applied to these variables described by an effectively arbitrary distribution function [125], which transforms the observables from $\vec{E}^{\text{cur}}$ to $\vec{E}^{\text{new}}$.

This transition is then accepted with a probability $\min\left[1, f^0(\vec{E}^{\text{cur}})/f^0(\vec{E}^{\text{new}})\right]$, in which case $\vec{a}^{\text{cur}}$ and $\vec{E}^{\text{cur}}$ are equated to $\vec{a}^{\text{new}}$ and $\vec{E}^{\text{new}}$. Otherwise the subsequent time step instead again employs $\vec{a}^{\text{cur}}$ and $\vec{E}^{\text{cur}}$. In both cases, however, the histogram entry $H(\vec{E}_l)$ corresponding to the new $\vec{E}^{\text{cur}}$ is increased by one. After M transitions, the bias introduced through the acceptance rule is removed and an improved estimate, $f^1(\vec{E})$, of $f(\vec{E})$ is generated according to $f^1(\vec{E}) = cf^0(\vec{E})H(\vec{E})$, where c appropriately normalizes $f^1(\vec{E})$.

The elements of $H(\vec{E})$ are then reinitialized to unity and the process repeated with $f^1(\vec{E})$ replacing $f^0(\vec{E})$. The transition rule in this and subsequent iterations increases the sampling probability of states with small $f(\vec{E})$ in such a manner that the likelihood of visiting configurations within a region $\vec{E}_0 - \delta\vec{E} < \vec{E} < \vec{E}_0 + \delta\vec{E}$ becomes independent of $\vec{E}_0$ with an increasing number of samples and iterations [122].

A difficulty associated with applying the multicanonical method to outage calculations is that small changes in the DGD τ_m are multiplied in Eq. (4.1) by the optical frequency ω, decorrelating the two successive Markov chain states and reducing algorithmic efficiency. Accordingly, we employ a modified emulator for which the unit vectors $\hat{r}_{m,k}$ evolve independently according to Eq. (4.12) for a spherically uniform initial distribution of $\hat{r}_{m,0}$, while each $\vec{\tau}_m$ represents a time-independent Gaussian distributed random vector with variance σ_τ^2. The PMD vector is then obtained from

$$\vec{\Omega}_m(k) = \vec{\tau}_m + \mathbf{U}_m(k)\,\vec{\Omega}_{m-1}(k), \tag{4.35}$$

where $\mathbf{U}_m(k) = \exp\left(\psi\,\hat{r}_{m,k}\times\right)$ and ψ equals the m-independent constant specified below. The final, emulated PMD vector is obtained by applying an overall random rotation $\vec{\Omega}'_N(k) = \mathbf{U}_0(k)\,\vec{\Omega}_N(k)$ after the final emulator section that suppresses the background PMD autocorrelation and therefore improves the emulated temporal correlations for small N.

Applying the results of Section 4.2.1 yields $\mathrm{E}\left\{\mathbf{U}_m(k)^{\mathrm{T}}\mathbf{U}_m(k+n)\right\} = h(n)\mathbf{I}_3$ in which

$$\begin{aligned} h(n) = {} & \frac{1}{9}\left(2\cos\psi + 1\right)^2 + \frac{2}{3}\sin^2\psi\;\left[P_1(\cos\alpha)\right]^n \\ & + \frac{2}{9}\left(1-\cos\psi\right)^2\left[P_2(\cos\alpha)\right]^n \end{aligned} \tag{4.36}$$

and the corresponding PMD autocorrelation function

$$\begin{aligned} C'_N(n) &= \mathrm{E}\left\{\vec{\Omega}'_N(k)\cdot\vec{\Omega}'_N(k+n)\right\} \\ &= \frac{\tau_{\mathrm{rms}}^2}{N}\,h(n)\left[\frac{1-h^N(n)}{1-h(n)}\right]. \end{aligned} \tag{4.37}$$

The variable ψ is computed from Eq. (4.37) by minimizing $C'_N(n)$ in the limit $n \to \infty$ yielding $\psi = 2\pi/3$. The decorrelation time t_d of the modified emulator is then obtained as in preceding sections from $h^N(n)$ in the $n \ll 1/\alpha^2$ limit. Since with $\psi = 2\pi/3$ $h(n) \approx \exp(-\alpha^2 n)\cosh(\alpha^2 n/2)$ and $h^N(n) \approx \exp(-N\alpha^2 n)$, we find $t_d = 1/(N\alpha^2)$ and the decorrelation time of the modified emulator is smaller by a factor of 2/3 than Eq. (4.32) for identical angular step sizes α.

The initial state of the Markov chain, $\vec{a}^{\rm cur}$, is generated by selecting $\hat{r}^{\rm cur}_{m,0}$ from a uniform distribution together with a series of Gaussian distributed random vectors $\vec{w}^{\rm cur}_m$ with variance σ^2_τ for each emulator section m. The unit vectors $\hat{r}^{\rm cur}_{m,k}$ for all intermediate times $k = 0, \ldots, N_t$ are determined from Eq. (4.12). The state $\vec{a}^{\rm cur}$ is therefore parametrized by a total of $N_a = (N_t + 1)(N + 1) + N$ stochastic vectors containing $N_t + 1$ elements for each emulator section, as well as $\mathbf{U}_0(k)$ and $\vec{w}^{\rm cur}_m$. The initial, $k = 0$, PMD vector $\vec{\Omega}'_N(0)$ is limited to the desired region of output variables by applying the Brownian bridge method to $\vec{w}^{\rm cur}_m$ to calculate the current $\vec{\tau}_m$, while the output system observable $E^{\rm cur} = \tau_\perp(N_t)$ is obtained from Eq. (4.35). The updated $\vec{w}^{\rm new}_m$ are then obtained by adding the Gaussian distributed random vectors $\Delta\vec{w}_m$ to the current state vector $\vec{w}^{\rm new}_m = \vec{w}^{\rm cur}_m + \epsilon\Delta\vec{w}_m$ with $\epsilon = 0.2$ and each component $w^{\rm new}_{m,a}$ with $a = x, y, z$ of $\vec{w}^{\rm new}_m$ is accepted with probability $\min\left[1, f_w(w^{\rm cur}_{m,a})/f_w(w^{\rm new}_{m,a})\right]$, where $f_w(x) = \exp(-x^2/2\sigma^2_\tau)/\sqrt{2\pi\sigma^2_\tau}$ is the Gaussian pdf associated with each $w_{m,a}$. The distribution of the elements of $\vec{w}^{\rm new}_m$ then approaches $f_w(x)$ [53] after a sufficiently large number of transitions.

Biasing the PMD temporal dynamics after N_t time steps given a specified initial PMD vector requires that the emulator state be altered at all intermediate times $k = 0, \ldots, N_t$ while preserving the constant angular step size between temporal increments. After initializing $\hat{r}^{\rm new}_{m,k}$ we subsequently rotate the sequence of unit vectors $\hat{r}^{\rm new}_{m,n}$ for steps with $n > k$ by a random angle about the unit vector $\hat{r}^{\rm new}_{m,k}$ obtained at time index k. Since these random rotations operate equally on all members of the time series, the relative angle between $\hat{r}^{\rm new}_{m,k}$ and $\hat{r}^{\rm new}_{m,k+1}$ is preserved. That is, after generating a series $\theta_{m,k}$ of zero-mean Gaus-

sian distributed angles with standard deviation $\sigma_\theta = \pi/60$ for each emulator section m and temporal index $k = 0, \ldots, N_t$, in the first step of the algorithm a random rotation $\mathbf{P}_m(0) = \exp\left(\theta_{m,0}\, \hat{u}_m \times\right)$ is applied to the unit vectors for all $k = 0, \ldots, N_t$ to form a new sequence $\hat{r}^{\rm new}_{m,k} = \mathbf{P}_m(0)\, \hat{r}^{\rm cur}_{m,k}$ in which each $\hat{u}_m$ is selected from a spherically uniform distribution. In the subsequent $k = 1$ step the random rotation $\mathbf{P}_m(1) = \exp\left(\theta_{m,1}\, \hat{r}^{\rm new}_{m,0} \times\right)$ is applied only to $n \geq 1$ replacing $\hat{r}^{\rm new}_{m,n}$ with $\mathbf{P}_m(1)\, \hat{r}^{\rm new}_{m,n}$. Similarly, to implement k:th step of the recursion $\mathbf{P}_m(k) = \exp\left(\theta_{m,k}\, \hat{r}^{\rm new}_{m,k-1} \times\right)$ is evaluated and $\hat{r}^{\rm new}_{m,n}$ is then replaced by $\mathbf{P}_m(k)\, \hat{r}^{\rm new}_{m,n}$ for all $k \leq n \leq N_t$. The resulting $\hat{r}^{\rm new}_{m,k}$ together with the $\vec{w}^{\rm new}_m$ form the transformed state $\vec{a}^{\rm new}$ of the emulator. The updated $\vec{\tau}_m$ are finally obtained by applying the Brownian bridge method to $\vec{w}^{\rm new}_m$ and $E^{\rm new} = \tau_\perp(N_t)$ is computed from Eq. (4.35).

Fig. 4.6 displays the pdf of $\tau_\perp$ after $N_t = t_d/5$ and $N_t = t_d$ temporal steps of an $N = 50$ section PMD emulator with $\tau_{\rm rms} = 21.7$ ps and $t_d = 15$ steps such that the the initial PMD vector at $k = 0$ is oriented in the $\hat{s} = [1, 0, 0]^{\rm T}$ direction with a magnitude $3\tau_{\rm rms}$. The dashed lines in Fig. 4.6 are obtained from the modified PMD emulator model Eq. (4.35) with three 3.33×10^5 sample multicanonical iterations, while the solid lines are associated with 10^6 random realizations of Eq. (4.1). Similarly, the contour plot of Fig. 4.7 depicts $\log_{10}$ of the cumulative distribution function (cdf) of $\tau_\perp(N_t)$. This calculation employs three 3.33×10^5 sample multicanonical iterations with $N_t \leq 8t_d$ random walk steps and an outage threshold $\tau_{\rm outage} = 2.77\tau_{\rm rms}$. The outage probability as expected increases rapidly for $N_t < 3t_d$ and subsequently converges to the anticipated Rayleigh distribution [92, 93]. For an identical number of samples, the modified emulator method evaluates regions of the pdf that are 7 orders of magnitude smaller than those obtained from the standard Monte-Carlo procedure while accurately preserving the temporal correlations of Eq. (4.1).

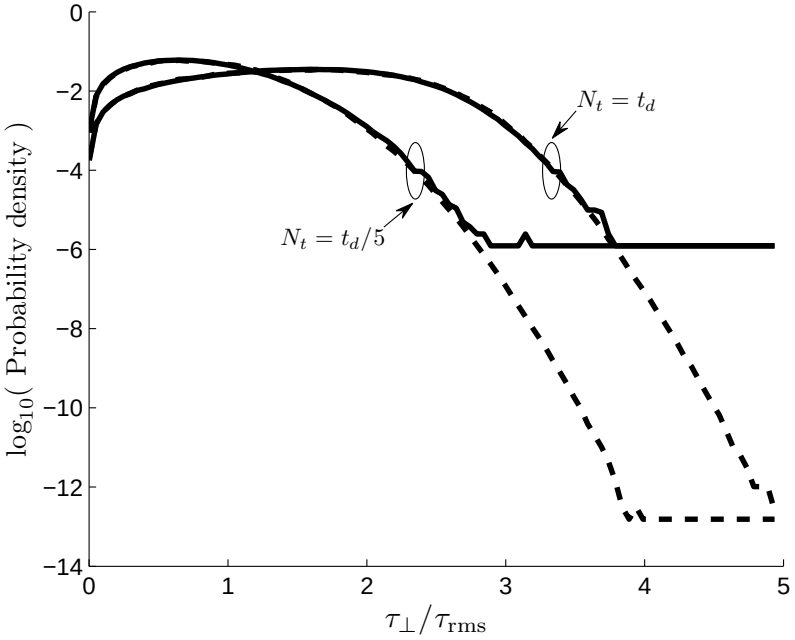


Figure 4.6: The pdf of $\tau_\perp$ after $N_t = t_d/5$ and $N_t = t_d$ time steps of an $N = 50$ section PMD emulator with $\tau_{\rm rms} = 21.7$ ps and $t_d = 15$ steps for a simulation with three multicanonical iterations of 3.33×10^5 random walks (dashed line) and 10^6 realizations of Eq (4.1) (solid line).

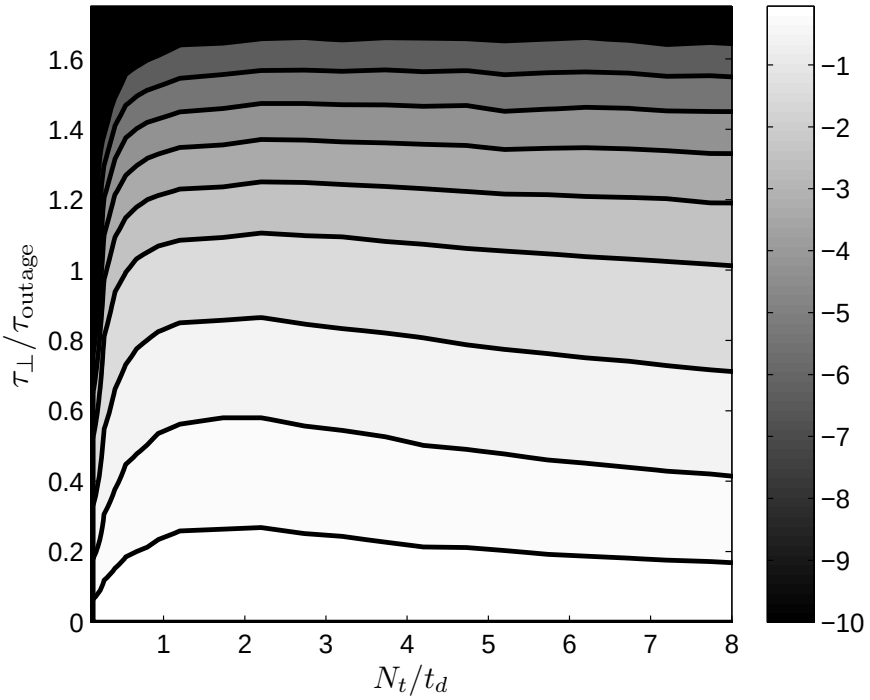


Figure 4.7: $\log_{10}$ of the cumulative distribution function computed with three 3.33×10^5 sample multicanonical iterations for $\tau_\perp(N_t)$, normalized by $\tau_{\rm outage} = 2.77\,\tau_{\rm rms}$, with $0 < N_t < 8\,t_d$ random walk steps and the emulator configuration of Fig. 4.6.

4.4 Conclusions

We have presented a model for the temporal evolution of PMD in optical fibers derived from a random walk on the unit sphere that accurately emulates the temporal autocorrelation statistics of PMD with relatively few fiber segments. The increased computational efficiency afforded by this method allows simulations of the PMD temporal dynamics to be preferentially biased towards regions of low probability using the multicanonical method, providing a potential significant improvement to estimates of for example the probability of a network outage and its duration. Finally, extensions of our approach to more general optical systems including PDL or a finite number of active polarization hinge sites is anticipated to provide more insight into the polarization temporal statistics and dynamics for improved network equalization and control.

Chapter 5

Transition Matrix Analysis of System Outages

We adapt the transition matrix method to time-dependent communication systems. We then calculate the distribution of the outage times of an optical fiber system impaired by stochastically varying polarization-mode dispersion.

5.1 Introduction

In communication system simulations, biased Monte-Carlo methods, such as the multicanonical or importance sampling techniques, have been employed to determine static quantities such as the probability distribution functions (pdfs) of the differential group delay (DGD) or of the eye-opening penalty [8,74,123]. Often, however, the dynamic behaviour of a system, for example the distribution of the times required to transition between two groups of states with different physical properties must be calculated. Here we show that the multicanonical algorithm, reformulated as a "transition matrix" method [34, 110, 115],

can be applied to such dynamic system calculations (note, however, that the error of static multicanonical polarization-mode dispersion (PMD) calculations was examined with this procedure in [72]).

This chapter is organized as follows. In Section 5.2, we will review multicanonical procedures for sampling physically unlikely system configurations. We then apply the technique to evaluate the relative probability for transitions between any two states of a system over a single time step. From the resulting formalism, we compute the pdf of the time duration of outages in an optical system affected by first-order PMD. This method extracts information from the low-probability regions of the pdf far more efficiently than conventional Monte-Carlo techniques.

5.2 The Multicanonical procedure

As in Chapter 4, we consider a general physical system described by a vector formed from N_E system observables $\vec{E}(\vec{a})$ that in turn depend on the values of N_a stochastically varying parameters $\vec{a}$. The multicanonical method rapidly estimates stochastic functions $f(\vec{E})$ in regions of sample space with a small probability of occurrence where a standard Monte Carlo calculation would require long computation times. This is accomplished through an iterative statistical procedure for biasing the sample space. In the standard case that $f(\vec{E})$ represents the pdf of $\vec{E}$, the physically relevant region of the system observables (solution space) $\vec{E}$ is first divided into N_B histogram bins labeled with index $m = 1, 2, \ldots, N_B$. All elements of two histograms, one for an estimate of the unnormalized pdf, $f_m^{(i)}$, with $i = 0$ initially, and a second for storing the intermediate results of each iteration H_m are initialized to unity. A random set of system variables $\vec{a}^{\,\mathrm{cur}}$ is generated and the observables $\vec{E}(\vec{a}^{\,\mathrm{cur}})$ with bin index m_{cur} computed. Next, these variables are randomly perturbed according to $\vec{a}^{\,\mathrm{new}} = \vec{a}^{\,\mathrm{cur}} + \delta\vec{a}$, where $\delta\vec{a}$ represents a small random perturbation that can be chosen from

an effectively arbitrary probability distribution, cf. Ref. [125], resulting in new observables $\vec{E}(\vec{a}^{\,\mathrm{new}})$ and m_{new} (in the calculations of this chapter, the distribution is associated with the system evolution over a small time step). This transition is accepted with probability $\min\{1, f^{(0)}_{m_{\mathrm{cur}}}/f^{(0)}_{m_{\mathrm{new}}}\}$, in which case $\vec{a}^{\,\mathrm{cur}}$ is set to $\vec{a}^{\,\mathrm{new}}$; otherwise, $\vec{a}^{\,\mathrm{cur}}$ is reemployed in the following iteration. In either case, the histogram element $H_{m_{\mathrm{cur}}}$ containing the new value of $\vec{E}^{\mathrm{cur}}$ is incremented by unity. After M transitions, a new estimate $f^{(1)}_m$ of $f(\vec{E})$ is obtained by correcting for the bias in the system variables according to $f^{(1)}_m = c\, f^{(0)}_m H_m$ (the normalization factor c is chosen such that $f^{(1)}_m$ describes a probability distribution). In the following iteration, $f^{(1)}_m$ is substituted for $f^{(0)}_m$, the elements of H_m are reset to unity, and the above steps repeated. The likelihood of generating states with small $f(\vec{E})$ then increases such that the sampling probability becomes independent of $\vec{E}$ after a sufficiently large number of samples and iterations [123].

5.3 Transition matrix method

The multicanonical method can be employed to determine the relative probability for transitions from one state of the system variables to a second. In particular, after each multicanonical system realization, the (l,k):th element, $\mathcal{T}_{lk}$, of an unbiased and unnormalized transition matrix $\boldsymbol{\mathcal{T}}$ is incremented by unity for every *accepted or rejected* transition from a state in the k:th to a state in the l:th histogram bin. The columns of $\boldsymbol{\mathcal{T}}$ are subsequently normalized to unity since the sum of the probabilities of all transitions out of the initial state k is one, yielding the unbiased, normalized transition matrix $\mathbf{T}$. After the last, L:th iteration, a second *biased multicanonical transition matrix* $\mathbf{B}$ is constructed by multiplying each element T_{lk} of $\mathbf{T}$ with the multicanonical acceptance probability $\min\{1, f^{(L)}_l/f^{(L)}_k\}$ and normalizing each column of this matrix separately to unity, so that B_{lk} corresponds to the likelihood of a transition from the k:th to the l:th state in the multicanonical procedure.

While the above procedure is based on the standard multicanonical formalism, the acceptance rule can be formulated instead in terms of a dynamically updated estimate of $\boldsymbol{\mathcal{T}}$, such that regions of low transition probability are preferentially sampled [115]. In this case, $\boldsymbol{\mathcal{T}}$ is initialized such that all elements $\mathcal{T}_{lk} = 1$. A transition between two states $\vec{E}_k$ and $\vec{E}_l$ generated in the Markov chain $\vec{a}^{\,\mathrm{new}} = \vec{a}^{\,\mathrm{cur}} + \delta\vec{a}$ is then accepted with probability $\min\{1, T_{kl}/T_{lk}\}$, for which $T_{lk} = \mathcal{T}_{lk}/\sum_l \mathcal{T}_{lk}$; in either case, $\mathcal{T}_{lk}$ or $\mathcal{T}_{kl}$ is incremented by unity. This has also been modified by fixing the estimate of $\mathbf{T}$ during a given iteration as in the multicanonical procedure and then updating this estimate before the next iteration. Since the acceptance criterion is then unchanged within each multicanonical iteration, such a procedure more accurately implements the detailed balance condition. That is, if f_k denotes the stationary probability of finding the system in the k:th histogram bin, then $T_{lk}\, f_k = T_{kl}\, f_l$ for each k and l, which follows from the properties of the state distribution in an equilibrium condition [8]. A far simpler procedure, to be discussed in the next chapter, is to accept the transition from k to l only if the state k as been previously visited at least as many times as the state l. In all cases, however, detailed balance should be imposed at the end of the calculation to avoid unphysical complex eigenvalues of the transition matrix. We have accordingly developed an iterative procedure, based on the observation that since $\mathbf{T}$ is normalized according to $\sum_l T_{lk} = 1$, detailed balance requires that $\sum_l T_{kl}\, f_l = f_k$, i.e., f_k is an eigenvector of the transition matrix $\mathbf{T}$ with unit eigenvalue. Hence, if we compute $f_k^{(0)}$ from the transition matrix $\mathbf{T}^{(0)} = \mathbf{T}$, a recursive series of approximations for $\mathbf{T}$ follow from the associated symmetry condition $T_{lk}^{(i+1)} = [T_{kl}^{(i)}\, f_l^{(i)} \;+\; T_{lk}^{(i)}\, f_k^{(i)}] \;/\; \left(2\, f_k^{(i)}\right)$.

To extend the above procedure to time-dependent problems, the magnitude of the random perturbation $\delta\vec{a}$ that yields an average change in $|\vec{E}^{\,\mathrm{new}} - \vec{E}^{\,\mathrm{cur}}|$ equal to the observed change over a time interval Δt is first determined. This can be accomplished, for example, by evaluating the average number of random perturbations required for the emulator to reach a state $\vec{E}$ that is significantly decorrelated from the initial state $\vec{E}_0$ and assigning to this value the experimentally determined drift time of the installed fiber [61]. In this

analysis, we assume that the underlying random process is Markovian, since memory effects would otherwise have to be incorporated into the transition matrix formalism. Retaining knowledge of the system evolution through previous states would then considerably expand the effective transition matrix dimensionality.

In one set of procedures for modeling time evolution, multiple series of biased transitions (paths) are generated by randomly selecting values of the system observables either according to the probabilities stored in the biased transition matrix or by applying e.g., the multicanonical acceptance rule together with the unbiased transition matrix probabilities. The resulting random walk avoids numerous direct calculations of the observables from the system variables, significantly reducing the simulation time. A path corresponding to a transition over a time $n\Delta t$ from the state k to the state l over the intermediate states $1, 2, \ldots, n$ is then assigned the weight $w = R_{1k}R_{21}\cdots R_{ln}$, where $R_{lk} = T_{lk}/B_{lk}$ designates the ratio of the unbiased to biased transition probabilities, and added to an appropriate histogram element.

While the transition matrix corresponds to a certain time step Δt, an approximate transition matrix for any time step τ is obtained by writing the time evolution of the pdf as $f(\vec{E}; t + \Delta t) = \mathbf{T} f(\vec{E}; t)$, so that $\mathrm{d}f(\vec{E}; t)/\mathrm{d}t \approx [(\mathbf{T} - \mathbf{I})/\Delta t]\, f(\vec{E}; t) = K f(\vec{E}; t)$, and $f(\vec{E}; t + \tau) \approx \mathrm{e}^{\tau K} f(\vec{E}; t)$ (a more exact procedure would require the logarithm of $\mathbf{T}$). For integer multiples of the time step $n\Delta t$, however, $f(\vec{E}; t + n\Delta t) = \mathbf{T}^n f(\vec{E}; t)$, which can be determined in as few as $\log_2(n)$ matrix multiplications. Of course, as $n \to \infty$, $f(\vec{E}; t+n\Delta t)$ tends towards the limiting stationary distribution, $f(\vec{E})$, with transient corrections that decay with time constants proportional to the eigenvalues of the transition matrix, cf. Fig 5.2.

5.4 Outage times

To illustrate the computation of global system properties with the above procedures, we examine the pdf of outage time durations induced by PMD in a single-mode optical fiber. A single-mode fiber supports two nearly degenerate orthogonally polarized modes that exit after slightly different propagation times $\Delta\tau$ as a result of birefringence induced by asymmetric perturbations such as core ellipticity or internal or external stress. The relative orientation of the birefringence axes of adjacent segments changes, however, along the fiber length and with time. Therefore, over a long propagation distance, the pulse width and shape fluctuate randomly. To lowest order, however, the pulse can be modeled as a superposition of two orthogonally polarized pulses with a group delay difference that is termed the DGD. When the ratio of the DGD to the average DGD of the fiber is three or greater, the bit-error-rate of the overall optical system typically exceeds 10^{-9}, which is the value required for acceptable line behaviour. The time intervals during which the DGD exceeds a certain value are termed outage times. Although in this chapter we consider first order effects, pulse distortion due to chromatic dispersion and higher-order PMD [106] can be incorporated into our models by instead constructing the matrix composed of the relative probabilities of transitions between states with differing system penalties [74].

Our fiber model consists of a series of $N = 100$ birefringent fiber segments, with an overall mean DGD of $\tau_{mean} = 20$ ps, separated by polarization rotators that can generate arbitrary rotations of the incoming polarization vector $\vec{\Omega}_i$ on the Poincaré sphere. The vector of system variables $\vec{a}$ is, therefore, composed of the $2N$-dimensional relative angles between the birefringence vectors of successive sections, while the system observables $\vec{E}$ are described by a single scalar DGD value. The outage time is normally determined by performing many time steps $\vec{a}^{\,\mathrm{new}} = \vec{a}^{\,\mathrm{cur}} + \delta\vec{a}$, and recording the number of steps for which the system remains in the outage region.

The Markov chain can also be obtained by sampling values of the system observables

according to the transition probabilities stored in the unbiased or biased transition matrices and then incrementing the histogram element for the time duration of each system path in the outage region by either 1 or w of Section 5.3, respectively. More efficiently, an initial state distribution given by $\tilde{f}_k$ for k in the non-outage region and zero elsewhere can be repeatedly multiplied by the transition matrix. After the n:th multiplication, the sum of the values of $\tilde{f}_k$ for all $\vec{E}_k$ that fall outside the outage region corresponds to the likelihood that the system returns to a non-outage state after an approximate time $n\Delta t$. These elements of $\tilde{f}_k$ are then equated to zero before the subsequent time step. Alternatively, since the paths between the initial and final state are repeatedly multiplied by the submatrix formed from the transition matrix entries for states in the outage region, the eigenvalues of this submatrix yield directly the decay rate of the probability density of the system outage duration. Further, the outage time histogram can be obtained by projecting $\tilde{f}$ onto the eigenvectors of the submatrix and summing the resulting distributions.

5.5 Numerical results

In Fig. 5.1, we show the probability that the system experiences an outage of a given duration, where the outage condition is defined as a fiber DGD greater than $2\tau_{mean}$ or $3\tau_{mean}$ (in the second case fewer than 10^4 unbiased samples were recorded in the outage region for 10^9 realizations). The circles in the figure indicate the results of an unbiased Markov chain calculation that recorded the outage times observed during 10^9 time steps while the solid line was generated by first constructing the unbiased transition matrix from both the accepted and rejected transitions during three 5×10^6 sample multicanonical iterations and then applying the multicanonical acceptance rule to this transition matrix. We have found that employing the acceptance rule based on the ratio of transition matrix elements yields similar accuracy in regions of adequate statistics but does not sample as far into the tail of the pdf for a fixed number of realizations. Repeatedly multiplying the

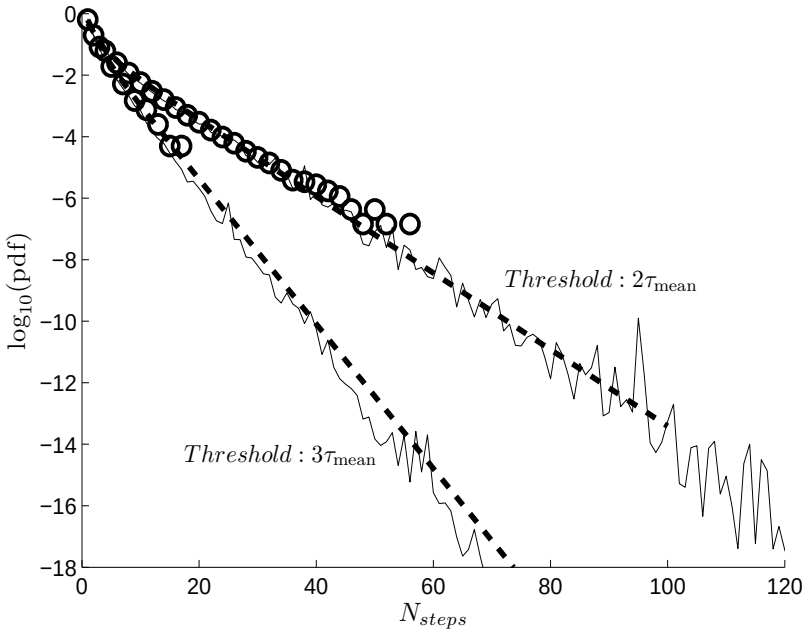


Figure 5.1: The pdf of the outage durations of an optical fiber emulator for thresholds of $2\tau_{mean}$ and $3\tau_{mean}$. Circles: standard method with 10^9 samples. Solid line: multicanonical reweighting with 3 iterations of 5×10^6 samples. Dashed line: repeated multiplication of the initial state vector by the unbiased transition matrix obtained after two 2×10^5 sample multicanonical iterations.

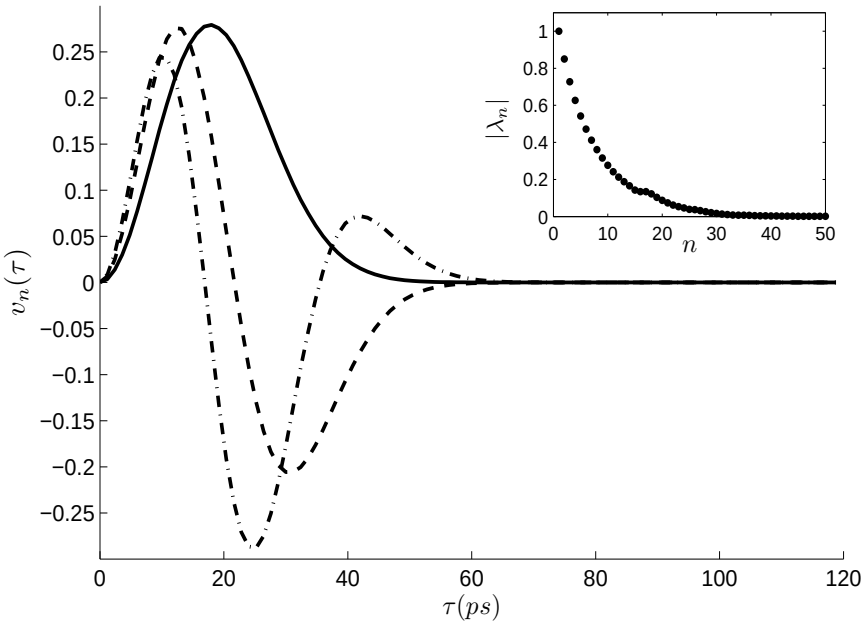


Figure 5.2: The first (solid line), second (dashed line) and third (dashed-dotted line) eigenvectors of the transition matrix $\mathbf{T}$ after three multicanonical iterations of 5×10^6 samples. Inset: the magnitudes of the first 50 eigenvalues.

initial state vector by the unbiased transition matrix obtained after two 2×10^5 sample multicanonical iterations instead yielded the dashed line of Fig. 5.1. Fig. 5.2. contains the three eigenvectors with largest eigenvalues of the transition matrix obtained from three 5×10^6 sample multicanonical iterations together with the magnitude of the highest 50 eigenvalues (inset). The first eigenvector (solid line) with unit eigenvalue corresponds to the asymptotically converged pdf of the DGD while the other eigenvectors have eigenvalues less than one and therefore influence the evolution of the pdf of the DGD at moderate time intervals.

5.6 Conclusions

We have established that the time-dependence of communication systems as well as of general physical systems can be rapidly and accurately modeled, even for very unlikely configurations by transition matrix methods. Of course some accuracy, which is model-dependent, is inevitably sacrificed since in the standard method the system variables vary continuously from one realization to the next, preserving higher-order correlations between the system observables and the underlying system variables.

Chapter 6

A Comparison of Transition Matrix Sampling Procedures

In this chapter we compare the accuracy of the multicanonical procedure with that of transition-matrix models of static and dynamic communication system properties incorporating different acceptance rules. For appropriate ranges of the underlying numerical parameters, we find that algorithmically simple yet highly accurate procedures can be employed in place of the standard multicanonical sampling algorithm.

6.1 Introduction

Multicanonical methods [8] have found numerous applications to optical systems since they were first adapted to communications theory in [76, 122, 123]. Although only static quantities, such as the probability distribution function (pdf), $f(\vec{E})$, that a system is in a configuration characterized by values, $\vec{E}$, of its observables can be calculated with multicanonical techniques, this restriction was removed by a suitable implementation of the

transition matrix method [34, 115]. The analysis employed different techniques for generating statistical samples. The first of these is based upon a modified multicanonical acceptance rule that retains the accuracy of the standard multicanonical method but converges faster and is simpler to program. A second procedure uniformly samples different regions of the pdf but is generally less accurate. Finally, this procedure was modified to increase the algorithmic precision without affecting the sampling probabilities. This chapter examines in greater detail the relative advantages of these three procedure with emphasis on discretization errors associated with the histogram bin widths.

6.2 Transition matrix methods

We have demonstrated in the preceding chapter that the multicanonical procedure can be adapted to dynamic system evolution if the frequency of all *accepted and rejected* transitions is retained [115]. In particular, specializing to a single observable so that $\vec{E}$ is replaced by E, the elements, $\mathcal{T}_{lk}$, of an unnormalized transition matrix, $\boldsymbol{\mathcal{T}}$ (and additionally in certain methods a histogram, V, of visited states) are initialized to zero while the initial $\vec{a}^{\,\mathrm{cur}}$ are randomly generated. For each accepted or rejected transition from an initial k:th histogram bin to a final, l:th, bin corresponding to the observables $E(\vec{a}^{\,\mathrm{cur}})$ and $E(\vec{a}^{\,\mathrm{new}})$, respectively, $\mathcal{T}_{lk}$, and where relevant, V_k, are incremented by unity. Following an accepted transition, the state $\vec{a}^{\,\mathrm{new}}$ in bin l replaces $\vec{a}^{\,\mathrm{cur}}$ as the initial state for the subsequent time step. The normalized transition matrix $\mathbf{T}$ is generated by scaling each column of $\boldsymbol{\mathcal{T}}$ such that $\sum_l T_{lk} = 1$ since the probabilities of all transitions from a state k must sum to unity.

From the above discussion, we observe that transition matrix methods are distinguished by the choice of acceptance rule. Although since all unbiased transitions out of a state are recorded any transition rule is permissible, as demonstrated below if the change in E over a Markov step is comparable to the histogram bin widths, the discretization error is

dependent on the rule selected. In the previous chapter, two procedures were considered. Method (1) employed the standard multicanonical acceptance rule to populate the transition matrix while method (2) accepted a transition only if the final state was previously visited fewer times than the initial state, i.e. $V_l < V_k$. In this chapter, we further refine method 1 by updating the estimate of $f(E)$ after every small number, N_u, of steps. Here the detailed balance condition between each pair of adjacent histogram bins is rewritten in the form of a recursion relation for the pdf, [41, 105]

$$f_{m+1} = f_m \, \frac{T_{m+1,m}}{T_{m,m+1}}. \tag{6.1}$$

Thus, starting from an arbitrary initial value for the probability of the first histogram bin and assuming e.g. that f_m is independent of m if either $T_{m+1,m} = 0$ or $T_{m,m+1} = 0$, we obtain an estimate of $f(E)$ that can be employed in the multicanonical acceptance rule. Since this calculation requires negligible computation time, f_m can be regenerated after any desired number of Markov steps (although preferably the procedure should be initialized with a Monte Carlo calculation). Dynamic system evolution can be modeled by repeatedly multiplying an initial state distribution by the transition matrix [127]. Additionally, multiplication by the matrix $\mathbf{T}$ is analogous to evolving the system through a simulated time interval, Δt, so that the eigenvector of $\mathbf{T}$ with unit eigenvalue corresponds to the probability distribution function, f_m [128].

In general, if the width of the m:th histogram bin is comparable to or greater than the average change in E over a single Markov step into or within this bin, the accuracy of f_m is affected by the Markov chain dynamics within the bin (in fact the calculations below suggest that this is the dominant source of error under the given conditions). That is, ideally, the probability of visiting states within a bin should follow the physical distribution that is obtained in a Markov chain calculation in the absence of an acceptance rule. Otherwise states are spuriously depleted toward one side of the histogram bin, which significantly alters the population of the states within the bin for large bin sizes.

However, if the acceptance rule is formulated to preserve detailed balance, as in the case of the multicanonical method, the depletion of states out of one side of a histogram bin is compensated by an equal number of incoming transitions from the neighbouring bin. This prevents states in the Markov chain from on average entering bins preferentially from e.g. lower values of E and exiting towards higher values, which affects the relationship, Eq. (6.1), between the pdf and the transition matrix elements. Therefore, for procedures that preserve detailed balance, the discretization error associated with the finite bin size is considerably reduced, as verified numerically below. The magnitude of the reduction, however, depends on a potentially large number of computational parameters.

Considering next method 2, which accepts transitions only to less visited states, suppose that at a certain stage of the calculation the acceptance rule only permits transitions from the current bin to bins with a smaller pdf. Then within a bin, a larger than average fraction of states with high pdf will result from incoming transitions (note that for a large bin width with high probability one or more Markov steps are required for the state to transfer from the larger to the smaller probability region of the bin). Therefore, the ratio between the frequency of transitions out of the bin to large pdf states and the transition frequency to small pdf states is smaller than in the standard Monte-Carlo procedure. However approximately half of the time, the acceptance rule will instead only permit transitions from the current bin to higher instead of lower pdf states. In this case, however, the above bias is not fully compensated since most Monte-Carlo transitions occur in any case to higher pdf states. Thus averaging over both possibilities for the acceptance rule, we conclude that the effective transition probability will on average be enhanced in the direction of large pdf values.

We have indeed observed in a number of different computational contexts involving a single observable E that the states in the Markov chain on average evolve preferentially at a constant velocity from one region of low $f(E)$ to the opposing region after which the simulation sometimes enters into a previously unsampled state with low $f(E)$ (in

the unusual case of several isolated large $f(E)$ regions, this behaviour will still occur modified by infrequent transitions between different high $f(E)$ regions). The Markov chain then remains in the new bin until the number of samples in the bin equals that of the adjacent histogram bin within the higher $f(E)$ region. The Markov chain then re-traverses the problem domain toward the original starting point. This ensures that on average the acceptance rule excludes transitions to higher pdf states as many times as transitions to lower pdf states. Consequently, method 2 yields a spurious bias that augments the computed slope of $f(E)$. To restore the correct transition probabilities, in method 3, following a transition into a given bin we discard all Markov steps associated with transitions out of a histogram bin until a certain number of confined steps, N_c, have been executed. The steady-state statistical distribution of states is then restored within the bin. The value of N_c required to ensure a given level of accuracy, however, can in general only be determined empirically since the average number of steps that the Markov chain spends in a histogram bin before exiting depends in a complicated fashion on the bin number and the computational and physical details of the problem.

6.3 Numerical results

We now quantify the accuracy of the three procedures by evaluating the pdf of the differential group delay (DGD), τ, in a polarization mode dispersion (PMD) emulator composed of $N = 10$ polarization maintaining (PM) fiber sections with $\tau_s = 1.0$ ps, separated by randomizing polarization controllers. This yields an average DGD of $\tau_{mean} = \tau_s\sqrt{8N/3\pi} = 2.91$ ps for the emulator. The histogram is then formed by dividing the interval $[0, 3.5]$ of normalized DGD values, τ/τ_{mean}, into 100 equal width segments. Thus the system variable E is identified with the DGD τ, while the bin designated by the index m corresponds to the range $3.5(m-1)/100 < E < 3.5m/100$.

The main graph (left axis) of Fig. 6.1 displays the ratio $f_m^{numerical}/f_m^{analytic}$ for a standard multicanonical calculation after three 1.67×10^6 step iterations ($\triangle$ markers), the transition matrix technique with a multicanonical acceptance rule (method 1, $\circ$ markers), a transition method procedure that rejects transitions to more sampled histogram bins (method 2, dashed-dotted line), and method 3 with N_c, the number of steps in which the Markov chain is confined to a bin before transitions out of the bin are permitted, equal to 40 ($+$ markers). By repeating our calculation with different values of N_c our method 3 curves are found to be nearly indistinguishable for $N_c > 20$; in general, the minimal value for a given computation and desired accuracy can only be determined empirically in this fashion. The right axis of the figure displays the analytic pdf of the DGD of the fiber emulator [59]. Here we have employed 5×10^7 emulator realizations in which the relative angle between each pair of emulator segments is randomly varied by an average of $\pi/80$ degrees between successive realizations. Further, in method 1 the pdf is updated after every step according to the current estimate from Eq. (6.1) of the transition matrix. As discussed in Section 6.2, method 2 predicts a reduced slope and therefore a ratio > 1 near the pdf maximum and < 1 for DGD values occurring with low probability. Further, although method 1 yields an accuracy comparable to the standard multicanonical method, we have found that it exhibits improved convergence with fewer samples and reduced programming complexity.

Considering next in Fig. 6.2, the total number of times, V_k, that a state in bin k is visited for the standard multicanonical procedure ($\triangle$ markers), method 1 ($\circ$ markers) and method 2 (dashed-dotted line) we observe that while the numerical error of method 2 is large, this procedure samples the pdf most evenly. Method 3 samples the pdf almost identically to method 2 and is therefore omitted in order to increase the legibility of the graph. Further, the sum over many iterations of the number of samples in each bin for the standard multicanonical procedure is not uniform as a result of variations in the sampling frequency in small pdf regions during each iteration. However, the multicanonical procedure and its variants concentrate samples in regions of very low pdf and therefore estimate the pdf

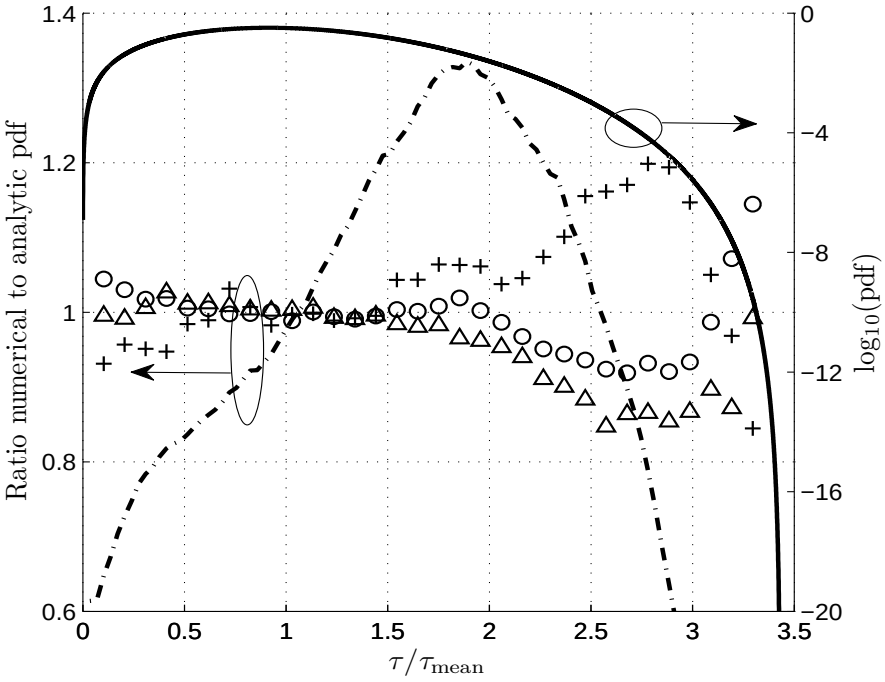


Figure 6.1: The ratio between the numerical and analytic pdfs for the standard multicanonical procedure ($\triangle$), the modified transition matrix procedure with a multicanonical acceptance rule (method 1, $\circ$), an acceptance rule that rejects transitions to more visited histogram bins (method 2, dashed-dotted line) and a procedure that restricts transitions out of a recently visited bin (method 3, $+$) as functions of the normalized DGD for a $N_{\text{sec}} = 10$ segment fiber emulator. The analytic result for the pdf is displayed against the right axis.

rapidly in these regions. Method 1 further improves on this feature since the intermediate pdf estimates are frequently revised.

Finally, in Fig. 6.3, we display the error averaged over all histogram bins according to [125]

$$\mathrm{E}\left\{\sum_{m=1}^{N_B}\ \left|\log_{10}\left(\frac{f_m^{\,numerical}}{f_m^{\,analytic}}\right)\right|\ f_m^{\,analytic}\right\}, \tag{6.2}$$

in which $\mathrm{E}\{\ldots\}$ denotes an ensemble average over 100, 2×10^6-sample calculations as a function of the width of the uniform distribution of the change, $\delta\vec{a}$, in the relative angles of two adjacent emulator segments. The results of the multicanonical procedure and methods 1-3 are denoted by $\triangle$, $\circ$, a dashed-dotted line and $+$, respectively. The accuracy of all methods greatly increases when the mean variation of the system observable over a single step in the Markov chain is large compared to the size of a histogram bin. Further, the improvement afforded by method 3 over method 2 is clearly visible in the figure, where for the smallest step sizes shown, the precision of method 3 approaches that of the multicanonical procedure.

6.4 Conclusions

Transition matrix procedures, which extend multicanonical techniques to dynamic problems, possess numerous implementations that differ considerably in accuracy and sensitivity to variations in numerical parameters. Here, we achieved a high degree of efficiency and accuracy by employing the multicanonical acceptance rule while continually updating the pdf estimates. We then interpreted the apparent violation of detailed balance and hence of numerical accuracy at large histogram bin sizes in terms of the change in probabilities of transitions to the two neighbouring bins compared to the values expected for an unbiased calculation. Acceptance rules that increase the probability of asymmetric transitions out of a bin generally enhance this source of error. Procedures of this nature can however be

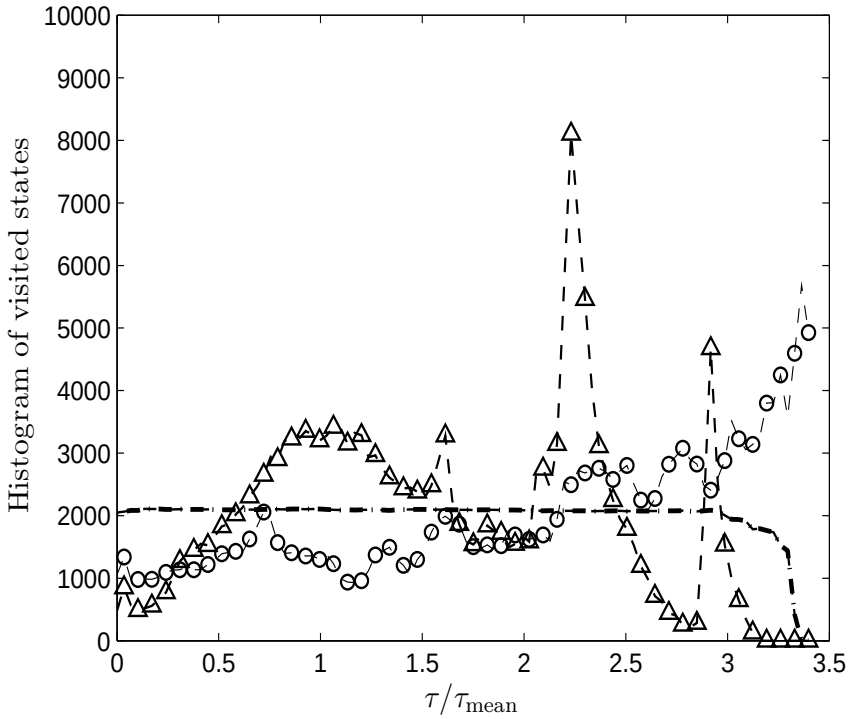


Figure 6.2: The total number of times each histogram bin is visited for the standard multicanonical procedure (triangles), method 1 (circles), and method 2 (dashed-dotted line).

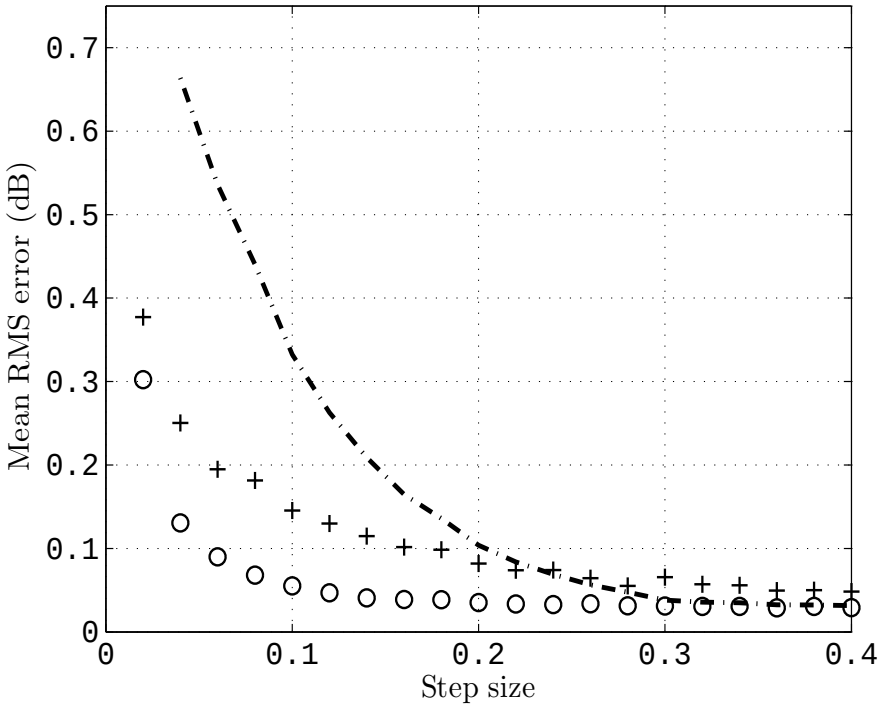


Figure 6.3: The variation of the error, Eq. (6.2), weighted by the histogram bin probability as a function of the average DGD change over one Markov step for the standard multicanonical method ($\triangle$), method 1 ($\circ$), method 2 (dashed-dotted line) and method 3 ($+$).

simple to program and computationally efficient if the numerical parameters ensure a small self-transition probability within each histogram bin. Further pursuing this line of reasoning could yield even more efficient transition matrix methods with possible experimental relevance [75, 76].

Chapter 7

Transition Matrix Analysis of the Hinge Model

In this chapter we analyze within the hinge model the time-dependence of polarization mode dispersion (PMD) induced by stochastic birefringence fluctuations. This procedure employs a modified transition matrix method approach that increases modeling accuracy for unlikely system configurations.

7.1 Introduction

The hinge model of PMD temporal evolution [13,15] provides a framework for calculations of the outage probability of fiber optic communication systems [4, 65]. Often, however, time-dependent quantities such as outage duration statistics must be calculated. In the preceding chapters, we analyzed the outages of optical channels characterized by a highly mode-coupled emulator model with a transition matrix procedure [127]. Here we present a simplified version of this method to classify the average probability of exceeding the outage

threshold within a specified time interval in the hinge model for stochastic, memoryless birefringence fluctuations.

7.2 Hinge model

The hinge model represents a fiber system as a series of N_S fiber spans joined by polarization scrambling regions with negligible differential group delay (DGD). Static and dynamic statistical quantities such as the pdf of the DGD or the average time spent in outage states can then be determined by averages over the polarization rotation angles. A "hinge model realization" is generated by associating the span DGDs with the absolute value of each of a set of N_S Gaussian distributed random values with zero norm and a standard deviation of σ_G (the results are effectively independent of the particular form of this distribution function). Each scrambler is modeled by three successive rotations drawn from a uniform random distribution on $[0, 2\pi]$ about the x, y and z axes and the DGD is determined with a standard recursive procedure [47]. Since the time variation of the local birefringence along a fiber link has not been adequately characterized experimentally and is in any case presumably specific to a fiber link, in analogy to statistical mechanics, here we consider the most random time variation. In particular, to approximate system evolution through a "time step", a random value in the interval $[-\delta, \delta]$ is added to each rotation angle. While these results are effectively independent of the form of this random distribution, they are of course affected if the birefringence values at successive time steps are significantly correlated.

Since in the hinge model, the span DGD values are assumed nearly constant over the observation interval, transitions between two hinge model realizations are precluded. The relative probability of transitions between any two states of the system separated by a single time step must accordingly be independently evaluated for each hinge realization.

Since such an approach precludes a statistical analysis, we instead introduce a "hinge metric" which we subsequently associate with the sum, here called the "link DGD", τ_m, of the magnitudes of the N_S span DGD values. A hinge model realization is then assigned to the m:th of N_M histogram bins if $(m-1)\Gamma_M/N_M < \tau_m < m\Gamma_M/N_M$ where m is referred to as the link index and Γ_M is the largest link DGD considered in the calculation. Although the statistics of the different instances of the hinge model in each such interval may vary significantly from the averages (presented below) over all systems in the interval, our methodology will clearly illustrate the expected hinge-model behaviour.

To extend the above procedure to more general time-dependent problems, the magnitude of the random perturbation, δ, is first determined by, for example, evaluating the average number of random perturbations required for the emulator to reach a DGD state that is significantly decorrelated from the initial state and assigning to this value the experimentally determined drift time of the installed fiber [61]. This analysis assumes that the underlying random process is Markovian, since memory effects would otherwise generally have to be incorporated into the transition matrix formalism. Retaining knowledge of the system evolution through previous states would then considerably expand the effective transition matrix dimensionality. Further, good agreement has been observed elsewhere between the measured outage statistics on a buried, installed fiber link and numerical simulation [129], suggesting a Markovian mode-coupling model is valid for buried fiber and, should be even more applicable to the rapid time evolution of aerial fiber.

7.3 Modified transition matrix method

The transition matrix procedure requires a description of the physical system by a Markov chain formed by the instantaneous DGD, $\Omega(\vec{a})$, of a hinge model realization with

link index m that in turn depends on a vector formed by the values of N_a stochastically varying rotation angles, or "system variables", $\vec{a}$. An initial random set of system variables $\vec{a}^{\,\mathrm{cur}}$ is generated and the DGD $\Omega(\vec{a}^{\,\mathrm{cur}})$ computed. Next, the system variables are randomly perturbed according to $\vec{a}^{\,\mathrm{new}} = \vec{a}^{\,\mathrm{cur}} + \delta\vec{a}$, where $\delta\vec{a}$ represents N_a small random perturbations drawn uniformly from the interval $[-\delta, \delta]$ yielding the updated DGD $\Omega^{\,\mathrm{new}}$. The resulting range of allowed DGD values, Γ_K, is subdivided into N_K histogram bins.

Since standard Monte-Carlo and multicanonical methods [122] do not retain the relative frequency of transitions between each pair of states $\Omega^{\,\mathrm{cur}}$ and $\Omega^{\,\mathrm{new}}$ in the Markov chain previous chapters have considered an alternative technique, adapted from Refs. [105, 115], that accumulates in an array of matrices $[\mathcal{T}_m]_{lk}$ the frequency of all *accepted and rejected* transitions from an initial DGD value, $\Omega^{\,\mathrm{cur}}$, within the k:th histogram bin, $(k-1)\Gamma_K/N_K < \Omega^{\,\mathrm{cur}} < k\Gamma_K/N_K$, to a final value in the l:th bin, $(l-1)\Gamma_K/N_K < \Omega^{\,\mathrm{new}} < l\Gamma_K/N_K$, in one time step. This transition is accepted (the "modified transition rule") if $[V_m]_l \leq [V_m]_k$ where V_m is a histogram whose k:th element is incremented after each transition. In this case, the l:th state is employed as the starting point for the next time step. If $k = l$, the new state should be chosen as the updated starting point to improve accuracy for small time steps. The resulting modified transition rule ensures uniform statistical sampling across all columns k of the matrix $\boldsymbol{\mathcal{T}}_m$. For small mean time steps, typically N_K must be large to avoid an artificial enhancement of transitions that join states near the boundaries of adjacent histogram bins. Our procedure can then be implemented as follows:

1. Obtain the first of N hinge model realizations by assigning the absolute value of Gaussian distributed random values to the DGD of each of the fiber spans and compute the link DGD, τ. If τ falls in the m:th histogram bin, $(m-1)\Gamma_M/N_M < \tau < m\Gamma_M/N_M$, the subsequent steps populate the m:th matrix of an array of N_m transition matrices, $\boldsymbol{\mathcal{T}}_m$.

2. Select the random coupling angles $\vec{a}^{\,\mathrm{cur}}$ between each pair of spans, compute the DGD

$\Omega^{\rm cur}$ of the initial emulator state and determine the associated DGD histogram bin (here k).

3. For each hinge model realization, perform the first of N_T time steps by setting $\vec{a}^{\,\rm new} = \vec{a}^{\,\rm cur} + \delta\vec{a}$ such that if the time step generates a transition from k:th to l:th DGD histogram bin the transition matrix element $[\mathcal{T}_m]_{lk}$ is incremented by unity.

4. This transition is accepted if $[V_m]_l \leq [V_m]_k$ in which case $[V_m]_l$ is incremented by one and the l:th state is employed as the starting point for the next time step, with $[V_m]_k$ incremented otherwise. If $k = l$, the new state should be chosen as the updated starting point to improve accuracy for small time steps.

5. After N_T repetitions of Steps (3)-(4), the subsequent hinge model realization is generated and steps (1)-(4) are repeated N times.

To transform each $\boldsymbol{\mathcal{T}}_m$ into the normalized transition matrix $\mathbf{T}_m$ each column of $\boldsymbol{\mathcal{T}}_m$ is individually normalized to ensure $\sum_l [T_m]_{lk} = 1$ since the probabilities of all transitions from a state k must sum to unity. Accordingly, if a system with hinge metric index m initially possesses a DGD Ω within the p:th histogram bin, i.e. $(p-1)\Gamma_K/N_K < \Omega < p\Gamma_K/N_K$, it can be represented as the unit "system" column vector of size $N_K \times 1$ whose p:th element is identically one, i.e. $\vec{s}^{\,(0)} = [0, \ldots, (s_p = 1), 0 \ldots, 0]^{\rm T}$. Multiplication by the transition matrix $\mathbf{T}_m$ then yields the system vector $\vec{s}^{\,(1)} = \mathbf{T}_m \vec{s}^{\,(0)}$ quantifying the state occupation probability after a single time step. Further, if $[f_m]_k = f_m(\Omega_k)$ denotes the stationary probability of finding the hinge model realization with link index m in the k^{th} DGD histogram bin, then $[T_m]_{lk}\,[f_m]_k = [T_m]_{kl}\,[f_m]_l$ for each k and l, which follows from the properties of the state distribution in an equilibrium condition [8]. Therefore, since $\sum_l [T_m]_{lk} = 1$, we have $[f_m]_k = \sum_l [T_m]_{kl}\,[f_m]_l$ implying that $f_m(\Omega)$ is an eigenvector of the transition matrix, $\mathbf{T}_m$, with unit eigenvalue [127].

It is now straightforward to determine the outage statistics under the simplified but illustrative assumption that a system outage occurs when Ω exceeds a certain threshold value (employing instead, for example, the power penalty [27] would yield more precise results but with the same general features). If the outage boundary is located between the j and $j-1$:st DGD histogram bins, the probability of a system outage after a single time step for an initial state within bin p is then given by $P^{(1)} = \sum_{k>j} s_k^{(1)}$ where $s_k^{(1)}$ is the k:th element of $\vec{s}^{(1)}$. Subsequently, we set $s_k^{(1)} = 0$ for all $k > j$ and employ this as the initial system vector, $\vec{s}^{(2)}$, for the subsequent iteration. Iterating this procedure yields the probability $P^{(i)}$ of the system remaining in an outage after i time steps.

7.4 Results

We now consider 4×10^3 hinge model realizations with $N_S = 10$ fiber spans generated from a Gaussian distribution with a standard deviation $\sigma_G = 2.0$ ps and evolve each realization $N_T = 10^4$, $\delta = \pi/5$ time steps. All DGD values are normalized to the same value, namely the mean DGD of a 10 section standard fiber emulator with 2.0 ps DGD in each section, namely $\sigma_E \equiv \sigma_G\sqrt{8N_S/3\pi}$, in order to remove the dependence on σ_G while facilitating comparisons between different curves. This calculation further employs $\Gamma_M = \Gamma_K = 7$, $N_M = 30$ and $N_K = 50$.

To illustrate the modified transition matrix formalism, Fig. 7.1 displays the pdf of the DGD for $\tau_m = 3\sigma_E$, computed as the eigenvector of $\mathbf{T}_m$ with unity eigenvalue in the modified transition matrix formalism (solid line) and with a Monte-Carlo calculation (+ markers). Clearly the transition matrix procedure greatly increases the accuracy of the pdf in the low probability, large DGD regions, as it distributes the number of samples nearly uniformly among the DGD bins. In Fig. 7.2, we depict the structure of the transition matrix, again for $\tau_m \approx 3$. The horizontal and vertical axes display the normalized DGD of

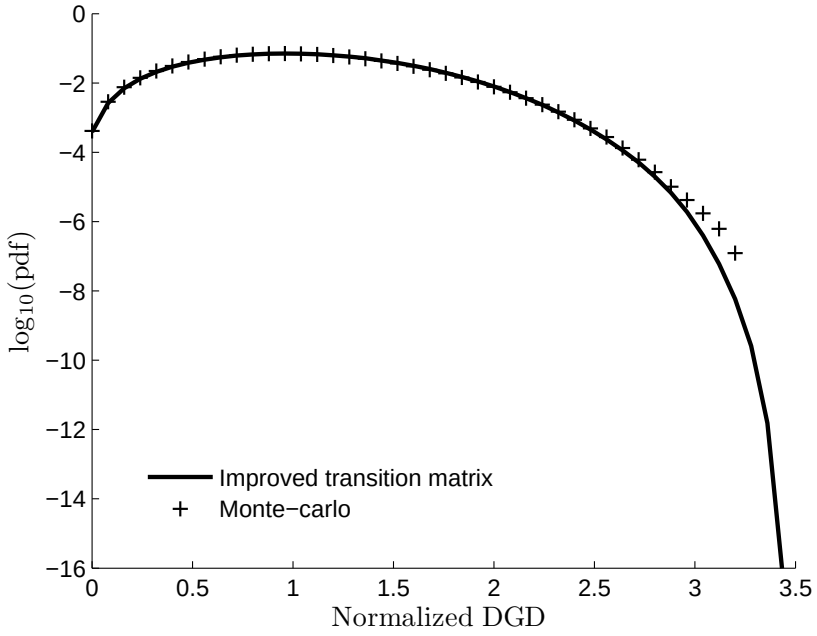


Figure 7.1: The eigenvector of the transition matrix with unity eigenvalue, i.e. the pdf of the DGD, for our new transition method procedure (solid line) and the corresponding Monte-Carlo calculation (+ markers) for a 10 section hinge-model emulator and $\tau_m = 3.0$.

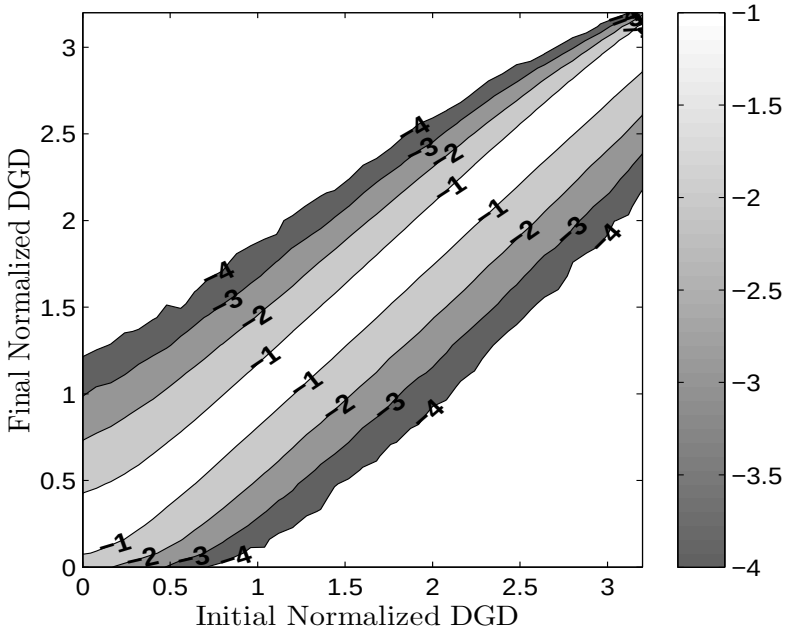


Figure 7.2: The transition matrix elements corresponding to Fig. 7.1. The x and y axes are labeled with the normalized DGD of the initial and final states, respectively, while the contours indicate the base 10 logarithm of the transition probability.

the initial state and final state while the contours indicate the logarithm of the probability of transition to a state. Since the largest DGD value attainable by the fiber link for a given hinge model realization is τ_m, the transition matrix elements only extend to this value. Although system configurations with small Ω evolve preferentially to higher DGD states, the DGD of the large DGD states changes less over a time step, presumably because of the lack of states with still greater DGD values. As noted in Ref. [129], however, the transition probability between two states is a function of the DGD difference and exhibits only a weak dependence on the DGD of the initial state.

We now employ the transition rule to compute the base 10 logarithm of the pdf averaged over all hinge models realizations with the same hinge metric, τ_m. These distributions are displayed as the vertical cross sections of the contour plot in Fig. 7.3. Evidently, the functional form of the pdf does not vary significantly with τ_m, although the width of the pdf does scale approximately with τ_m. These results also imply that the improvement afforded by fast mid-link polarization scrambling coupled with forward error correction coding is limited in systems with large hinge metric values resulting from a high transition probability between outage states. In contrast, systems with small hinge metric cannot evolve into the outage region, even if scrambling is absent.

Finally, we display the average probability that a system configuration with a given link and fiber DGD value will enter the outage region after 20 time steps, where an outage state is defined by $\Omega > 3.5\sigma_E$. A contour plot of the base 10 logarithm of this probability is presented in Fig. 7.4. As expected, if the hinge metric value (link DGD) is less than the outage value, the link is error-free with respect to DGD. Fig. 7.5 instead displays the base 10 logarithm of the number of steps after which a given state has developed a 50% probability of having evolved into the outage region. The number of required steps grows nearly logarithmically as the hinge metric decreases towards the outage value.

This analysis suggests that if the instantaneous DGD, Ω, can be continuously moni-

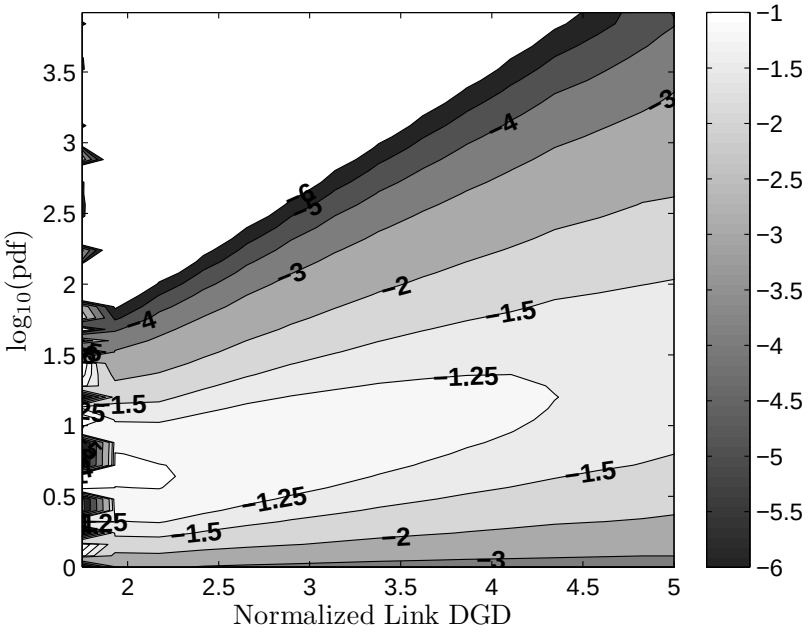


Figure 7.3: $\log_{10}(\text{pdf})$ for an ensemble of 10 section emulators, displayed as vertical cross sections, grouped according to the link DGD:s specified on the x-axis.

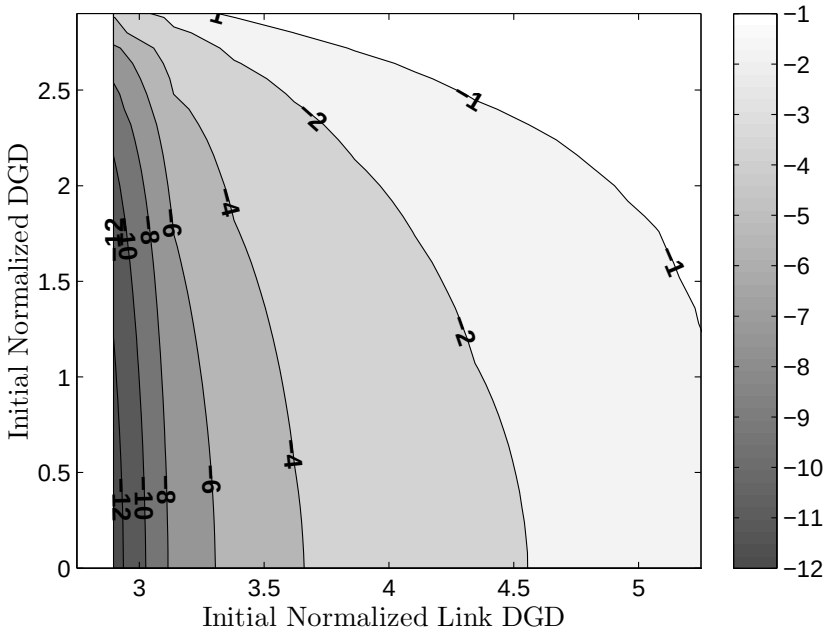


Figure 7.4: $\log_{10}$(pdf) from the modified transition matrix procedure that a system configuration with the indicated link and fiber (x and y axis) DGD values evolves after 20 time steps into an outage state for which the average DGD is 3.5 times the mean DGD of a standard emulator.

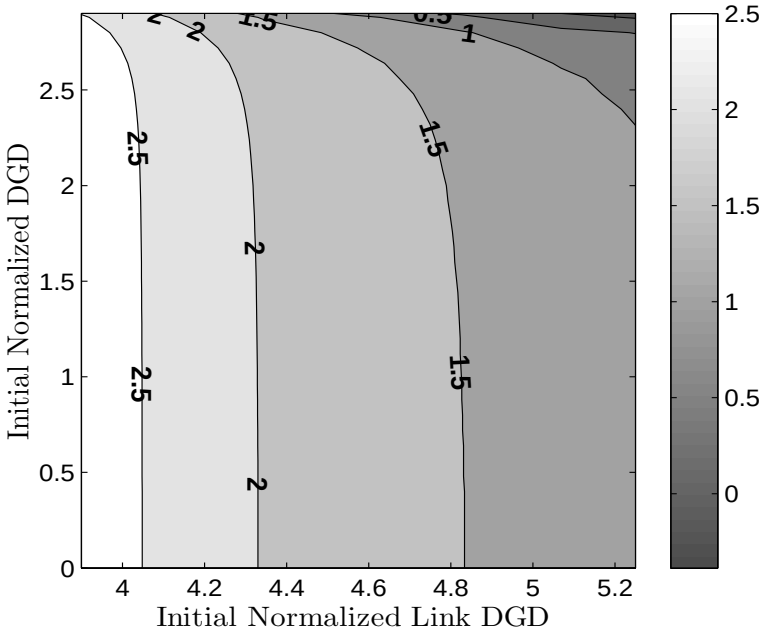


Figure 7.5: As in the previous figure, but for the base 10 logarithm of the average number of steps required until the probability of an outage exceeds 50%.

tored, a comparison of the experimentally measured pdf with graphs such as Fig. 7.3, or an analysis of the behaviour of the pdf for large Ω, could yield the likelihood that the system possesses an overall link DGD of τ_m, after which time-dependent statistical quantities such as those of Figs. 7.4-7.5 may then be simulated directly using the transition matrix procedure of Section 7.3. Conversely, a comparison of the measured outage statistics of an installed link in which the span DGD values have been measured with the results of Figs. 7.4-7.5 should indicate the degree of long-term correlations of the birefringence, especially in regions of system space with relatively few measurement events. Such a measurement, which could have significant system implications, would also either verify the Markovian approximation for the hinge angles for a given system or would provide alternative transition probability models with time-dependent behaviour that would reflect the underlying physical forces acting on the system.

7.5 Conclusions

This chapter has summarized a calculation of outage probabilities in the hinge model that, together with measurements of model parameters along the general lines of Ref. [129], could assist in reconfiguring certain fiber systems when the outage event probability is large. Conversely, if the overall DGD but not the individual hinge metric values can be continuously monitored, a comparison of the experimentally measured probability density function with graphs such as those in this paper might assist in determining the spatial and time-correlation properties of the birefringence fluctuations.

While the transition matrix formalism is both efficient and easily implemented, more complex formulations could, for example, confine sample generation to limited regions of system space or to histogram bins that have been sampled fewer than a specified number of times, as in the biased multicanonical method. The advantages afforded by such techniques,

however, are highly problem-dependent

This formalism can be immediately extended to quantities such as the average number of steps until an outage state is exited or the percentage of hinge/axis rotations that produce an increase in the DGD for states with a given DGD value. We have also analyzed other hinge model metrics such as the magnitude of the longest PMD vector, but these results appear to have limited practical interest.

Chapter 8

Modified Transition Matrix Simulations of Communication Systems

We apply appropriately enhanced transition matrix and multicanonical methods to communication systems. This procedure not only predicts time-independent quantities such as the bit-error-probability but can also be applied to dynamic effects such as the distribution of fading times.

8.1 Introduction

While importance sampling [132] is often employed to calculate communication system quantities such as the bit-error probability (BEP), the bias function can be difficult to determine. In contrast, the multicanonical method, which was first adapted to communication systems in [122, 123], automates this process while a closely related transition

matrix technique that additionally could be extended to dynamic system quantities yields improved computational efficiency with reduced algorithmic complexity [105, 115]. In this chapter we extend the results of Refs. [26, 69] by examining fading channels in wireless communication systems with both of the above procedures.

8.1.1 Modified transition matrix method

We recall from previous chapters that the multicanonical procedure requires a description of a physical system by a vector of N_E observables, $\vec{E}(\vec{a})$, that in turn depend on N_a stochastic parameters $\vec{a}$. To estimate a stochastic function $f(\vec{E})$ such as the probability density function (pdf) for statistically unlikely values of $\vec{E}$, the relevant region of $\vec{E}$ is first partitioned into N_B histogram bins centered at $\vec{E}_m$ with $m = 1, 2, \ldots, N_B$. All elements of two histograms, one for an estimate of the unnormalized pdf, $f_m^{(0)}$ and a second for storing the intermediate results of each iteration H_m are initialized to unity.

In the first iteration of the procedure, a random set of system variables $\vec{a}^{\,\mathrm{cur}}$ is selected. A small random perturbation selected from an effectively arbitrary distribution function [125] is then added, so that $\vec{a}^{\,\mathrm{new}} = \vec{a}^{\,\mathrm{cur}} + \delta\vec{a}$ and the associated observables $\vec{E}(\vec{a}^{\,\mathrm{new}})$ and $\vec{E}(\vec{a}^{\,\mathrm{cur}})$ with bin indices m_{new} and m_{cur} are computed. This perturbation or "transition" is accepted with a probability $\min\{1, f_{m_{cur}}^{(0)}/f_{m_{new}}^{(0)}\}$, in which case $\vec{a}^{\,\mathrm{cur}}$ is equated to $\vec{a}^{\,\mathrm{new}}$. Otherwise the subsequent step again employs $\vec{a}^{\,\mathrm{cur}}$. In both cases, however, the histogram entry $H_{m_{cur}}$ corresponding to the new $\vec{E}^{\,\mathrm{cur}}$ is increased by one. After M steps, the bias introduced through the acceptance rule is removed and an improved estimate, $f_m^{(1)}$, of $f(\vec{E})$ is generated according to $f_m^{(1)} = c\, f_m^{(0)}\, H_m$, where c is an appropriate normalization constant. The next iteration then proceeds similarly with H_m reset to unity, and $f_m^{(1)}$ replacing $f_m^{(0)}$. The transition rule in this and subsequent iterations increases the sampling probability of states with small $f(\vec{E})$ such that the likelihood of visiting configurations within a region $\left|\vec{E} - \vec{E}_0\right| \leq \delta\vec{E}$ becomes independent of $\vec{E}_0$ as the

number of samples and iterations increases.

Since the multicanonical method does not retain the relative frequency of transitions between each pair of states in the Markov chain we have previously considered an alternative technique, adapted from [105], that accumulates in a matrix $\boldsymbol{\mathcal{T}}$ with elements $\mathcal{T}_{lk}$ the frequency of all *accepted and rejected* transitions between the initial, k:th, histogram bin and the final, l:th, bin. The columns of $\boldsymbol{\mathcal{T}}$ are subsequently normalized according to $\sum_l \mathcal{T}_{lk} = 1$ since the probabilities of all transitions from a state k must sum to unity, yielding the transition matrix $\mathbf{T}$ with elements T_{lk}.

Since

$$f_{n+1} = f_n \frac{T_{n+1,n}}{T_{n,n+1}} \tag{8.1}$$

as a consequence of the detailed balance condition [41, 105], by setting $f_1 = 1$ as well as $f_m = f_{m-1}$ for bins in which $T_{j+1,j} = 0$ or $T_{j,j+1} = 0$, and subsequently normalizing $\sum_m f_m = 1$, we can rapidly construct an estimate of $f(\vec{E})$ that can be employed in the Markov acceptance rule. This intermediate pdf estimate can be regenerated after any desired number of Markov steps while the final improved result for $f(\vec{E})$, obtained after M Markov steps follows from $f_k = \sum_l T_{kl} f_l$ implying that $f(\vec{E})$ is an eigenvector of the transition matrix, $\mathbf{T}$, with unit eigenvalue [127].

8.1.2 Wireless fading channels

The above procedures can be applied to the method of exact Doppler spread (MEDS) model [88] for the time dynamics of the complex gain, $\mu(t)$, of a Rayleigh fading channel for which $\mu(t) \approx \hat{\mu}_1(t) + i\hat{\mu}_2(t)$ and, for the maximum Doppler frequency, $f_{\max}$, the number

of scatterers, N_a, and $a = 1, 2$,

$$\hat{\mu}_a(t) = \sqrt{\frac{2}{N_a}} \sum_{b=1}^{N_a} \cos\left(2\pi f_{ab} t + \theta_{ab}\right) \tag{8.2}$$

$$f_{ab} = f_{\max} \sin\left[\frac{\pi}{2N_a}\left(b - \frac{1}{2}\right)\right]. \tag{8.3}$$

The initial state $\vec{a}^{\,\mathrm{cur}} = [\theta_{11}, \ldots, \theta_{1N_1}, \theta_{21}, \ldots, \theta_{2N_2}]$ of the Markov chain incorporates a set of phases, θ_{ab}, that are uniformly distributed in $[0, 2\pi]$. The dynamic (time-dependent) channel quantities for the state $\vec{a}^{\,\mathrm{cur}}$ are subsequently evaluated according to Eq. (8.2). The phases are then changed slightly and the Markov acceptance rule is applied after the channel properties are recalculated.

We first determine the pdf, $f(E)$, with E identified with the number of fading events N_F occurring within a time interval $0 \leq t \leq T_F$, and, in a subsequent calculation, estimate the fade duration distribution of a mobile channel, which is typically incorrectly sampled by Markov procedures because of correlations between consecutive short fading events. In this case, for each $\vec{a}^{\,\mathrm{cur}}$ the system variables $\vec{E}^{\,\mathrm{cur}} = [\tau_1, \tau_2]$ are selected with τ_1 and τ_2 the calculated durations of two consecutive fading events. This yields a large, $N^2 \times N^2$, but sparse transition matrix. The conditional probability $f(\tau_2|\tau_1)$ is then determined from the calculated joint pdf $f(\tau_2, \tau_1)$ according to $f(\tau_2|\tau_1) = f(\tau_2, \tau_1)/\sum_{\tau_2} f(\tau_2, \tau_1)$. The fade duration distribution, $f(\tau)$, then corresponds to the eigenvector with unit eigenvalue of the $f(\tau_2|\tau_1)$ matrix since $f(\tau_2|\tau_1)f(\tau_1) = p(\tau_1|\tau_2)f(\tau_2)$ and $f(\tau_2) = \sum_{\tau_1} f(\tau_2|\tau_1)f(\tau_1)$. For sufficiently low threshold levels, however, correlations between consecutive short fading events can be neglected, and we instead apply the one-dimensional multicanonical and transition matrix procedures.

Next, we consider a standard model of fading channels [69] in which the signals broadcast by $L + 1$ users are detected by N_R independent reception branches. The desired user transmits a signal at a power level P_S while the L interfering transmitters transmit at the same power level P_I. Additionally, additive white Gaussian noise, $\vec{n}$, with time-averaged

intensity σ^2 is present at each receiver so that if s and s_l represent the modulated symbol of the desired user and the l:th interferer, $\vec{r} = \left[r_1, r_2, \ldots, r_N\right]^{\mathrm{T}}$ is a column vector representing the received signal vector, and $\vec{c}$, $\vec{c}_l$ are the complex channel gains for the desired user and the l:th user respectively, [69]

$$\vec{r} = \sqrt{P_s}\,\vec{c}\,s + \sqrt{P_I}\sum_{l=1}^{L}\vec{c}_l\,s_l + \vec{n}. \tag{8.4}$$

The channel gains are modeled by independent-identically-distributed (i.i.d.) symmetric complex Gaussian variables with variance $1/2$ in the real and imaginary components separately, yielding Rayleigh fading. The optimum weighting of the N_R signals r_l that maximizes the signal-to-interference ratio is then [69]

$$\vec{w} = \left[P_I\sum_{l=1}^{L}\vec{c}_l\,\vec{c}_l^{\,\dagger} + \sigma^2\,\mathbf{I}_N\right]^{-1}\vec{c}, \tag{8.5}$$

where $\mathbf{I}_N$ signifies the $N \times N$ identity matrix, $\dagger$ is the Hermitian transpose, and identify $D = \mathrm{Re}\{\,\vec{w}^{\,\dagger}\vec{r}\,\}$ with the decision variable to be consistent with Ref. [26], although other authors include an additional factor of 2 into D [69].

Next, we examine BPSK modulation for a source with equal 1 and -1 bit probabilities. In this case, since the BPSK constellation is symmetric $s = s_l = 1$, implying that a decision error occurs if $D < 0$. Then to implement the multicanonical or transition matrix formalism all stochastic quantities are varied in such a manner as to preserve the Gaussian distribution of the system variables (although other distributions are possible, cf. [125]). Thus we can vary v and u in the Box-Muller transformation for the real and imaginary parts of a Gaussian complex random variable according to

$$z = \sqrt{-2\ln(u+\epsilon_u)}\exp\left[i2\pi(v+\epsilon_v)\right], \tag{8.6}$$

in which ϵ_u and ϵ_v are selected from a uniform distribution on $[-\delta/2, \delta/2]$. While this transformation is periodic in v, $u + \epsilon_u$ must remain in the interval $[0, 1]$. To minimize the

impact of the boundary discontinuities, if $u + \epsilon_u$ takes the form $-\lambda$ or $1 + \lambda$ with $\lambda > 0$ after the perturbation, we respectively substitute λ or $1 - \lambda$.

Finally, we observe that the numerical technique of Section 8.1.1 may be applied to a filter based Rayleigh channel simulator [88] in which $\vec{a}^{\,\mathrm{cur}}$ is associated with a N_T dimensional Gaussian noise vector. A fading channel realization is obtained by filtering the elements of $\vec{a}^{\,\mathrm{cur}}$ as in Ref. [88]. The variables $\vec{a}^{\,\mathrm{cur}}$ are then varied according to Eq. (8.6) to generate the succeeding Markov chain state. Although calculations of, for example, the distribution of the number of fading events per unit time systematically over-sample the tail region of the pdf as a result of rapid transitions across the outage threshold induced by statistical noise, such a procedure can accurately predict the statistics of small channel gains and long fade durations.

8.2 Numerical results

In Fig. 8.1 we display the pdf of the number of fading events during a $T_F = 1$ sec interval for a $N_1 = 25$, $N_2 = 26$ MEDS channel simulator with $f_{\max} = 10$ Hz. This calculation employs two 10^5-sample iterations of the multicanonical procedure for outage levels, R, of 5 dB (dashed-dotted line), 10 dB (dashed-line) and 15 dB (solid line) below the mean channel gain, respectively, together with the modified transition matrix method (+ markers), and from a single realization of Eq. (8.2) ($\circ$ markers). While both the multicanonical and transition matrix methods concentrate samples in regions of low probability, the modified transition matrix procedure continuously revises the intermediate pdf estimate and is therefore more numerically efficient.

Next, Fig. 8.2 illustrates the distribution of fade durations obtained from three 2×10^6-sample iterations of the multicanonical method (solid lines) for outage levels of $R =$ 5 dB and 15 dB below the mean channel gain, together with the modified transition

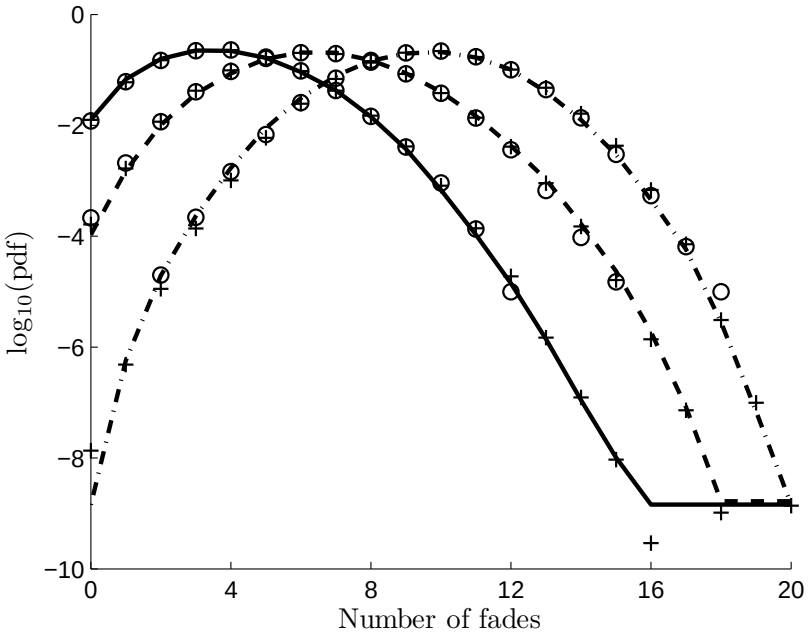


Figure 8.1: The base 10 logarithm of the pdf of the number of fading events for $f_{\max} = 10$ Hz within $T_F = 1$ sec for outage threshold levels of $R = 5$ dB (dashed-dotted line), 10 dB (dashed line) and 15 dB (solid line) below the mean channel gain as evaluated with two 10^5-sample iterations of the multicanonical method as well as for the modified transition matrix procedure (+) with a 2×10^5 samples and a direct evaluation of Eq. (8.2) ($\circ$).

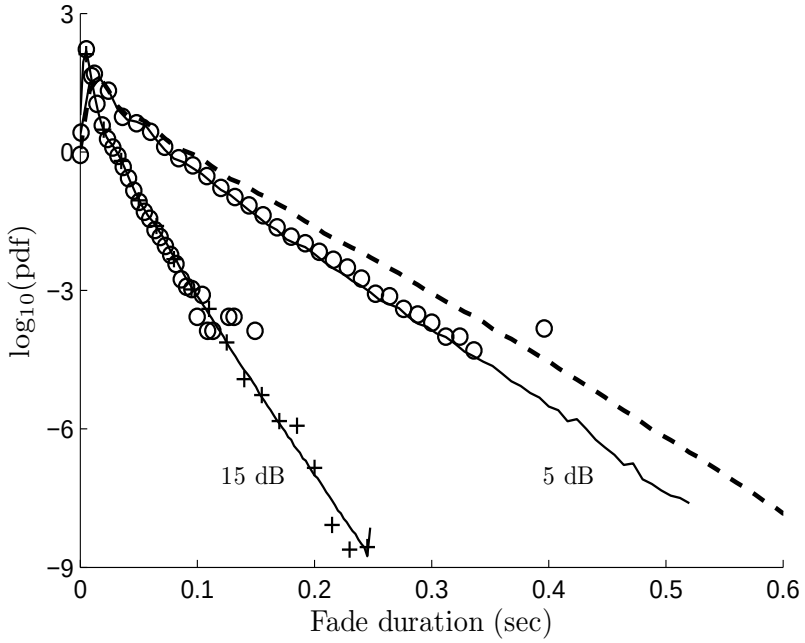


Figure 8.2: The fade duration distribution for three 2×10^6-sample iterations of the multicanonical method (solid lines) with $N = 100$ bins for outage levels of $R = 5$ dB and 15 dB below the mean channel gain, respectively, and for 6×10^6 samples of the modified transition matrix procedure (+) and Eq. (8.2) ($\circ$). The dashed line indicates the sampling bias introduced by the one-dimensional multicanonical calculation for $R = 5$ dB.

matrix (+ markers). Here 100 bins were employed to discretize the system observables over the x-axis interval. The dashed line in the figure is calculated with the standard one-dimensional multicanonical procedure for a threshold 5 dB below the mean channel gain. This clearly oversamples long fade durations since considerable correlation exists between consecutive short fading events. On the other hand, the two dimensional procedure far more accurately reproduces the correct fade duration distribution evaluated directly from Eq. (8.2) ($\circ$ markers) for thresholds $R > 5$ dB. For $R = 15$ dB, we employed one-dimensional procedures as mentioned in the preceding section.

The final calculation considers the BEP and the pdf of the decision variable D for $L = 2$, $N = 4$, an interference power level $P_I = 1$ (the power unit is arbitrary), a signal-to-interference (SIR) ratio, defined as P_R/P_I of 50 and a SNR, σ^2/P_I, of 12 dB. This simulation further employs 32×10^6 time steps of size $\delta = \pi/20$ and 2000 bins covering the interval $[-5, 20]$ in the decision variable D. Note here that for the accurate determination of the BEP, the right endpoint of one histogram interval must be located at the origin. With these values, the BEP calculated is 7.05×10^{-6}, which compares to the analytic value of 7.14×10^{-6} (the computed value varies by an amount on the order of 0.1×10^{-6} with differing random number sequences). The probability that a random realization possesses a decision variable value D (the pdf of D) is given by the solid line of Fig. 8.3, where all realizations with $D > 20$ are grouped into the last histogram bin. The dashed line of Fig. 8.3 displays the corresponding multicanonical result for 4 iterations of 5×10^6 samples.

Both the resulting BEP of 6.9×10^{-6} and the numerical pdf are comparable to the transition method result, although a larger number of multicanonical iterations with fewer samples will generally accurately predict the pdf over a wider range of values at the cost of a somewhat diminished overall level of numerical accuracy. The multicanonical and transition matrix methods can be combined as in [127] to extend this feature to the transition matrix procedure; however, the round-off error of the numerical subroutines must be sufficiently low to accommodate the large differences in the output magnitudes.

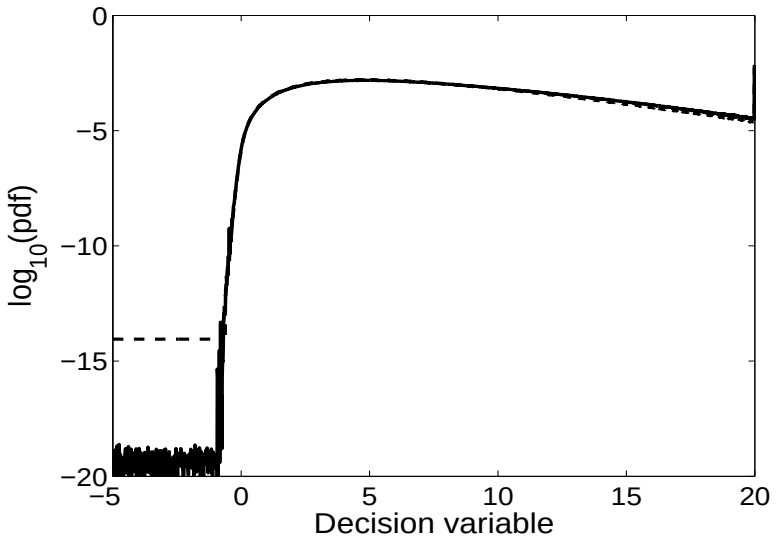


Figure 8.3: The pdf of the decision variable D associated with an optimum combining receiver and the system described in the text as computed with the modified transition matrix procedure (solid line) and the multicanonical procedure (dashed line).

8.3 Conclusions

This chapter has advanced new calculational procedures for time-independent and time-dependent wireless fading channels that can be extended to numerous problems throughout communications theory [127]. However, non-stochastic systems with memory or systems that are described by two or more observables require large transition matrix dimensions and presumably computation times.

Additionally, transition matrix compiles a history of changes in certain quantities that themselves must adequately encapsulate the relevant system behaviour. For example, if a particular region of system configuration space contributes significantly to the final result, these observables must vary significantly over this region. However, such limitations are shared by most other biasing techniques and therefore should not restrict the applicability of these methods.

Chapter 9

Conclusions

In this thesis, I have developed a general formalism for the rapid and accurate estimation of the frequency and temporal dynamics of optical polarization in the presence of PMD and PDL. While this formalism reproduces power-series expansion and differential equation solution techniques of previous authors, it also preserves the required Lorentz group symmetry of the Mueller matrix in every expansion order. This significantly improves the bandwidth of high PMD estimation accuracy, making the approach suited to the stochastic analysis of PMD and PDL induced system penalty, while also yielding physically realizable operator expansions that could with future research facilitate the design of joint PMD and PDL compensators [98,124,126]. Further, an analysis of data obtained from linear channel equalization filters in coherent optical communication systems in terms of Magnus orders of the Jones matrix could provide accurate real time monitoring of higher order PMD and its temporal statistics for the forecasting of PMD induced outage events and improved network control.

Next, I have demonstrated experimentally that low-amplitude mechanical excitations of commercially available dispersion compensation modules can excite high-frequency, $> 75 \times 10^3$ rotations/s, polarization transients that are nearly invariant between succes-

sive impacts. As a result, these mechanical excitations can induce transient degradations of the system performance in which an error-free system briefly exceeds acceptable performance thresholds before relaxing to its initial error-free state. Tracking the state of polarization throughout these transient events is expected to place challenging requirements on the speed and complexity of electronic compensation filters in modern coherent optical systems. Conversely, this work demonstrates the possibility of constructing high-frequency, reproducible PMD (and PDL) scramblers for network testing applications by placing a DCM or similar coils of optical fiber in series with one or more sections of PMD or polarization maintaining fiber. While simple to construct, this apparatus could provide a straightforward means of testing and verifying network survivability during rapid polarization transient events.

In the remainder of this thesis, I have examined two models of the temporal evolution of polarization characterized by a stochastic evolution of the PMD vector over the Poincaré sphere involving (1) a fiber emulator model that reproduces the expected temporal PMD autocorrelation statistics with a small number of emulator sections and (2) a general transition matrix formalism that additionally applies to other time-dependent communication systems [128].

While previous numerical simulations of the PMD temporal dynamics have been limited to qualitative models or complex implementations involving the evolution of PMD in the distributed limit, the random walk of method (1) accurately reproduces the PMD temporal autocorrelation function with model parameters that relate in a simple manner to the expected autocorrelation time of the PMD vector. The increased computational efficiency afforded by this approach allows simulations of the PMD temporal dynamics to be preferentially biased towards regions of low probability using the multicanonical method, permitting for example the calculation of PMD induced outage probabilities for thresholds that are otherwise intractable for standard Monte-Carlo simulations. Extensions of this approach to more general optical systems, including a finite number of active polarization

hinge sites, should allow the accurate simulation of PMD and PDL temporal dynamics for improved network equalization and control.

Finally, I have demonstrated that transition matrix methods can rapidly and accurately model the temporal dynamics of statistically unlikely but physically significant configurations in both optical and general communication systems. Transition matrix procedures, which extend multicanonical techniques to dynamic problems, possess numerous implementations that differ considerably in accuracy and sensitivity to variations in numerical parameters. In particular, in this thesis I have achieved a high degree of efficiency and accuracy by employing the multicanonical acceptance rule while continually updating the pdf estimate from the transition matrix. However, acceptance rules that increase the probability of asymmetric transitions out of a bin can be simple to program and computationally efficient if the numerical parameters ensure a small self-transition probability within each histogram bin. Further pursuing this line of reasoning could yield even more efficient transition matrix methods with possible experimental relevance, as in Refs. [75, 76]. Finally, I have demonstrated the applicability of transition matrix sampling procedures by calculating the outage dynamics associated with the hinge model of polarization evolution and separately wireless fading channels. Calculations such as these could assist in the active reconfiguration of communication systems in the event of a large outage probability or, conversely, for accurately determining the distribution of the outage duration.

Appendix A

Matrix Identities

A.1 Pauli matrix identities

The following identities, established through direct calculation, are valid for an arbitrary polarization Jones vector, $|A\rangle = [A_x, A_y]^{\mathrm{T}}$, and 3×1 Stokes vector $\vec{A} = \langle A|\vec{\boldsymbol{\sigma}}|A\rangle$, respectively [6, 39, 47, 55, 64, 98].

$$\begin{aligned} \mathbf{I}_2 &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \boldsymbol{\sigma}_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \\ \boldsymbol{\sigma}_2 &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \boldsymbol{\sigma}_3 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \end{aligned} \tag{A.1}$$

The matrices $\boldsymbol{\sigma}_a$, $a = 1, 2, 3$, satisfy the orthogonality relation

$$\frac{1}{2}\mathrm{Trace}\,\{\boldsymbol{\sigma}_a \boldsymbol{\sigma}_b\} = \delta_{ab} \tag{A.2}$$

where δ_{ab} is the Kronecker delta function. Consequently, $\boldsymbol{\sigma}_a$ together with the identity matrix $\mathbf{I}_2$ form a basis for the space of 2×2 matrices in that an arbitrary Jones matrix $\mathbf{T}$

can be written as the linear combination

$$\mathbf{T} = t_0 \mathbf{I}_2 + t_1 \boldsymbol{\sigma}_1 + t_2 \boldsymbol{\sigma}_2 + t_3 \boldsymbol{\sigma}_3 \tag{A.3}$$

involving the complex numbers, t_0 and t_n, $n = 1, 2, 3$, where $t_0 = \text{Trace}\{\mathbf{T}\}/2$ and $t_n = \text{Trace}\{\boldsymbol{\sigma}_n \mathbf{T}\}/2$ [47]. The coefficients of this expansion are real valued if $\mathbf{T}$ is hermitian, while for unitary $\mathbf{T}$, t_0 is real and t_n is purely imaginary.

The following identities can be established by applying

$$\boldsymbol{\sigma}_a \boldsymbol{\sigma}_b = \delta_{ab} \, \mathbf{I}_2 + i \epsilon_{abc} \, \boldsymbol{\sigma}_c \tag{A.4}$$

in which ϵ_{abc} is the Levi-Civita tensor and $a, b, c = 1, 2, 3$.

$$|A\rangle\langle A| = \frac{1}{2}|\vec{A}| \; \mathbf{I}_2 + \frac{1}{2}\vec{A} \cdot \vec{\boldsymbol{\sigma}} \tag{A.5}$$

$$\langle A|(\vec{B} \cdot \vec{\boldsymbol{\sigma}})|A\rangle = \vec{B} \cdot \vec{A} \tag{A.6}$$

$$\langle A|(\vec{B} \times \vec{\boldsymbol{\sigma}})|A\rangle = \vec{B} \times \vec{A} \tag{A.7}$$

$$(\vec{A} \cdot \vec{\boldsymbol{\sigma}})^2 = |\vec{A}|^2 \; \mathbf{I}_2 \tag{A.8}$$

$$(\vec{A} \cdot \vec{\boldsymbol{\sigma}})(\vec{B} \cdot \vec{\boldsymbol{\sigma}}) = (\vec{A} \cdot \vec{B}) \; \mathbf{I}_2 + i(\vec{A} \times \vec{B}) \cdot \vec{\boldsymbol{\sigma}} \tag{A.9}$$

$$(\vec{A} \cdot \vec{\boldsymbol{\sigma}})\vec{\boldsymbol{\sigma}}(\vec{A} \cdot \vec{\boldsymbol{\sigma}}) = 2\vec{A}(\vec{A} \cdot \vec{\boldsymbol{\sigma}}) - |\vec{A}|^2 \vec{\boldsymbol{\sigma}} \tag{A.10}$$

$$(\vec{A} \cdot \vec{\boldsymbol{\sigma}})\vec{\boldsymbol{\sigma}}(\vec{B} \cdot \vec{\boldsymbol{\sigma}}) = \vec{A}(\vec{B} \cdot \vec{\boldsymbol{\sigma}}) + \vec{B}(\vec{A} \cdot \vec{\boldsymbol{\sigma}}) - (\vec{A} \cdot \vec{B})\vec{\boldsymbol{\sigma}} - i(\vec{A} \times \vec{B})\mathbf{I}_2 \tag{A.11}$$

In the following, $\vec{W} = \vec{A} + i\vec{B}$ is a complex three-dimensional vector, $\vec{W}^* = \vec{A} - i\vec{B}$ and

$|\vec{W}|^2 \equiv \vec{W}^\dagger \vec{W}$.

$$(\vec{W} \cdot \vec{\boldsymbol{\sigma}})^\dagger = \vec{W}^* \cdot \vec{\boldsymbol{\sigma}} \tag{A.12}$$

$$\vec{\boldsymbol{\sigma}}(\vec{W} \cdot \vec{\boldsymbol{\sigma}}) = \vec{W}\mathbf{I}_2 + i\vec{W} \times \vec{\boldsymbol{\sigma}} \tag{A.13}$$

$$(\vec{W} \cdot \vec{\boldsymbol{\sigma}})^\dagger \vec{\boldsymbol{\sigma}} = \vec{W}^*\mathbf{I}_2 - i\vec{W}^* \times \vec{\boldsymbol{\sigma}} \tag{A.14}$$

$$(\vec{W} \cdot \vec{\boldsymbol{\sigma}})^\dagger \vec{\boldsymbol{\sigma}}(\vec{W} \cdot \vec{\boldsymbol{\sigma}}) = \vec{W}(\vec{W}^* \cdot \vec{\boldsymbol{\sigma}}) + \vec{W}^*(\vec{W} \cdot \vec{\boldsymbol{\sigma}}) - |\vec{W}|^2\vec{\boldsymbol{\sigma}} - i(\vec{W}^* \times \vec{W})\mathbf{I}_2 \tag{A.15}$$

A.2 Random Jones rotation matrices

Random polarization rotations are often required in numerical simulation for example when randomizing the input polarization states of channels in a WDM system, scrambling the polarization at each step of a "coarse-step" split-step Fourier (SSF) calculation [109] or simulating strong mode coupling between segments of a fiber emulator. Quite often, both in software and in the literature, rotations of the form

$$\mathbf{U} = \begin{bmatrix} \cos\theta & \sin\theta\, \mathrm{e}^{i\phi} \\ -\sin\theta\, \mathrm{e}^{-i\phi} & \cos\theta \end{bmatrix} \tag{A.16}$$

are employed where the angles θ and ϕ are typically selected from a uniform distribution over $[0, 2\pi]$. However, Eq. (A.16) when applied to an input Jones vector $|A_{in}\rangle$ according to $|A_{out}\rangle = \mathbf{U}|A_{in}\rangle$ with a uniform distribution of angles produces an output Stokes vector $\vec{A}_{out} = \langle A_{out}|\vec{\boldsymbol{\sigma}}|A_{out}\rangle$ that tends to "cluster" in regions of higher probability density on the Poincaré sphere. For example, in Fig. A.1a regions of higher probability density are apparent near the s_2 axis when the output Stokes vector is calculated for 10^4 random realizations of Eq. (A.16) for a 45 deg polarized input SOP. In what follows, I will present a simple algorithm for generating uniformly distributed 2×2 unitary matrices that yield a uniform distribution of output Stokes vectors over the Poincaré sphere.

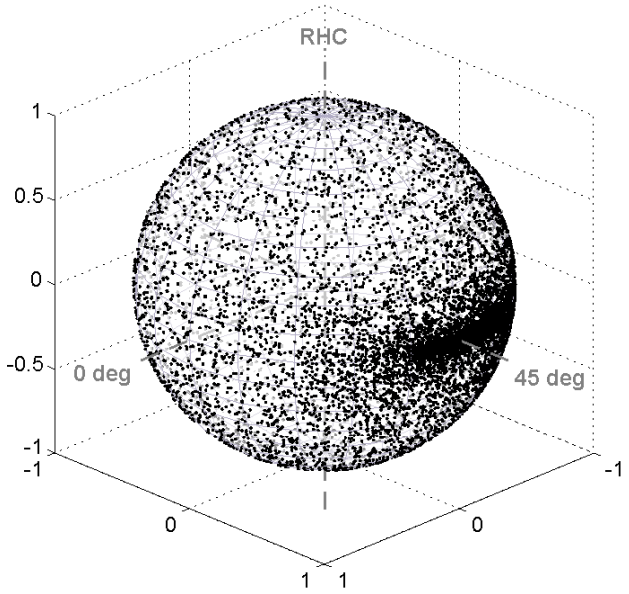
To proceed, I observe that

$$\mathbf{U} = \mathrm{e}^{i\delta} \begin{bmatrix} \cos\theta\, \mathrm{e}^{i\chi} & \sin\theta\, \mathrm{e}^{i\phi} \\ -\sin\theta\, \mathrm{e}^{-i\phi} & \cos\theta\, \mathrm{e}^{-i\chi} \end{bmatrix} \tag{A.17}$$

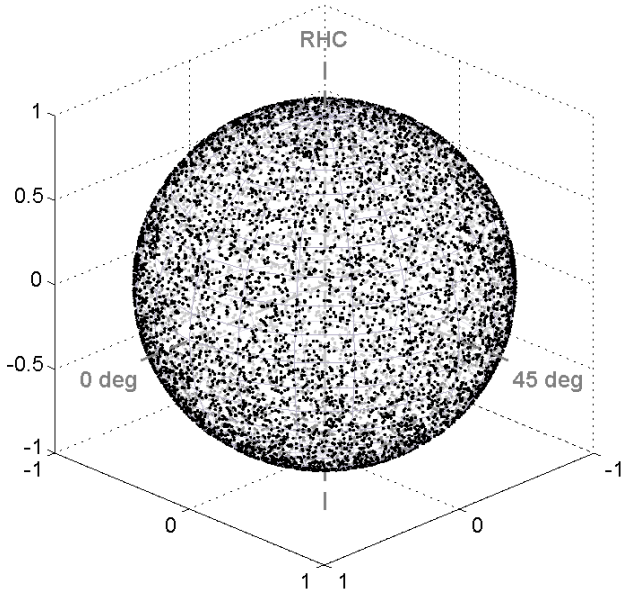
for rotation angles, δ, θ, ϕ and χ, is the most general form of a 2×2 unitary matrix and consequently covers the $U(2)$ group. Next, I select the probability densities of the various angles in Eq. (A.17) to ensure that the rotation matrix is uniformly distributed with respect to the Haar measure [32], that is, the invariant measure of volume for an algebraic group, for $U(2)$. The final result is remarkably straightforward:

1. The angles δ, ϕ and χ in Eq. (A.17) must be uniformly distributed over $[0, 2\pi]$
2. θ has probability density $f_\theta(x) = \sin(2x)$ with $0 \leq x \leq \pi/2$, and, consequently, $\cos(2\theta)$ is uniformly distributed over $[-1, 1]$.

Finally, to generate random values of $\cos\theta$ and $\sin\theta$ with the correct distribution I set $\cos\theta = \sqrt{u}$ and $\sin\theta = \sqrt{1-u}$ for a uniform random variable u over $[0, 1]$. Fig. A.1b illustrates the output Stokes vector on the Poincaré sphere resulting from this algorithm and clearly shows the uniform distribution of the output vector on the Poincaré sphere. This algorithm is easily implemented in, e.g. Matlab, and generates 10^6 random rotation matrices in < 0.5 sec on a standard Pentium Core-2 processor.



(a) Standard method



(b) Uniformly distributed rotation matrix

Figure A.1: A Monte-Carlo simulation of the output Stokes vector $\vec{A}_{out} = \langle A_{out}|\vec{\boldsymbol{\sigma}}|A_{out}\rangle$ with $|A_{out}\rangle = \mathbf{U}|A_{in}\rangle$ and a 45 deg polarized input field, $|A_{in}\rangle = [1,1]^{\mathrm{T}}/\sqrt{2}$, for 10^4 random rotation matrices $\mathbf{U}$ generated according to (a) Eq. (A.16) and (b) Eq. (A.17).

Bibliography

[1] G. Agrawal. *Fiber-Optic communications systems*. John Wiley & Sons, second edition, 1997.

[2] A. Aiello and J.P. Woerdman. Linear algebra for Mueller calculus, 2004. arXiv:math-ph/0412061.

[3] C. Antonelli, C. Colamarino, A. Mecozzi, and M. Brodsky. A model for temporal evolution of PMD. *IEEE Photon. Technol. Lett.*, 20(9–12):1012–1014, 2008.

[4] C. Antonelli and A. Mecozzi. Theoretical characterization and system impact of the hinge model of PMD. *J. Lightwave Technol.*, 24(11):4064–4074, 2006.

[5] C. Antonelli, A. Mecozzi, K. Cornick, M. Brodsky, and M. Boroditsky. PMD-Induced penalty statistics in fiber links. *IEEE Photon. Technol. Lett.*, 17(5):1013–1015, 2005.

[6] R. Barakat. Exponential versions of the Jones and Mueller-Jones polarization matrices. *J. Opt. Soc. Amer. A*, 13(1):158–163, 1996.

[7] B. Berg. Introduction to multicanonical Monte Carlo simulations. *Fields Inst. Commun.*, 26(1):1–24, 2000.

[8] B. Berg and T. Neuhaus. Multicanonical algorithms for first-order phase transitions. *Phys. Lett. B*, 267(2):249–253, 1991.

[9] N. Bergano, C. Poole, and R. Wagner. Investigation of polarization dispersion in long lengths of single-mode fiber using multilongitudinal mode lasers. *J. Lightwave Technol.*, 5(11):1618–1622, 1987.

[10] G. Biondini and W. Kath. Polarization-mode dispersion emulator with Maxwellian lengths and importance sampling. *IEEE Photon. Technol. Lett.*, 16(2):789–791, 2004.

[11] S. Blanes, F. Casas, J. Oteo, and J. Ros. The Magnus expansion and some of its applications. *Phys. Lett.*, 470(5–6):151–238, 2009.

[12] M. Born and E. Wolf. *Principles of Optics*. Cambridge University Press, seventh edition, 1999.

[13] M. Boroditsky, M. Brodsky, N. Frigo, P. Magill, C. Antonelli, and A. Mecozzi. Outage probabilities for fiber routes with finite number of degrees of freedom. *IEEE Photon. Technol. Lett.*, 17(2):345–347, 2005.

[14] M. Boroditsky, K. Cornick, C. Antonelli, M. Brodsky, and S. Dods. Comparison of system penalties from first- and multiorder polarization mode dispersion. *IEEE Photon. Technol. Lett.*, 17(8):1650–1652, 2005.

[15] M. Brodsky, M. Boroditsky, P. Magill, and N. Frigo. Persistence of spectral variations in DGD statistics. *Opt. Expr.*, 13(11):4090–4095, 2005.

[16] H. Bulow, W. Baumert, H. Schmuch, F. Mohr, T. Schulz, F. Kuppers, and W. Weiershausen. Measurement of the maximum speed of PMD fluctuations in installed field fiber. In *WE4, OFC Techn. Digest*. OFC/NFOEC, 1999.

[17] C. Antonelli and M. Shtaif and M. Brodsky. Sudden death of entanglement induced by polarization mode dispersion. *Phys. Rev. Lett.*, 106(8):1950–1954, 2011.

[18] J. Cameron, L. Chen, X. Bao, and J. Stears. Time evolution of polarization mode dispersion in optical fibers. *IEEE Photon. Technol. Lett.*, 10(9):1265–1267, 2003.

[19] H. Chen, R. Jopson, and H. Kogelnik. On the bandwidth of higher-order polarization-mode dispersion: the Taylor series expansion. *Opt. Exp.*, 11(11):1270–1282, 2003.

[20] L. Chen, O. Chen, S. Hadjifaradji, and X. Bao. Polarization-mode dispersion measurement in a system with polarization dependent loss or gain. *IEEE Photon. Technol. Lett.*, 16(1):206–208, 2004.

[21] E. Collett. *Polarized light in fiber optics*, pages 379–431. The PolaWave Group, Lincroft, New Jersey, 2003.

[22] J. Damask. *Polarization optics in telecommunications*. Springer, 2005.

[23] G. Dattoli, J. Gallardo, and A. Torre. An algebraic view to the operatorial ordering and its applications to optics. *Rivista Del Nuovo Cimento*, 11(11):1–79, 1988.

[24] A. Djupsjobacka. On differential group-delay statistics for polarization-mode dispersion emulators. *J. Lightwave Technol.*, 19(2):285–290, 2001.

[25] H. Dong, S. Yan, J. Zhou, G. Ning, Y. Gong, and C. Wu. Generalized Mueller matrix method for polarization mode dispersion measurement in a system with polarization-dependent loss or gain. *Opt. Expr.*, 14(12):5067–5072, 2006.

[26] X. Dong. Multicanonical simulation of communication systems. In *IEEE Radio and Wireless Symposium*, pages 205–207, 2007.

[27] A. dos Santos, M. Jimenez, J. von der Weid, and A. Djupsjobacka. BER and power penalty fluctuations due to PMD statistics in 10 Gbit/s optical transmissions. *Electron. Lett.*, 38(12):577–579, 2002.

[28] D. Dumas. Polarization effects of mechanical impacts on DCMs. Master's thesis, University of Waterloo, 2009.

[29] W. Eickhoff, Y. Yen, and R. Ulrich. Wavelength dependence of birefringence in single-mode fiber. *Appl. Opt.*, 20(19):3428–3435, 1981.

[30] A. Eyal, W. K. Marshall, M. Tur, and A. Yariv. Representation of second-order polarization mode dispersion. *Electron. Lett.*, 41(5):1658–1659, 1999.

[31] A. Eyal and M. Tur. A modified Poincaré sphere technique for the determination of polarization-mode dispersion in the presence of differential gain/loss, ThR1. In *OFC Tech. Digest*, page 340. OFC/NFOEC, 1998.

[32] J. Faraut. *Analysis on Lie groups: an introduction*. Cambridge Univ. Press, 2008.

[33] J. Ferreira, G.Vilela de Faria, and J. P. von der Weid. Simulating temporal evolution of PMD-related variables in optical fibers and emulators. *Micr. and Opt. Technol. Lett.*, 50(7):1950–1954, 2008.

[34] M. Fitzgerald, R. Picard, and R. Silver. Monte Carlo transition dynamics and variance reduction. *J. Stat. Phys.*, 98(1–2):321–345, 2000.

[35] C. Fludger, T. Duthel, D. van den Borne, C. Schulien, E. Schmidt, T. Wuth, J. Geyer, E. De Man, G. Khoe, and H. de Waardt. Coherent equalization and POLMUX-RZ-DQPSK for robust 100-GE transmission. *J. Lightwave Technol.*, 26(1):64–72, 2009.

[36] E. Forestieri and L. Vincetti. Exact evaluation of the Jones matrix of a fiber in the presence of polarization mode dispersion of any order. *J. Lightwave Technol.*, 19(12):1898–1909, 2001.

[37] G. Foschini and C. Poole. Statistical theory of polarization dispersion in single mode fibers. *J. Lightwave Technol.*, 9(11):1439–1456, 1991.

[38] C. Francia, F. Bruyere, D. Penninckx, and M. Chbat. PMD second-order effects on pulse propagation in single-mode optical fibers. *IEEE Photon. Technol. Lett.*, 10(12):1739–1741, 1998.

[39] N. Frigo. A generalized geometrical representation of coupled mode theory. *IEEE J. Quant. Electron.*, QE-22(11):2131–2140, 1986.

[40] A. Galtarossa and C. Menyuk, editors. *Polarization mode dispersion*. Springer, 2005.

[41] R. Ghulghazaryan, S. Hayryan, and C. Hu. Efficient combination of Wang-Landau and transition matrix Monte Carlo methods for protein simulations. *J. Comp. Chem.*, 28(3):715–726, 2006.

[42] N. Gisin, J. Von der Weid, and J. Pellaux. Polarization mode dispersion of short and long single-mode fibers. *J. Lightwave Technol.*, 9(7):821–827, 1991.

[43] N. Gisin and B. Huttner. Combined effects of polarization mode dispersion and polarization dependent losses in optical fibers. *Opt. Commun.*, 142(1–3):119–125, 1997.

[44] M. Glasner, D. Yevick, and T. Fysik. Generalized propagation techniques for longitudinally varying refractive index distributions. *Math. Comp. Model.*, 16(10):177–182, 1992.

[45] M. Glasner, D. Yevick, and B. Hermansson. Generalized propagation formulas of arbitrarily high order. *J. Chem. Phys.*, 95(11):8266–8272, 1991.

[46] M. Glasner, D. Yevick, and B. Hermansson. High-order generalized propagation techniques. *J. Opt. Soc. Amer. B*, 8(2):413–415, 1991.

[47] J. Gordon and H. Kogelnik. PMD fundamentals: Polarization mode dispersion in optical fibers. *Proc. Nat. Acad. Sci.*, 97(9):4541–4550, 2000.

[48] D. Gottlieb. Skew symmetric bundle maps on space-time. *Cont. Math.*, 220(11):117–141, 1998.

[49] H. Haunstein, T. Schorr, A. Zottmann, W. Sauer-Greff, and R. Urbansky. Performance comparison of MLSE and iterative equalization in FEC systems for PMD channels with respect to implementation complexity. *J. Lightwave Technol.*, 24(11):4047–4054, 2006.

[50] B. Heffner. Automated measurement of polarization mode dispersion using Jones matrix eigenanalysis. *IEEE Photon. Technol. Lett.*, 4(9):1066–1068, 1992.

[51] F. Heismann. Accurate Jones matrix expansion for all orders of polarization mode dispersion. *Opt. Lett.*, 28(21):2013–2015, 2003.

[52] F. Heismann. Comparison study of the average transfer matrix of first- and second-order PMD. *J. Lightwave Technol.*, 24(4):1929–1936, 2006.

[53] R. Holzlöhner and C. Menyuk. Use of multicanonical Monte Carlo simulations to obtain accurate bit error rates in optical communications systems. *Opt. Lett.*, 28(20):1894–1896, 2003.

[54] W. Huang. *Fiber birefringence modeling for polarization mode dispersion*. PhD thesis, University of Waterloo, 2007.

[55] B. Huttner, C. Geiser, and N. Gisin. Polarization-induced distortions in optical fiber networks with polarization-mode dispersion and polarization-dependent losses. *J. Sel. Top. Quant. Electron.*, 6(2):317–329, 2000.

[56] B. Huttner and N. Gisin. Anomalous pulse spreading in birefringent optical fibers with polarization-dependent losses. *Opt. Lett.*, 22(8):504–506, 1997.

[57] J. Ferreira and G. Vilela de Faria and J. P. von der Weid. Simulating temporal evolution of PMD-related variables in optical fibers and emulators. *Micro. and Opt. Technol. Lett.*, 50(7):1950–1954, 2008.

[58] R. Jopson, L. Nelson, and H. Kogelnik. Measurement of second-order polarization-mode dispersion vectors in optical fibers. *IEEE Photon. Technol. Lett.*, 11(9):1553–1555, 1999.

[59] M. Karlsson. Probability density functions of the differential group delay in optical fiber communication systems. *J. Lightwave Technol.*, 19(3):324–331, 2001.

[60] M. Karlsson and J. Brentel. Autocorrelation function of the polarization-mode dispersion vector. *Opt. Lett.*, 24(14):939–941, 1999.

[61] M. Karlsson, J. Brentel, and P. Andrekson. Long-term measurement of PMD and polarization drift in installed fibers. *J. Lightwave Technol.*, 18(7):941–951, 2000.

[62] S. Klarsfeld and J. Oteo. Analytic properties of the Magnus operator for two solvable Hamiltonians. *Phys. Lett. A*, 142(6–7):393–397, 1989.

[63] H. Kogelnik, L. Nelson, and P. Winzer. Second-order PMD outage of first-order compensated fiber systems. *IEEE Photon. Technol. Lett.*, 16(4):1053–1055, 2004.

[64] H. Kogelnik, L. E. Nelson, and J. P. Gordon. Emulation and inversion of polarization-mode dispersion. *J. Lightwave Technol.*, 21(2):482–495, 2003.

[65] H. Kogelnik, P. Winzer, L. Nelson, R. Jopson, M. Boroditsky, and M. Brodsky. First-order PMD outage for the hinge model. *IEEE Photon. Technol. Lett.*, 17(6):1208–1210, 2005.

[66] P. Krummrich and K. Kotten. Extremely fast (microsecond timescale) polarization changes in high speed long haul WDM transmission systems. In *FI3, OFC Tech. Digest*. OFC/NFOEC, 2003.

[67] P. Krummrich, E.-D. Schmidt, W. Weiershausen, and A. Mattheus. Field trial results on statistics of fast polarization changes in long haul WDM transmission systems. In *OThT6, OFC Tech. Digest*. OFC/NFOEC, 2005.

[68] T. Kudou, M. Iguchi, M. Masuda, and T. Ozeki. Theoretical basis of polarization mode dispersion equalization up to the second order. *J. Lightwave Technol.*, 18(4):614–617, 2000.

[69] D. Lao and A. Haimovich. Exact closed-form preformance analysis of optimum combining with multiple cochannel interferers and Rayleigh fading. *IEEE Trans. Commun.*, 51(6):995–2003, 2003.

[70] J. Lee, M. Kim, and Y. Chung. Statistical PMD emulator using variable DGD elements. *IEEE Photon. Technol. Lett.*, 15(1):54–56, 2003.

[71] A. Lima, I. Lima, T. Adah, and C. Menyuk. A novel polarization diversity receiver for PMD mitigation. *IEEE Photon. Technol. Lett.*, 14(4):465–457, 2002.

[72] A. Lima, I. Lima, and C. Menyuk. Error estimation in multicanonical Monte Carlo simulations with applications to polarization-mode dispersion emulators. *J. Lightwave Technol.*, 23(11):3781–3789, 2005.

[73] P. Lu, L. Chen, and X. Bao. Statistical distribution of polarization-dependent loss in the presence of polarization-mode dispersion in single-mode fibers. *IEEE Photon. Technol. Lett.*, 13(5):451–453, 2001.

[74] T. Lu, W. Huang, D. Yevick, M. O'Sullivan, and M. Reimer. A multicanonical comparison of PMD compensator performance. *J. Opt. Soc. Amer. A*, 22(12):2804–2809, 2005.

[75] T. Lu, D. Yevick, B. Hamilton, D. Dumas, and M. Reimer. An experimental realization of biased multicanonical sampling. *IEEE Photon. Technol. Lett.*, 17(12):1583–2585, 2005.

[76] T. Lu, D. Yevick, L. Yan, B. Zhang, and A. E. Willner. An experimental approach to multicanonical sampling. *IEEE Photon. Technol. Lett.*, 16(8):1978–1980, 2004.

[77] M. Shtaif and C. Antonelli and M. Brodsky. Nonlocal compensation of polarization mode dispersion in the transmission of polarization entangled photons. *Opt. Expr.*, 19(3):1728–1733, 2011.

[78] W. Magnus. On the exponential solution of differential equations for a linear operator. *Comm. Pure Appl. Math.*, VII:649–673, 1954.

[79] B. Marks, I. Lima, and C. Menyuk. Autocorrelation function for polarization mode dispersion emulators with rotators. *Opt. Lett.*, 27(13):1150–1152, 2002.

[80] A. Mecozzi, C. Antonelli, and M. Brodsky. Statistics of the polarization mode dispersion dynamics. *Opt. Lett.*, 32(20):3032–3034, 2007.

[81] A. Mecozzi and M. Shtaif. The statistics of polarization-dependent loss in optical communication systems. *IEEE Photon. Technol. Lett.*, 14(3):313–315, 2002.

[82] L. Nelson, M. Birk, S. Woodward, and P. Magill. Field measurements of polarization transients on a long-haul terrestrial link. Number ThT5 in 2011 IEEE Photon. Conf., 2011.

[83] L. Nelson, S. Woodward, S. Foo, X. Zhou, M. Feuer, D. Hanson, D. McGhan, H. Sun, M. Moyer, M. O'Sullivan, and P. Magill. Performance of a 46-Gbps dual-polarization QPSK transceiver with real-time coherent equalization over high PMD fiber. *J. Lightwave Technol.*, 27(3):158–167, 2009.

[84] L. Nelson, X. Zhou, N. Mac Suibhne, A. Ellis, and P. Magill. Experimental comparison of coherent polarization-switched QPSK to polarization-multiplexed QPSK for 10×100 km WDM transmission. *Opt. Expr.*, 19(11):10849–10856, 2011.

[85] R. Noé. PMD in high-bit-rate transmission and means for its mitigation. *J. Sel. Topics Quant. Electron.*, 10(2):341–355, 2004.

[86] A. Orlandini and L. Vincetti. A simple and useful model for Jones matrix to evaluate higher order polarization-mode dispersion effects. *IEEE Photon. Technol. Lett.*, 13(11):1176–1178, 2001.

[87] J. Oteo and J. Ros. From time ordered products to Magnus expansion. *J. of Math. Phys.*, 41(5):3268–3277, 2000.

[88] M. Pätzold. *Mobile fading channels*. John Wiley and Sons, 2002.

[89] P. B. Phua, J. M. Fini, and H. A. Haus. Real-time first- and second-order PMD characterization using averaged state-of-polarization of filtered signal and polarization scrambling. *J. Lightwave Technol.*, 21(4):982–989, 2003.

[90] C. Poole. Measurement of polarization-mode dispersion in single-mode fibers with random mode coupling. *Opt. Lett.*, 14(10):523–525, 1989.

[91] C. Poole, N. Bergano, R. Wagner, and H. Schulte. Polarization dispersion and principal states in a 147 km undersea lightwave cable. *J. Lightwave Technol.*, 6(7):1185–1190, 1988.

[92] C. Poole and T. Darcie. Distortion related to polarization-mode dispersion in analog lightwave systems. *J. Lightwave Technol.*, 11(11):1749–1759, 1993.

[93] C. Poole, R. Tkach, A. Chraplyvy, and D. Fishman. Fading in lightwave systems due to polarization-mode dispersion. *IEEE Photon. Technol. Lett.*, 3(1):68–70, 1991.

[94] M. Reimer. Modeling and simulation of polarization mode dispersion and polarization dependent loss. Master's thesis, University of Waterloo, 2007.

[95] M. Reimer and D. Yevick. A Clifford algebra analysis of polarization-mode dispersion and polarization-dependent loss. *IEEE Photon. Technol. Lett.*, 18(6):734–736, 2006.

[96] M. Reimer and D. Yevick. Application of the Magnus expansion to polarization-mode dispersion and polarization-dependent loss. In *OWI40, OFC Tech. Digest.* OFC/NFOEC, 2006.

[97] M. Reimer and D. Yevick. Least-squares analysis of the Mueller matrix. *Opt. Lett.*, 31(16):2399–2401, 2006.

[98] M. Reimer and D. Yevick. Mueller matrix description of polarization mode dispersion and polarization dependent loss. *J. Opt. Soc. Amer. A*, 23(6):1503–1508, 2006.

[99] M. Reimer and D. Yevick. Frequency interpolation of Jones matrices. *IEEE Photon. Technol. Lett.*, 19(9-12):801–803, 2007.

[100] K. Roberts, M. O'Sullivan, K.T. Wu, H. Sun, A. Awadalla, D. Krause, and C. Laperle. Performance of dual-polarization QPSK for optical transport systems. *J. Lightwave Technol.*, 27(16):3546–3559, 2009.

[101] P. Roberts and H. Ursell. Random walk on a sphere and on a Riemannian manifold. *Phil. Trans. Royal Soc. Lond.*, 252(1012):317–356, 1960.

[102] W. Salzman. An alternative to the Magnus expansion in time-dependent perturbation theory. *J. Chem. Phys.*, 82(2):822–826, 1985.

[103] W. Salzman. Convergence of Magnus and Magnus-like expansions in the Schrodinger representation. *J. Chem. Phys.*, 85(8):4605–4613, 1986.

[104] D. Sandel, V. Mirvoda, S. Bhandare, F. Wust, and R. Noé. Some enabling techniques for polarization mode dispersion compensation. *J. Lightwave Technol.*, 21(5):1198–1210, 2003.

[105] M. Shell, P. Debenedetti, and A. Panagiotopoulos. An improved Monte Carlo method for direct calculation of the density of states. *J. Chem. Phys.*, 119(18):9406–9411, 2003.

[106] W. Shieh. On the second-order approximation of PMD. *IEEE Photon. Technol. Lett.*, 12(3):290–292, 2000.

[107] M. Shtaif. The Brownian-bridge method for simulating polarization mode dispersion in optical communications systems. *IEEE Photon. Technol. Lett.*, 15(1):51–53, 2003.

[108] M. Shtaif and A. Mecozzi. Study of the frequency autocorrelation of the differential group delay in fibers with polarization mode dispersion. *Opt. Lett.*, 25(10):707–709, 2000.

[109] O. Sinkin, R. Holzlöhner, J. Zweck, and C. Menyuk. Optimization of the Split-Step Fourier method in modeling optical-fiber communications systems. *J. Lightwave Technol.*, 21(1):61–68, 2003.

[110] G. Smith and A. Bruce. A study of the multi-canonical Monte Carlo method. *J. Phys. A.*, 28(23):6623–6643, 1995.

[111] G. Soliman, M. Reimer, and D. Yevick. Measurement and simulation of polarization transients in dispersion compensation modules. *J. Opt. Soc. Amer. A*, 27(12):2532–2541, 2010.

[112] B. Tromborg, M. Reimer, and D. Yevick. Multicanonical evaluation of the tails of the probability density function of semiconductor optical amplifier output power fluctuations. *J. Quant. Electron.*, 46(1):57–61, 2010.

[113] D. Waddy, L. Chen, and X. Bao. Polarization effects in aerial fiber. *Opt. Fiber Technol.*, 11(3):1–19, 2005.

[114] P. Wai and C. Menyuk. Polarization mode dispersion, decorrelation, and diffusion in optical fibers with randomly varying birefringence. *J. Lightwave Technol.*, 14(2):148–157, 1996.

[115] J-S. Wang and R. Swendsen. Transition matrix Monte Carlo method. *J. Stat. Phys.*, 106:245–285, 2002.

[116] Z. Wang, C. Xie, and X. Ren. PMD and PDL impairments in polarization division multiplexing signals with direct detection. *Opt. Expr.*, 17(10):7993–8004, 2009.

[117] P. Wijesinghe, U. Gunawardana, and R. Liyanapathirana. An efficient algorithm for capacity and outage probability estimation in MIMO channels. *IEEE Commun. Lett.*, 15(6):644–646, 2011.

[118] R. Wilcox. Exponential operators and parameter differentiation in quantum physics. *J. Math. Phys.*, 8(4):962–982, 1967.

[119] P. Williams. PMD measurement techniques - avoiding measurement pitfalls. Venice summer school on PMD, pages 1–12, Venice, Italy, 2002.

[120] C. Xie, D. Werner, H. Haunstein, R. Jopson, S. Chandrasekhar, X. Liu, Y. Shi, S. Gronbach, T. Link, and K. Czotscher. Optical polarization mode dispersion compensators for high-speed optical communication systems. *Bell Labs Tech. Journ.*, 14(4):115–129, 2010.

[121] H. Yaffe and D. L. Peterson. Experimental determination of system outage probability due to first and second order PMD. *J. Lightwave Technol.*, 24(11):4155–4161, 2006.

[122] D. Yevick. Multicanonical communication system modeling - application to PMD statistics. *IEEE Photon. Technol. Lett.*, 14(11):1512–1514, 2002.

[123] D. Yevick. The accuracy of multicanonical system models. *IEEE Photon. Technol. Lett.*, 15(2):224–226, 2003.

[124] D. Yevick, M. Chanachowicz, M. Reimer, M. O'Sullivan, W. Huang, and T. Lu. Chebyshev and Taylor approximations of polarization mode dispersion for improved compensation bandwidth. *J. Opt. Soc. Amer. A*, 22(8):1662–1667, 2005.

[125] D. Yevick and T. Lu. Improved multicanonical algorithms. *J. Opt. Soc. Amer. A*, 23(11):2912–2918, 2006.

[126] D. Yevick, T. Lu, W. Huang, and W. Bardyszewski. Operator expansions for PMD analysis and compensation. *J. Opt. Soc. Amer. A*, 23(2):455–460, 2006.

[127] D. Yevick and M. Reimer. Transition matrix analysis of system outages. *IEEE Photon. Technol. Lett.*, 19(17-20):1529–1531, 2007.

[128] D. Yevick and M. Reimer. Modified transition matrix simulations of communication systems. *IEEE Commun. Lett.*, 12(10):755–757, 2008.

[129] D. Yevick, M. Reimer, H. Yaffe, P. Leo, D. Peterson, S. Wang, and K. Rochford. Frequency and duration of communication system outages resulting from polarization mode dispersion. *J. Lightwave Technol.*, 26(13-16):1901–1908, 2008.

[130] L. Yi, A. Eyal, P. Hedekvist, and A. Yariv. Measurement of high-order polarization mode dispersion. *IEEE Photon. Technol. Lett.*, 12(7):861–863, 2000.

[131] M. Yu, C. Kan, M. Lewis, and A. Sizmann. Statistics of polarization-dependent loss, insertion loss, and signal power in optical communication systems. *IEEE Photon. Technol. Lett.*, 14(12):1695–1697, 2002.

[132] W. Zhuang. Adaptive importance sampling for bit error rate estimation over fading channels. *J. Franklin Inst. - Eng. and Appl. Math.*, 331B(3):285–297, 1994.