Summer 2012

Analytical and computational methods for the study of rare event probabilities in dispersive and dissipative waves

Daniel S. Cargill
New Jersey Institute of Technology

ABSTRACT

ANALYTICAL AND COMPUTATIONAL METHODS FOR THE STUDY OF RARE EVENT PROBABILITIES IN DISPERSIVE AND DISSIPATIVE WAVES

by
Daniel S. Cargill

The main focus of this dissertation is the application of importance sampling (IS) to calculate the probabilities associated with rare events in nonlinear, large-dimensional lightwave systems that are driven by noise, including models for fiber-based optical communication system and mode-locked lasers. Throughout the last decade, IS has emerged as a valuable tool for improving the efficiency of simulating rare events in such systems. In particular, it has shown great success in simulating various sources of transmission impairments found in optical communication systems, with examples ranging from large polarization fluctuations resulting from randomly varying fiber birefringence to large pulse-width fluctuations resulting from imperfections in the optical fiber. In many cases, the application of IS is guided by a low-dimensional reduction of the system dynamics. Combining the low-dimensional reduction with Monte Carlo simulations of the original system has been shown to be an extremely effective scheme for computing, for example, the probability with which a pulse deviates significantly from its initial form due to a random forcing. In the context of nonlinear optics, this might represent a transmission error where the propagation model is the nonlinear Schrödinger equation (NLSE) with additive or multiplicative noise.

A shortcoming of this method is that the efficiency of the IS technique depends strongly on the accuracy of the low-dimensional reduction used to guide the simulations. These low-dimensional reductions are often derived from a formal perturbation theory, referred to as soliton perturbation theory (SPT) for the case of soliton propagation under the forced NLSE. As demonstrated here, such reduction methods

are often inadequate in their description of the pulse's dynamics. In particular, the interaction between a propagating pulse and dispersive radiation leads to a radiation-induced drift in a pulse's phase, which is largely unaccounted for in the reduced systems currently in use.

The first part of this dissertation is devoted to understanding the interaction between a pulse and dispersive radiation, leading to the derivation of an improved reduced system based on a variational approach. Once this system is derived and verified numerically, it serves as the basis for an improved IS method that incorporates the dynamics of the radiation, which is subsequently extended to more realistic propagation models. Of particular interest is the case of the NLSE with a periodic modulation of the dispersion constant, referred to as dispersion management (DM), and a related model where this modulation is averaged to give an autonomous, nonlocal equation. Following the nomenclature commonly use in literature, the former (nonautonomous) equation will be referred to as the NLSE+DM and the latter (autonomous) equation as the DMNLSE. A complicating aspect of these more realistic models is that, unlike the NLSE, exact solutions only exist as numerical objects rather than as closed-form solutions, which introduces an addition source of error in the derivation of a reduced system for the pulse dynamics.

In the second part of this dissertation, the IS method is extended to the calculation of phase-slip probabilities in mode-locked lasers (MLL). Realistic models for pulse propagation in MLL include the dissipative effects of gain and loss, in addition to nonlocal saturation effects. As a result most of the reduced systems derived for pulse dynamics are extremely complicated, which diminishes their applicability as guides for IS simulations. Therefore, a MLL operating in the soliton propagation regime is considered, where the effects of gain, loss and saturation are treated perturbatively. A simple reduced system for the pulse dynamics is derived for this MLL model, allowing the IS technique to be effectively applied.

ANALYTICAL AND COMPUTATIONAL METHODS FOR THE STUDY OF RARE EVENT PROBABILITIES IN DISPERSIVE AND DISSIPATIVE WAVES

by
Daniel S. Cargill

A Dissertation
Submitted to the Faculty of
New Jersey Institute of Technology and
Rutgers, The State University of New Jersey – Newark
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Mathematical Sciences

Department of Mathematical Sciences, NJIT
Department of Mathematics and Computer Science, Rutgers-Newark

August 2012

APPROVAL PAGE

ANALYTICAL AND COMPUTATIONAL METHODS FOR THE STUDY OF RARE EVENT PROBABILITIES IN DISPERSIVE AND DISSIPATIVE WAVES

Daniel S. Cargill

Dr. Richard O. Moore, Dissertation Advisor — Date
Associate Professor of Mathematical Sciences, New Jersey Institute of Technology

---

Dr. John Bechtold, Committee Member — Date
Professor of Mathematical Sciences, New Jersey Institute of Technology

---

Dr. Roy H. Goodman, Committee Member — Date
Associate Professor of Mathematical Sciences, New Jersey Institute of Technology

---

Dr. David J. Horntrop, Committee Member — Date
Associate Professor of Mathematical Sciences, New Jersey Institute of Technology

---

Dr. Colin McKinstrie, Committee Member — Date
Scientist, Advanced Photonics Department, Bell Labs, Alcatel-Lucent

BIOGRAPHICAL SKETCH

Author: Daniel S. Cargill

Degree: Doctor of Philosophy

Date: August 2012

Undergraduate and Graduate Education:

- Doctor of Philosophy in Applied Mathematics,
  New Jersey Institute of Technology, Newark, NJ, 2011
- Masters of Science in Industrial Mathematics,
  Rutgers University, Camden, NJ, 2007
- Bachelor of Science in Mathematics and Economics,
  University of Delaware, Newark, DE, 2005

Major: Applied Mathematics

Presentations and Publications:

Daniel S. Cargill, Richard O. Moore, and C. J. McKinstrie, "Noise bandwidth dependence of soliton phase in simulations of stochastic nonlinear Schrödinger equations," *Optics Letters*, Vol. 36, No. 9, pp. 1659-1661, 2011.

C. J. McKinstrie, D. S. Cargill, "Simultaneous frequency conversion, regeneration and reshaping of optical signal pulses," *Optics Express*, Vol. 26, No. 20, pp. 6881-6886, 2012.

D. S. Cargill, R. O. Moore, and C. J. McKinstrie, "Radiation-driven phase drift in stochastic nonlinear Schrödinger equations," *CLEO:2011*, Baltimore, MD, May 2011.

D. S. Cargill, R. O. Moore, and C. J. McKinstrie, "A variational approach for determining the effects of dispersive radiation on the parameters of Schrödinger solitons," *Nonlinear Evolution Equations and Wave Phenomena: Computation and Theory*, Athens, GA, April 2011.

D. S. Cargill, R. O. Moore, and C. J. McKinstrie, "A variational approach for determining the effects of dispersive radiation on the parameters of Schrödinger solitons," *International Congress on Industrial and Applied Mathematics*, Vancouver, Can., July 2011.

D. S. Cargill, R. O. Moore, and C. J. McKinstrie, "Noise dependent soliton phase distributions in simulations of stochastic nonlinear Schrödinger equations," *10th International Conference on Mathematical and Numerical Aspects of Waves*, Vancouver, Can., July 2011.

To my son Harrison,

for the moments of joy you have already given me and for those yet to come.

ACKNOWLEDGMENT

I would like to extend my deepest appreciation to my advisor and friend, Dr. Richard O. Moore. Through guidance and example, he has taught me what it means to be an applied mathematician and for this I am extremely grateful. I would also like to acknowledge my committee members, Dr. David J. Horntrop, Dr. Roy H. Goodman, Dr. John K. Bechtold, Dr. Gregory A. Kriegsmann and Dr. Colin J. McKinstrie, all of whom contributed invaluable comments and suggestions. I would like to individually thank Dr. Colin J. McKinstrie, who graciously allowed me to work alongside him at Bell Labs, and Dr. William L. Kath, who generously supported me over the final months of this work. I would also like to thank Dr. Daljit S. Ahluwalia, Dr. Jonathan H. Luke and Dr. Michael R. Booty, for creating a wonderful environment in which to learn, and Dr. John A. Pelesko for introducing me to the world of applied mathematics and the beauty it contains.

I would like to thank my fellow graduate students Matt Maleji, Jacek Wrobel, Matthew Causley, Daniel Fong, Oleksiy Varfolomiyev, Jeffrey Pohlmeyer and Te-Sheng Lin. I enjoyed our numerous mathematical and philosophical discussions, and I hope you have learned as much from me as I have from you. Thank you for being great friends and superb colleagues. I wish you all the best in your future work.

Finally, I wish to thank my family for showing an unmeasurable amount of support and understanding. It is only through their individual and collective sacrifice that this achievement was possible and my gratitude extends beyond what words can express.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

Chapter **Page**

TABLE OF CONTENTS
(Continued)

Chapter **Page**

TABLE OF CONTENTS
(Continued)

Chapter **Page**

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

LIST OF FIGURES
(Continued)

Figure **Page**

LIST OF FIGURES
(Continued)

Figure **Page**

LIST OF FIGURES
(Continued)

Figure **Page**

LIST OF FIGURES
(Continued)

Figure **Page**

LIST OF FIGURES
(Continued)

Figure **Page**

LIST OF FIGURES
(Continued)

Figure **Page**

LIST OF FIGURES
(Continued)

Figure **Page**

CHAPTER 1

INTRODUCTION

Rare Events and the Importance Sampling Technique. In stochastic systems, the term *rare event* refers to a particular system configuration, i.e., state, that occurs with an exceedingly low probability. Despite their name, rare events are ubiquitous in nature and appear in many different contexts where stochastic effects play an important role in system dynamics. Although they will be discussed here in the context of light-wave systems [3, 4, 5, 6, 7, 8], rare events can also be found in problems from industrial routing [9], rogue waves [10, 11] and financial asset pricing [12]. In addition, they are not limited to complex systems, as evident in their relation to failures in the Gaussian elimination algorithm applied to random matrices [13].

Many different methods have been deployed to study rare events, the most common of which are based on modifications of the Monte Carlo method [14, 15, 16]. In its standard form however, the Monte Carlo method is computationally infeasible to implement for the study of rare events, simply because of the number of trials required to generate such an event is inversely proportional to the probability of the event occurring at random [17]. For example, one would expect to need on the order of 10^9 Monte Carlo trials to simulate a rare event occurring with a probability on the order of 10^{-9}. However, this problem can be addressed by *augmenting* the standard Monte Carlo method with the variance reduction technique known as importance sampling (IS) [15, 18, 19, 20, 21], resulting in what is commonly referred to as the importance sampled Monte Carlo (ISMC) method. The ISMC method works by biasing the noise realizations of the standard Monte Carlo method with the intention of producing rare events with greater frequency and thus increased efficiency. To account for the biasing, each simulation is weighted when constructing the histogram

of results. When implemented correctly, the advantages of the ISMC method are substantial, resulting in an increase in efficiency of several orders of magnitude over a standard Monte Carlo approach.

Throughout the last decade, the ISMC method has emerged as a valuable tool for improving the efficiency in simulating rare events [22]. In particular, it has shown great success in simulating rare events associated with transmission impairments found in optical communication systems [23]. However, one disadvantage of the ISMC technique is that it requires knowledge of the locations in sample space that are most likely to lead to rare events. This is usually not a problem for low-dimensional systems, unfortunately, most of the systems of interest in the context of rare events are high-dimensional due to the high dimensionality that is characteristic of quantitatively accurate models of physical systems. Because of this high-dimensionality, the application of the ISMC method is often supplemented with a low-dimensional reduction of the system's dynamics that is used to guide simulations [23]. This combination works well in most applications, however, a major shortcoming of this approach is that it couples the efficiency of the ISMC method to the accuracy of the low-dimensional reduction, and thus has the potential for failure in systems possessing complex dynamics, which can not be accurately captured by a low-dimensional approximation.

The low-dimensional reductions discussed above are usually derived from an asymptotic approach to the perturbations, e.g., soliton perturbation theory (SPT) for the case of soliton propagation under the stochastic nonlinear Schrödinger equation. As the author will demonstrate here, such reduction methods are often inadequate in their description of the pulse's dynamics. In particular, the interaction between a propagating pulse and dispersive radiation leads to a radiation-induced drift in the pulse's phase, which is unaccounted for in the reduced systems currently in use as guides to IS based methods.

Document Structure. The main focus of this dissertation is the application of the ISMC method to calculate the probabilities associated with rare events in nonlinear, high-dimensional light-wave systems that are driven by noise. This includes models for fiber-based optical communications systems, as well as mode-locked lasers.

Because IS is a central idea throughout this document, a detailed review of this technique is presented in Chapter 2. The ISMC method is also presented in this chapter, in the context of finding the probabilities associated with large pulse deviations due to propagation in a noisy optical fiber. This chapter begins by presenting a stochastic version of the nonlinear Schrödinger equation (NLSE) as a mathematical propagation model for noisy optical fiber, and continues with a discussion of the special solutions to the NLSE, known as solitons, and the perturbative technique known as soliton perturbation theory, which yields a low-dimensional approximation used to guide the ISMC method.

Chapter 3 of this dissertation is devoted to understanding the interaction between a pulse and dispersive radiation, leading to the derivation of an improved reduced system based on a variational approach. Once this system is derived and verified numerically, it serves as the basis for an improved ISMC method that incorporates the dynamics of the radiation, which is presented in Chapter 4.

Chapter 5 of this dissertation is devoted to extending the aforementioned improved ISMC method to more realistic propagation models. Of particular interest is the case of the NLSE with a periodic modulation of the dispersion constant, referred to as dispersion management (DM), which yields a varying coefficient version of NLSE (NLSE+DM), and a related model where this modulation is averaged to give an autonomous, nonlocal equation (DMNLSE). A complicating aspect of these more realistic models is that, unlike the NLSE, exact solutions only exist as numerical objects rather than as closed-form solutions, which introduces an addition source of error in the derivation of a reduced system for the pulse dynamics.

In Chapter 6, the ISMC method is extended to the calculation of phase-slip probabilities in mode-locked lasers (MLL). Importantly, realistic models for pulse propagation in MLL include dissipative effects such as gain and loss in addition to nonlocal effects such as saturation, resulting in extremely complicated systems for the reduced pulse dynamics, which severely limits their use as a guide for IS simulations. Therefore, a MLL operating in the soliton propagation regime is first considered, where gain, loss and saturation can be treated perturbatively. A reduced system for the pulse dynamics is derived for this MLL model, allowing for an effective ISMC method to be constructed. Chapter 6 concludes by considering MLL models for operation in the DM soliton regime, where pulses exhibit complex dynamics throughout propagation. Finally, Chapter 7 provides a summary of the document, along with a discussion of future directions in which this work can be extended.

CHAPTER 2

REVIEW OF IMPORTANCE SAMPLING FOR RARE EVENTS IN OPTICAL FIBER COMMUNICATIONS

As previously discussed, the ISMC method has been successfully applied to the direct simulation of transmission impairments caused by polarization mode dispersion [24, 25, 26, 27], noise induced perturbations [3, 28] and width fluctuations resulting from randomly varying dispersion [6]. Since the ISMC method is a fundamental idea throughout the remainder of this document, this chapter is dedicated to its detailed review, including its application to the calculation of rare event probabilities in soliton based optical fiber communication systems, as presented in [23]. The first part of this chapter reviews the stochastic propagation of solitons in optical fiber-based communication systems, with particular attention given to the formulation of soliton perturbation theory (SPT), which provides a method for finding a suitable biasing distribution when implementing the ISMC method. The second part outlines the ISMC method, first in a general setting, then extended to the specific application of finding probabilities associated with large deviations of the parameters of an optical soliton (optical pulse). The chapter concludes with results from numerical simulations and a discussion of an observed radiation-driven phase drift, which is investigated further in subsequent chapters.

2.1 Stochastic Soliton Evolution in Optical Fiber Communication Systems

An optical fiber is a flexible, cylindrical fiber made of nearly pure silica glass, which has an inner core and an outer cladding. The index of refraction in the core is slightly higher than in the cladding, which enables light to be trapped in the core

through total internal reflection [1]. The number of *spatial* modes a particular fiber can support is determined by the index of refraction and diameter of both the core and cladding. Optical fiber-based communication systems (almost) exclusively use single-mode fiber, i.e., fiber supporting only a single transverse mode, with typical core and cladding indices around 1.45 ± 0.005 and diameters of 5-10μm and 60-140μm, respectively [1].

Even with modern advances in both the manufacturing process and the ability to produce extremely pure silica glass, significant attenuation of optical signals can still occur over long distances. For this reason, long haul optical communication systems typically use pulses of relatively high optical power, which are maintained through periodic amplification based on stimulated emissions of a gain element [29, 30]. High optical power confined to a single mode (with small modal volume) leads to a polarization response which is nonlinearly dependent on the applied electric field. When combined with the linear effect of temporal dispersion, this results in an envelope modulation of the carrier wave into stable optical pulses which are governed by the well-studied nonlinear Schrödinger equation (NLSE) [31, 32] with a small linear loss term representing the attenuation. In addition, stimulated emission is always accompanied by *spontaneous* emission, taking the form of optical noise in the fiber. This is referred to as amplified spontaneous emissions noise (ASE noise) and is typically modeled by an augmentation of NLSE to include a stochastic forcing term.

The NLSE for optical fiber can be formally derived using a slowly-varying-envelope (SVE) approximation [31, 32, 33, 1] applied to Maxwell's equations, where nonlinear effects are included through a power series expansion for the polarization in terms of the electric field. The perturbations added to account for periodic amplification, result in periodic power variations in the propagating pulse [34, 35, 29]. If the period of these variations are small relative to the transmission distance, they can be effectively averaged out through asymptotic averaging (otherwise known as homogenization [36]),

yielding a stochastically forced NLSE as a first-order approximation for the evolution of the optical pulse envelopes [29]. In dimensionless form, this equation is given by

$$i\frac{\partial u}{\partial z} + \frac{1}{2}\frac{\partial^2 u}{\partial t^2} + |u|^2 u = i\sigma n(z,t), \tag{2.1}$$

where the variables z and t are, respectively, dimensionless distance and retarded time, and $u(z,t)$ is the complex electromagnetic field envelope. The derivation of the spatial mode profile in single-mode optical fiber and the SVE approximation that leads to NLSE with linear loss as the evolution equation for pulses in optical fiber is presented in Appendix A, along with the details of the homogenization method and the choice of dimensional quantities that lead to the dimensionless equation given in Equation (2.1).

Note that this equation provides no indication as to the modulation format, i.e., how information is encoded into optical pulses at the beginning of transmission, or the detection process, i.e., how the information is decoded to an electric signal, at the end of the transmission line. It does, however, provide a model for the evolution of the optical pulse, or equivalently for any associated measure of the pulse that may be used to encode information, e.g., amplitude or phase. The assumption here is that statistical information about these measures can be related to a bit error after a specific modulation format is chosen. Commonly used modulation formats and detector models are discussed in Appendix A, along with the implications of including these modulation/demodulation models as part of a more complete model for an optical communication system.

The term $n(z,t)$ in Equation (2.1) is a spatiotemporal stochastic forcing that represents additive ASE noise. Since the bandwidth of modern optical amplifiers is very large relative to the bandwidth of optical pulses, ASE noise is typically approximated by idealized (infinite bandwidth) white noise, which is Gaussian-distributed and mean-zero [37, 29]. When explicitly modeling individual amplifiers, the noise

takes the discrete form of

$$n(z,t) = \sum_{k=1}^{N_a} n_k(t)\,\delta(z - kz_a)\,, \tag{2.2a}$$

with

$$\mathbb{E}[n_k(t)] = 0 \quad \text{and} \quad \mathbb{E}[n_k(t)\bar{n}_j(t')] = \delta(t-t')\delta_{kj}, \tag{2.2b}$$

where N_a is the number of amplifiers, each of which is located at a point $z = kz_a$, with z_a the amplifier spacing. $\mathbb{E}[\cdot]$ represents the expected value, and $\delta(\cdot)$ and δ_{kj} are the Dirac and Kronecker delta functions, respectively, and the noise strength is given by σ^2. This noise model is appropriate for discretely placed amplifiers and therefore is the form used in all simulations, however, it is often mathematically convenient to take the limit as $z_a \to 0$ which yields a continuous noise representation, where the covariance is given by

$$\mathbb{E}[n(z,t)\bar{n}(z',t')] = \delta(t-t')\delta(z-z'), \tag{2.3}$$

where σ^2 is scaled appropriately such that the total noise power remains the same.

Note that both Equation (2.2) and (2.3) are mathematical idealizations, since noise in physical systems cannot have infinite bandwidth, or equivalently infinite power. However, the specific value of the noise bandwidth is not needed as long as it is larger than the soliton bandwidth (which is the case in practice) [38, 30], because the only noise components capable of *directly* affecting the soliton parameters are assumed to lie within the same spectral range as the soliton itself. On the other hand, *the noise bandwidth can indirectly affect the phase fluctuations*. This will be explicitly shown through numerical simulations in Section 2.4 of this chapter and is more thoroughly investigated in Chapter 3.

In the absence of noise ($\sigma = 0$), the NLSE in (2.1) admits a well-known four-parameter family of soliton solutions

$$u_{\mathrm{sol}}(z,t) = u_0(z,t)\exp(i\Theta(z,t))\,, \tag{2.4a}$$

where

$$u_0(z,t) = A\ \mathrm{sech}(A[t-T(z)])\,, \qquad \Theta(z,t) = \Omega\, t + \Phi(z), \tag{2.4b}$$

with

$$T(z) = T_0 + \Omega z \quad \text{and} \quad \Phi(z) = \frac{A^2-\Omega^2}{2}z + \Phi_0. \tag{2.4c}$$

The four soliton parameters A, Ω, T_0 and Φ_0 are *arbitrary* constants that can be traced to four invariances found within the NLSE[1] [29]. Each soliton parameter can be associated with a dimensionless physical value: the pulse amplitude A (and inverse width), the frequency Ω (relative to the carrier frequency and directly proportional to the group velocity), the *initial* mean timing T_0 and the *initial* phase offset Φ_0, which are both defined at $t = 0$. Each time noise is added to the system, part of the noise is incorporated into the soliton, where it produces small stochastic changes of the soliton parameters, resulting in a random walk of the four quantities A, Ω, T_0 and Φ_0 [39, 40, 29]. For typical system configurations, the noise power σ^2 is small, and thus, the noise-induced changes of the soliton parameters are also small. In rare cases, however, individual contributions can add coherently, resulting in large deviations of individual or multiple soliton parameters, thus creating the potential for transmission errors.

[1]Note that NLSE is integrable and therefore possesses an infinite number of conserved quantities. The implication of this point will be made clear in subsequent chapters when non-integrable and dissipative systems are considered.

2.1.1 Soliton Perturbation Theory

The effect of noise upon propagating pulses (more specifically on the pulse parameters), can be captured using the what is commonly referred to in the nonlinear optics literature as soliton perturbation theory (SPT) [29, 41]; it is essentially the method of multiple scales taken to first order in a small parameter characterizing the perturbation amplitude. The application of SPT to the perturbed NLSE in Equation (2.1) yields a first order approximation for the evolution of a soliton, in the form of a set of stochastic ODEs, one for each pulse parameter. These ODEs provide insight into the interactions between the pulse and the noise, and thus furnish the means to form an accurate biasing distribution, which is critical to effective implementation of the ISMC method.

The ability of noise to shift the soliton parameters is embedded in the relationship each parameter has with an underlying invariance in the NSLE equation. Since these parameters represent true degrees of freedom, any value of the parameters is permitted and the noise encounters no resistance when inducing these parameters to change. This allows small coherent perturbations to build upon each other, leading to large deviations from the initial parameter values. Furthermore, the two z-dependent terms in Equation (2.4c) indicate that small fluctuations in the frequency and amplitude integrate into large timing and phase shifts.

The assumption that the stochastic forcing manifest as adiabatic motion of the four soliton parameters suggests the introduction of the short length scale $z_1 = \sigma z$, which equivalently introduces an additional derivative in Equation (2.1), i.e.,

$$i\frac{\partial u}{\partial z} + i\sigma\frac{\partial u}{\partial z_1} + \frac{1}{2}\frac{\partial^2 u}{\partial t^2} + |u|^2 u = i\sigma n(z,t), \tag{2.5}$$

suggesting an expansion for the solution in the form

$$u = \left[v_0(t,z,z_1) + \sigma v_1(t,z,z_1)\right] \exp(i\Theta(t,z,z_1))\,. \tag{2.6}$$

The $O(1)$ equation is the NLSE applied to $v_0 \exp(i\Theta)$ and thus the first order solution is a soliton with parameters that now depend on the slow time scale,

$$v_0(t, z, z_1) \exp(i\Theta(t, z, z_1)) = u_{\text{sol}}(t, z, A(z_1), \Omega(z_1), T_0(z_1), \Phi_0(z_1)), \tag{2.7}$$

where u_{sol} is defined in Equation (2.4). The next order gives an evolution equation for the $O(\sigma)$ perturbation $v_1(t, z, z_1)$ (from hereafter referred to as the radiation), which takes the form

$$\begin{aligned} L_{\text{nls}}(v_1 \exp(i\Theta)\,; u_{\text{sol}}) = in(z,t) - i\Bigg[& \frac{\partial u_{\text{sol}}}{\partial A}\frac{dA}{dz_1} + \frac{\partial u_{\text{sol}}}{\partial T}\frac{\partial T(z,z_1)}{\partial z_1} \\ & + \frac{\partial u_{\text{sol}}}{\partial \Omega}\frac{d\Omega}{dz_1} + \frac{\partial u_{\text{sol}}}{\partial \Phi}\frac{\partial \Phi(z,z_1)}{\partial z_1}\Bigg], \end{aligned} \tag{2.8}$$

where $L_{\text{nls}}(\cdot; u_{\text{sol}})$ is the nonlinear Schrödinger operator linearized around the soliton solution in Equation (2.4). Since the $L_{\text{nls}}(v_1 \exp(i\Theta)\,; u_{\text{sol}})$ is in the form of a linear operator acting on $v_1 \exp(i\Theta)$, Fredholm theory [42] requires that the right hand side of Equation (2.8) be orthogonal to the zero eigenfunctions of the adjoint operator. Typically, this requirement is used to isolate the $O(\sigma)$ evolution of the pulse, which in this case is the evolution of the parameters. In the present form, the projections would be over both space and time variables. A simpler approach is to separate the evolutionary and transverse derivatives of $L_{\text{nls}}(v_1 \exp(i\Theta)\,; u_{\text{sol}})$ by rewriting Equation (2.8) as

$$\begin{aligned} \left(\frac{\partial}{\partial z} + \Omega\frac{\partial}{\partial t}\right) v_1 - L(v_1; u_0) = & \, n(z,t)\exp(-i\Theta(t,z,z_1)) \\ & - \left[\frac{\partial u_0}{\partial A}\frac{dA}{dz_1} + \frac{\partial u_0}{\partial T}\frac{\partial T}{\partial z_1} + iu_0 t\frac{d\Omega}{dz_1} + iu_0\frac{\partial \Phi}{\partial z_1}\right], \end{aligned} \tag{2.9a}$$

where

$$L(v_1; u_0) = \frac{i}{2}\frac{\partial^2 v_1}{\partial t^2} - \frac{i}{2}A^2 v_1 + 2i|u_0|^2 v_1 + i(u_0)^2 \bar{v}_1. \tag{2.9b}$$

The Fredholm condition can now be framed in terms of the zero eigenfunctions of L, with projections over only the t variable.

The price paid for this simplification is that the linear operator $L(\cdot; u_0)$ is non-normal, with a generalized nullspace that admits four eigenfunctions at the zero eigenvalue (two ordinary and two generalized), all four of which must be used in the projection. These eigenfunctions can be found by mapping each one back to its corresponding eigenfunction in L_{nls}, through the relation

$$\left(\frac{\partial}{\partial z} + \Omega\frac{\partial}{\partial t}\right) v - L(v; u_0) = -iL_{\text{nls}}(v\exp(i\Theta)\,; u_{\text{sol}})\exp(-i\Theta)\,. \tag{2.10}$$

By appealing to the four NLSE invariances that produced the free parameters A, Ω, T_0, Φ_0, it is easily shown that the solutions of $L_{\text{nls}}(v\exp(i\Theta)\,, u_{\text{sol}}) = 0$ are the derivatives of $u_{\text{sol}}(t, z; A, \Omega, T_0, \Phi_0)$ with respect to each of the four free parameters. To see this more clearly, let N_{nls} represent the nonlinear Schödinger operator, written as

$$N_{\text{nls}}(u) = i\frac{\partial u}{\partial z} + \frac{1}{2}\frac{\partial^2 u}{\partial t^2} + |u|^2 u, \tag{2.11}$$

and recall that the entire family of soliton solutions given in Equation (2.4) satisfies the equation $N_{\text{nls}}(u) = 0$. Now consider a solution of the form $u_{\text{sol}}^{(\epsilon)}(t, z; X+\epsilon)$, where X represents one of four free soliton parameters and ϵ is a small perturbation. Inserting this solution into Equation (2.11) gives

$$\begin{aligned} N_{\text{nls}}\left(u_{\text{sol}}^{(\epsilon)}\right) &= N_{\text{nls}}\left(u_{\text{sol}}^{(0)} + \epsilon\frac{\partial u_{\text{sol}}^{(0)}}{\partial X} + O\big(\epsilon^2\big)\right) \\ &= N_{\text{nls}}\left(u_{\text{sol}}^{(0)}\right) + \epsilon L_{\text{nls}}\left(\frac{\partial u_{\text{sol}}^{(0)}}{\partial X}; u_{\text{sol}}^{(0)}\right) + O\big(\epsilon^2\big)\,, \end{aligned} \tag{2.12}$$

implying that

$$L_{\text{nls}}\left(\frac{\partial u_{\text{sol}}^{(0)}}{\partial X}; u_{\text{sol}}^{(0)}\right) = 0. \tag{2.13}$$

Combining the results of Equation (2.13) with Equation (2.10), suggest the following definitions for the eigenfunctions of $L(\cdot; u_0)$:

$$v_A = \frac{1}{A}\frac{\partial}{\partial t}\left[(t-T)u_0\right], \quad v_T = -\frac{\partial u_0}{\partial t}, \quad v_\Omega = i(t-T)u_0, \quad v_\Phi = iu_0, \tag{2.14}$$

with corresponding relations,

$$L\left(v_A; u_0\right) = Av_\Phi, \quad L\left(v_T; u_0\right) = 0, \quad L\left(v_\Omega; u_0\right) = v_T, \quad L\left(v_\Phi; u_0\right) = 0. \tag{2.15}$$

Thus, v_Φ and v_T represent true eigenfunctions of L, whereas v_A and v_Ω represent generalized eigenfunctions; all four are plotted in Figure 2.1.

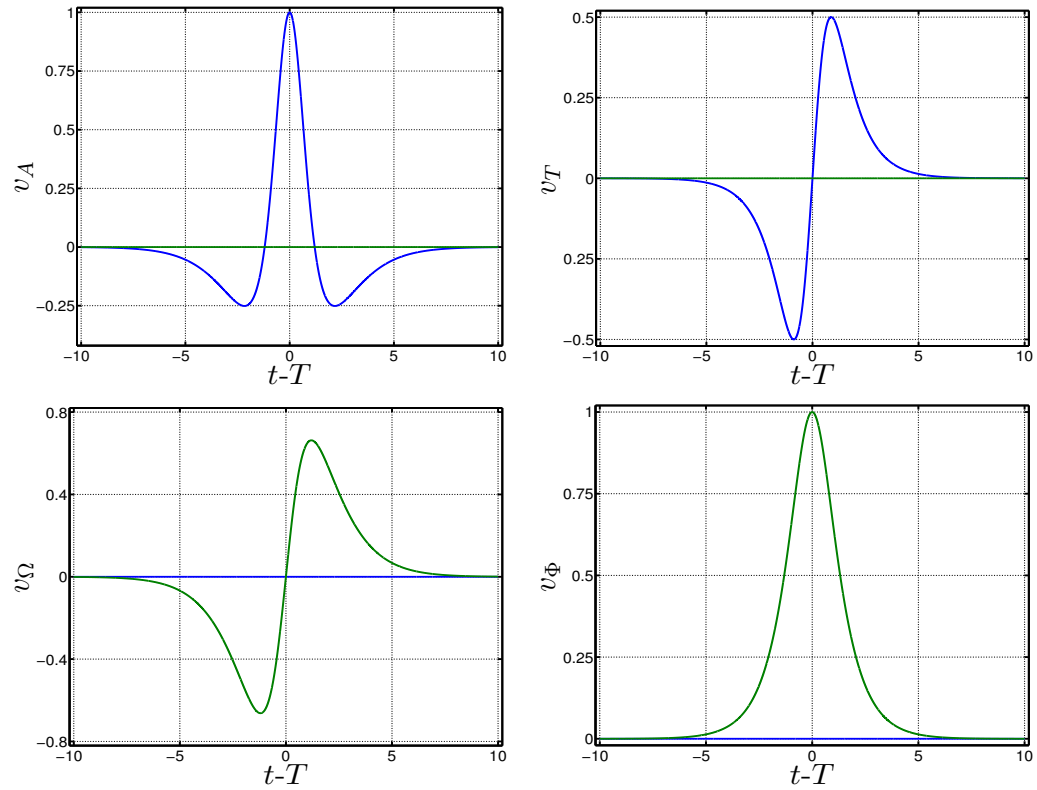


Figure 2.1 Real (blue) and imaginary (green) parts of the discrete soliton modes given in Equation (2.14), with $A = 1$ and Ω, T, $\Phi = 0$.

Under the standard definition for the inner product, i.e.,

$$\langle f, g\rangle = \mathrm{Re}\left[\int \bar{f}(t)\, g(t)\, dt\right], \tag{2.16}$$

the adjoint of $L(\cdot;u_0)$ is calculated to be

$$L^\dagger(v^\dagger;u_0) = -\frac{i}{2}\frac{\partial^2 v^\dagger}{\partial t^2} + \frac{i}{2}A^2 v^\dagger - 2i|u_0|^2 v^\dagger + i(u_0)^2\bar{v}^\dagger, \tag{2.17}$$

which shows that, in addition to being non-normal, $L(\cdot;u_0)$ it is also non-self-adjoint. Note that † denotes the adjoint space and, unless otherwise indicated, all integrals are over the entire real line. The eigenfunctions of $L^\dagger(\cdot;u_0)$ can be calculated by the relation $L^\dagger(v^\dagger;u_0) = iL(iv^\dagger;u_0)$, which gives

$$v_A^\dagger = -iv_\Phi, \quad v_T^\dagger = -i\frac{v_\Omega}{A}, \quad v_\Omega^\dagger = i\frac{v_T}{A}, \quad v_\Phi^\dagger = iv_A. \tag{2.18}$$

Under the inner product defined in Equation (2.16), the eigenfunctions of $L(\cdot;u_0)$ given in Equation (2.14) form an orthogonal set, i.e.,

$$\langle v_X, v_Y\rangle = \langle v_Y, v_X\rangle\,\delta_{XY}, \tag{2.19}$$

for $X, Y = A, T, \Omega, \Phi$. Likewise, the adjoint eigenfunctions given by Equations (2.18) are also mutually orthogonal, i.e.,

$$\left\langle v_X^\dagger, v_Y^\dagger\right\rangle = \left\langle v_Y^\dagger, v_X^\dagger\right\rangle\delta_{XY}. \tag{2.20}$$

More importantly, however, is that (as scaled) the eigenfunctions of $L(\cdot;u_0)$ together with their associated adjoint eigenfunctions form a biorthonormal basis for the generalized nullspace of $L(\cdot;u_0)$, i.e.,

$$\left\langle v_X^\dagger, v_Y\right\rangle = \delta_{XY}, \tag{2.21}$$

which can be used to solve for the evolution of each soliton parameter through projections in Equation (2.9). To see how this is done, first note that the radiation $(v_1(t,z,z_1))$ is orthogonal to the generalized nullspace of $L(\cdot;u_0)$, since it is initially zero and (as it will be shown) only grows from the portion of noise that is orthogonal

to each of the generalized eigenfunctions and its corresponding adjoint. Using the definitions in Equation (2.14), Equation (2.9) can be rewritten as

$$\begin{aligned}\left(\frac{\partial}{\partial z}+\Omega\frac{\partial}{\partial t}\right)v_1 - L(v_1;u_0) &= n(z,t)\exp(-i\Theta(t,z,z_1)) \\ &\quad - \left[v_A\frac{dA}{dz_1}+v_T\frac{\partial T}{\partial z_1}+\left(v_\Omega+Tv_\Phi\right)\frac{d\Omega}{dz_1}+v_\Phi\frac{\partial\Phi}{\partial z_1}\right].\end{aligned} \tag{2.22}$$

Projecting both sides of this equation against $v_X^\dagger$ gives

$$\begin{aligned}\left\langle v_X^\dagger, \left(\frac{\partial}{\partial z}+\Omega\frac{\partial}{\partial t}\right)v_1\right\rangle - \left\langle v_X^\dagger, L(v_1;u_0)\right\rangle &= \\ \frac{\partial}{\partial z}\left\langle v_1, v_X^\dagger\right\rangle - \left\langle v_1, \left(\frac{\partial}{\partial z}+\Omega\frac{\partial}{\partial t}\right)v_X^\dagger\right\rangle - \left\langle v_1, L^\dagger(v_X^\dagger;u_0)\right\rangle &= \\ \left\langle v_X^\dagger, n(z,t)\exp(-i\Theta(t,z,z_1))\right\rangle - \left[\frac{\partial X}{\partial z_1}+T\frac{d\Omega}{dz_1}\delta_{X\Phi}\right],&\end{aligned} \tag{2.23}$$

for $X = A$, T, Ω, Φ. Since $L^\dagger(v_X^\dagger;u_0)$ is either 0 or $v_Y^\dagger$ ($Y \neq X$) and v_1 is orthogonal to each of the generalized eigenfunctions, the first and third terms on the left side of this equation evaluate to zero. In addition, each adjoint eigenfunction $v_X^\dagger$ is only a function of $(t-\Omega z)$, which eliminates the second term on the left. This results in the following four equations:

$$\begin{aligned}\frac{dA}{dz_1} &= \mathrm{Re}\left[\int \bar{v}_A^\dagger \exp(-i\Theta)\, n(z,t)\, dt\right], \\ \frac{\partial T}{\partial z_1} &= \mathrm{Re}\left[\int \bar{v}_T^\dagger \exp(-i\Theta)\, n(z,t)\, dt\right], \\ \frac{d\Omega}{dz_1} &= \mathrm{Re}\left[\int \bar{v}_\Omega^\dagger \exp(-i\Theta)\, n(z,t)\, dt\right], \\ \frac{\partial\Phi}{\partial z_1} &= \mathrm{Re}\left[\int \left(\bar{v}_\Phi^\dagger - T\bar{v}_\Omega^\dagger\right) \exp(-i\Theta)\, n(z,t)\, dt\right].\end{aligned} \tag{2.24}$$

Combining these equations with the $O(1)$ derivative in z, yields evolution equations of the form,

$$\begin{aligned}
\frac{dA}{dz} &= \sigma \mathrm{Re}\left[\int \bar{v}_A^\dagger \exp(-i\Theta)\, n(z,t)\, dt\right], \\
\frac{dT}{dz} &= \Omega + \sigma \mathrm{Re}\left[\int \bar{v}_T^\dagger \exp(-i\Theta)\, n(z,t)\, dt\right], \\
\frac{d\Omega}{dz} &= \sigma \mathrm{Re}\left[\int \bar{v}_\Omega^\dagger \exp(-i\Theta)\, n(z,t)\, dt\right], \\
\frac{d\Phi}{dz} &= \frac{A^2-\Omega^2}{2} + \sigma \mathrm{Re}\left[\int \left(\bar{v}_\Phi^\dagger - T\bar{v}_\Omega^\dagger\right) \exp(-i\Theta)\, n(z,t)\, dt\right],
\end{aligned} \tag{2.25}$$

which are accurate up to $O(\sigma)$.

These equations constitute a system of stochastic ODEs that establish a projection from an infinite-dimensional noise process ($n(z,t)$) to a random walk for the four soliton parameters. Using the noise model in Equation (2.2) for discrete amplification, the stochasticity appears as discrete random jumps in the soliton parameters which take the form of projections between the noise realizations at each amplifier and the four adjoint modes of the linearization,

$$\frac{dA}{dz} = \sum_{k=1}^{N_a} \sigma \mathrm{Re}\left[\int \bar{v}_A^\dagger \exp(-i\Theta)\, n_k(t)\, dt\right] \delta(z - kz_a), \tag{2.26a}$$

$$\frac{dT}{dz} = \Omega + \sum_{k=1}^{N_a} \sigma \mathrm{Re}\left[\int \bar{v}_T^\dagger \exp(-i\Theta)\, n_k(t)\, dt\right] \delta(z - kz_a), \tag{2.26b}$$

$$\frac{d\Omega}{dz} = \sum_{k=1}^{N_a} \sigma \mathrm{Re}\left[\int \bar{v}_\Omega^\dagger \exp(-i\Theta)\, n_k(t)\, dt\right] \delta(z - kz_a), \tag{2.26c}$$

and

$$\frac{d\Phi}{dz} = \frac{A^2-\Omega^2}{2} + \sum_{k=1}^{N_a} \sigma \mathrm{Re}\left[\int \left(\bar{v}_\Phi^\dagger - T\bar{v}_\Omega^\dagger\right) \exp(-i\Theta)\, n_k(t)\, dt\right] \delta(z - kz_a). \tag{2.26d}$$

Because the noise is added at discrete points, these equations can be formally integrated giving

$$A(z) = A(0) + \sum_{k=1}^{N_a} \Delta A_k \ \mathrm{H}(z - kz_a)\,, \tag{2.27a}$$

$$\begin{aligned} T(z) = T(0) + \Omega(0)z &+ \sum_{k=1}^{N_a} \Delta\Omega_k \,(z - kz_a)\ \mathrm{H}(z - kz_a) \\ &+ \sum_{k=1}^{N_a} \Delta T_k \ \mathrm{H}(z - kz_a)\,, \end{aligned} \tag{2.27b}$$

$$\Omega(z) = \Omega(0) + \sum_{k=1}^{N_a} \Delta\Omega_k \ \mathrm{H}(z - kz_a)\,, \tag{2.27c}$$

and

$$\begin{aligned} \Phi(z) = \Phi(0) &+ \frac{A(0)^2}{2} z + \sum_{k=1}^{N_a} A(0)\Delta A_k \,(z - kz_a)\,\mathrm{H}(z - kz_a) \\ &+ \frac{1}{2}\sum_{k=1}^{N_a} (\Delta A_k)^2 \,(z - kz_a)\,\mathrm{H}(z - kz_a) \\ &+ \frac{1}{2}\sum_{k=1}^{N_a}\sum_{j=1}^{k-1} \Delta A_k \Delta A_j \,(z - kz_a)\,\mathrm{H}(z - kz_a) \\ &+ \frac{1}{2}\sum_{k=1}^{N_a}\sum_{j=k+1}^{N_a} \Delta A_k \Delta A_j \,(z - jz_a)\,\mathrm{H}(z - jz_a) \\ &- \frac{\Omega(0)^2}{2} z - \sum_{k=1}^{N_a} \Omega(0)\Delta\Omega_k \,(z - kz_a)\,\mathrm{H}(z - kz_a) \\ &- \frac{1}{2}\sum_{k=1}^{N_a} (\Delta\Omega_k)^2 \,(z - kz_a)\,\mathrm{H}(z - jz_a) \\ &- \frac{1}{2}\sum_{k=1}^{N_a}\sum_{j=1}^{k-1} \Delta\Omega_k \Delta\Omega_j \,(z - kz_a)\,\mathrm{H}(z - kz_a) \\ &- \frac{1}{2}\sum_{k=1}^{N_a}\sum_{j=k+1}^{N_a} \Delta\Omega_k \Delta\Omega_j \,(z - jz_a)\,\mathrm{H}(z - jz_a) \\ &+ \sum_{k=1}^{N_a} \Delta\Phi_k \ \mathrm{H}(z - kz_a)\,, \end{aligned} \tag{2.27d}$$

where

$$\begin{aligned}
\Delta A_k &= \sigma \mathrm{Re}\left[\int \bar{v}_A^\dagger(kz_a,t)\exp(-i\Theta(kz_a))\,n_k(t)\,dt\right],\\
\Delta \Omega_k &= \sigma \mathrm{Re}\left[\int \bar{v}_\Omega^\dagger(kz_a,t)\exp(-i\Theta(kz_a))\,n_k(t)\,dt\right],\\
\Delta T_k &= \sigma \mathrm{Re}\left[\int \bar{v}_T^\dagger(kz_a,t)\exp(-i\Theta(kz_a))\,n_k(t)\,dt\right],\\
\Delta \Phi_k &= \sigma \mathrm{Re}\left[\int \left(\bar{v}_\Phi^\dagger(kz_a,t) - T(kz_a)\bar{v}_\Omega^\dagger(kz_a,t)\right)\exp(-i\Theta(kz_a))\,n_k(t)\,dt\right].
\end{aligned} \tag{2.28}$$

The stochastic jump terms, ΔX_k, are mean zero and, assuming that the parameter values immediately after the kth amplifier are known, the variances at the $k+1$ amplifier are given by

$$\begin{aligned}
\mathbb{E}\left[(\Delta A_{k+1})^2\right] &= \frac{\sigma^2}{2}\|v_A^\dagger(kz_a)\|^2 = \sigma^2 A_k,\\
\mathbb{E}\left[(\Delta T_{k+1})^2\right] &= \frac{\sigma^2}{2}\|v_T^\dagger(kz_a)\|^2 = \frac{\sigma^2\pi^2}{12A_k^3},\\
\mathbb{E}\left[(\Delta \Omega_{k+1})^2\right] &= \frac{\sigma^2}{2}\|v_\Omega^\dagger(kz_a)\|^2 = \frac{\sigma^2 A_k}{3},\\
\mathbb{E}\left[(\Delta \Phi_{k+1})^2\right] &= \frac{\sigma^2}{2}\|v_\Phi^\dagger(kz_a) - T_k v_\Omega^\dagger(kz_a)\|^2\\
&= \sigma^2\frac{12\left(1+(T_k+\Omega_k z_a)^2 A_k^2\right)+\pi^2}{36A_k},
\end{aligned} \tag{2.29}$$

where the notation $X_k = X(kz_a)$ has been used for brevity.

2.2 The General Importance Sampled Monte Carlo Method

Here, the ISMC method is presented in general terms. Let x be a random variable (RV) which takes on values from set Ω and is distributed according to the PDF $p(x)$. In addition, suppose $y(x)$ is a RV which is dependent on the outcomes of x and that the set R represents the values of $y(x)$ that correspond to a specific event of interest. The probability of y taking on a value in a region R, is given by the integral,

$$P = \mathbb{P}[y \in R] = \int_{\Omega^*} p(x)\,dx = \int_\Omega I(y(x))p(x)dx = \mathbb{E}[I(y(x))]\,, \tag{2.30}$$

where Ω^* is the subset of Ω for which $y(x) \in R$ ($x \in \Omega^* \iff y(x) \in R$), and $I(y(x))$ is an indicator function which is 1 for $y(x) \in R$ and 0 otherwise. A standard Monte Carlo scheme would approximate this integral by drawing samples x_k according to the distribution $p(x)$ and forming the estimator

$$\hat{P}_{\mathrm{MC}} = \frac{1}{N} \sum_{k=1}^{N} I(y(x_k)). \tag{2.31}$$

When evaluated at the samples x_k, the indicator function $I(y(x_k))$ forms a Bernoulli RV, where $\mathbb{P}[I(y(x_k)) = 1] = P$ and $\mathbb{P}[I(y(x_k)) = 0] = 1 - P$, and thus the sum in Equation (2.31) forms a binomial RV with parameters (N, P), implying that the estimator $\hat{P}_{\mathrm{MC}}$ has variance given by

$$\mathbb{V}\left[\hat{P}_{\mathrm{MC}}\right] = \frac{P(1-P)}{N}, \tag{2.32}$$

with a relative error (coefficient of variation) of,

$$\mathbb{C}_{\mathrm{var}}\left[\hat{P}_{\mathrm{MC}}\right] = \frac{\sqrt{\mathbb{V}\left[\hat{P}_{\mathrm{MC}}\right]}}{\mathbb{E}\left[\hat{P}_{\mathrm{MC}}\right]} = \frac{\sqrt{1-P}}{\sqrt{NP}}. \tag{2.33}$$

Ensuring that Equation (2.31) is an accurate estimator for the true probability requires that $\mathbb{C}_{\mathrm{var}}\left[\hat{P}_{\mathrm{MC}}\right] \ll 1$. However, if the set R in Equation (2.30) is such that $P \ll 1$ (i.e., $y(x) \in R$ is a *rare event*), this requirement can be approximated by $N \gg 1/P$, which implies that a very large (and often unattainable) number of Monte Carlo runs is required to obtain a suitably accurate estimate.

To fully understand the reason why the standard MC method requires such a large number of samples, it is best to think in terms of the set Ω^* in the sample space which corresponds to the set R representing the (rare) event of interest. The MC method is designed to calculate the measure of the set Ω^* under the distribution p, i.e.,

$$\int_{\Omega^*} p(x)\, dx, \tag{2.34}$$

by randomly sampling under the p distribution and forming the estimator given in Equation (2.31). Since the set R represents a rare event, however, the measure of the set Ω^* under the p distribution is exceedingly small, thus requiring a large number of samples to resolve its measure accurately. In cases such as this, IS can be used to greatly improve the efficiency of the standard MC method, by drawing samples not from the distribution p, but rather from a new distribution p^*, which is carefully chosen to assign a much higher measure to set Ω^*, i.e.,

$$\int_{\Omega^*} p(x)\,dx \ll \int_{\Omega^*} p^*(x)\,dx. \tag{2.35}$$

Note that this requirement is commonly met in application by enforcing the stronger condition of

$$p(x) \ll p^*(x) \text{ for } x \in \Omega^*. \tag{2.36}$$

Of course, sampling according to the new distribution introduces a bias which must be accounted for by discounting each sample through the likelihood ratio $r(x) = p(x)/p^*(x)$. Using this approach, Equation (2.30) takes the equivalent form

$$\mathbb{P}[y \in R] = \int_\Omega I(y(x))\frac{p(x)}{p^*(x)}p^*(x)dx = \int_\Omega I(y(x))r(x)p^*(x)dx = \mathbb{E}^*\left[I(y(x))r(x)\right], \tag{2.37}$$

which implies a new ISMC estimator of the form

$$\hat{P}_{\text{IS}} = \frac{1}{N}\sum_{k=1}^{N} I(y(x_k^*))l(x_k^*), \tag{2.38}$$

with variance

$$\mathbb{V}^*\left[\hat{P}_{\text{IS}}\right] = \frac{\mathbb{E}^*\left[I(y(x))r^2(x)\right] - P^2}{N} \approx \frac{\mathbb{E}\left[I(y(x))r(x)\right])}{N}. \tag{2.39}$$

The expectation in Equation (2.37), the samples in Equation (2.38) and the variance in (2.39) are all taken with respect to the new distribution $p^*(x)$ as indicated by the

asterisk. Since the biasing distribution $p^*(x)$ was chosen with the aim of increasing the measure of the set Ω^* (i.e. $r(x) \ll 1$ for $x \in \Omega^*$), Equation (2.39) shows that a lower variance and relative error of the estimator can be obtained using the same number of samples under the IS implementation in Equation (2.38).

It should be stressed that the improvement found in the IS implementation in Equation (2.38) is *only* present if the biasing distribution $p^*(x)$ is chosen such that Equation (2.35) is satisfied. If the biasing distribution is poorly chosen, i.e.,

$$\int_{\Omega^*} p^*(x)\,dx < \int_{\Omega^*} p(x)\,dx, \tag{2.40}$$

the error of the new method can be orders of magnitude *higher* than that of the standard MC method using an equivalent number of samples. The IS provides no indication as to the form of biasing distribution other than (2.35). However, from Equations (2.34), (2.35) and a bit of intuition, it is seen that the optimal choice of $p^*(x)$ must satisfy

$$\int_{\Omega^*} p^*(x)\,dx = 1. \tag{2.41}$$

In addition, if the set Ω^* is high dimensional, there could be large variations in $p(x)$ with $x \in \Omega^*$. In such cases, it is reasonable to assume that the measure of the set Ω^* under $p(z)$ will be dominated by the contributions around the points that maximize $p(x)$. This fact discourages the use of uniform biasing distributions such as

$$p^*(x) = \begin{cases} \frac{1}{V_{\Omega^*}} & x \in \Omega^*, \\ 0 & x \notin \Omega^*, \end{cases} \quad \text{where } V_{\Omega^*} = \int_{\Omega^*} dx. \tag{2.42}$$

In fact, it can be shown that the optimal choice for $p^*(x)$ is $p(x|x \in \Omega^*)$, i.e., the original density conditioned on the event of interest [14]. Unfortunately, this distribution requires the very normalization constant one is trying to compute. Further,

even the simpler requirement of Equation (2.36) is hard to meet in application when an inadequate amount of information is avaliable concerning set Ω^*.

2.2.1 Multiple Importance Sampling

When computing the probability measure of sets containing multiple critical points of p, or when reconstructing entire probability distributions rather than measures associated with compact sets, it is often convenient to use multiple biased noise distributions whose estimators must then be resummed in an unbiased way. For example, if the noise configuration space is divided into the non-overlapping bins of a histogram, the ISMC method above would have to be applied across many sets, i.e., R_q for $q = 1, \ldots, Q$, the union of which gives all possible values of interest that the RV $y(x)$ can achieve. The ISMC method applied to each set individually would result in a large number of discarded runs that fall outside of the target set. A more efficient approach would be to aggregate the total number of runs across multiple biasing distributions, so that every run is utilized in the calculation of a probability. This can be done through an approach known as multiple importance sampling [24, 43, 44], which assigns a weight, $w_q(x)$, to each biasing distribution, resulting in a multi-distribution importance-sampled estimator of the form,

$$\hat{P}_{\text{MIS}} = \sum_{q=1}^{Q} \frac{1}{N_q} \sum_{k=1}^{N_q} w_q(x^*_{q,k}) I(y(x^*_{q,k})) r_q(x^*_{q,k}), \tag{2.43}$$

where Q is the total number of biasing distributions (each distribution labeled p^*_q), N_q is the number of samples drawn from $p^*_q(x)$, $x^*_{q,k}$ is the kth such sample, and $r_q(x) = p(x)/p^*_q(x)$. In addition, an unbiased estimator for the variance of P_{MIS} is

given by [24]

$$\mathbb{V}^*\left[\hat{P}_{\text{MIS}}\right] = \sum_{q=1}^{Q} \frac{1}{N_q - 1} \left[\frac{1}{N_q} \sum_{k=1}^{N_q} w_q^2(x_{q,k}^*) I(y(x_{q,k}^*)) r_q^2(x_{q,k}^*) - \left(\frac{1}{N_q} \sum_{k=1}^{N_q} w_q(x_{q,k}^*) I(y(x_{q,k}^*)) r_q(x_{q,k}^*) \right)^2 \right]. \tag{2.44}$$

There are several possible strategies for selecting the weights; however, in order to ensure that the estimator $\hat{P}_{\text{MIS}}$ stays unbiased, it must be that $\sum_{q=1}^{Q} w_q(x) = 1$ for all x. It is illustrative to consider two obvious choices which represent opposite ends of the possible spectrum of weighting schemes. The first is to take $w_q(x) = 1$ for all x, but only for $q = q^*$, where q^* is the index for the distribution chosen to explicitly target the outcome $I(y(x^*)) = 1$. This is exactly equivalent to the case of applying IS to the individual sets, which is inefficient for reasons already discussed. The second is to take $w_q(x) = 1/Q$ for all x and q, which is equivalent to weighting each biasing distribution equally regardless of the noise realization. This assumes that each distribution is equally good in resolving different areas of the sample space, which is obviously not the case since each biasing distribution is chosen to concentrate samples around a specific set.

A compromise between the two weighting schemes above, and the one implemented here, is the *balance heuristic* [44], given by

$$w_q(x) = \frac{N_q p_q^*(x)}{\sum_{q'=1}^{Q} N_{q'} p_{q'}^*(x)}. \tag{2.45}$$

Note that $N_q p_q^*(x)$ is proportional to the expected number of outcomes resulting in $I(y(x)) = 1$, where the samples are drawn from the qth distribution. Thus, the weight of a sample with the balance heuristic is given by the likelihood of realizing that sample with the qth distribution relative to the total likelihood of realizing the same sample with all distributions. Use of the balance heuristic assumes that the

weights $w_q(x)$ are trivial to compute, which is not guaranteed to hold in general, although does hold here.

2.3 ISMC Method for Rare Events in Soliton Communication Systems

In the context of soliton based optical communication systems, rare events are associated with large, noise-induced perturbations of the pulse parameters, which occur with exceedingly low probabilities. The ISMC method is effective in resolving these probabilities but, as discussed above, it requires knowledge of the specific noise realizations, i.e., points in sample space, that produce the rare event of interest, i.e., a large change in the pulse parameters from their deterministic values. Furthermore, it is not enough to simply know these points; one really needs knowledge of the most probable points, i.e., those noise realizations that produce prescribed parameter changes with *highest probability*, as these will dominate any approximation of the measure of this set of points, i.e., the probability of a large deviation.

Finding the specific noise realizations that induce a parameter to significantly deviate from its mean value is the central problem in applying the ISMC method, since these are the points around which one must form the biasing distributions. These points could eventually be found by conducting numerous standard MC runs of the stochastic NLSE given in Equation (2.1) and recording the noise realizations leading to the desired deviation; however, to find the biasing points this way would require as much computational cost as the original MC method for finding the desired probabilities.

Fortunately, an alternative method for finding these points is found in SPT, which provides a low dimensional system of stochastic ODEs for the approximate evolution of *just* the soliton parameters as opposed to the evolution of an entire pulse. Thus, SPT provides the means of finding the most likely noise realizations around which to form the biasing distributions essential to the ISMC method.

2.3.1 Numerical Implementation and the Biasing Scheme

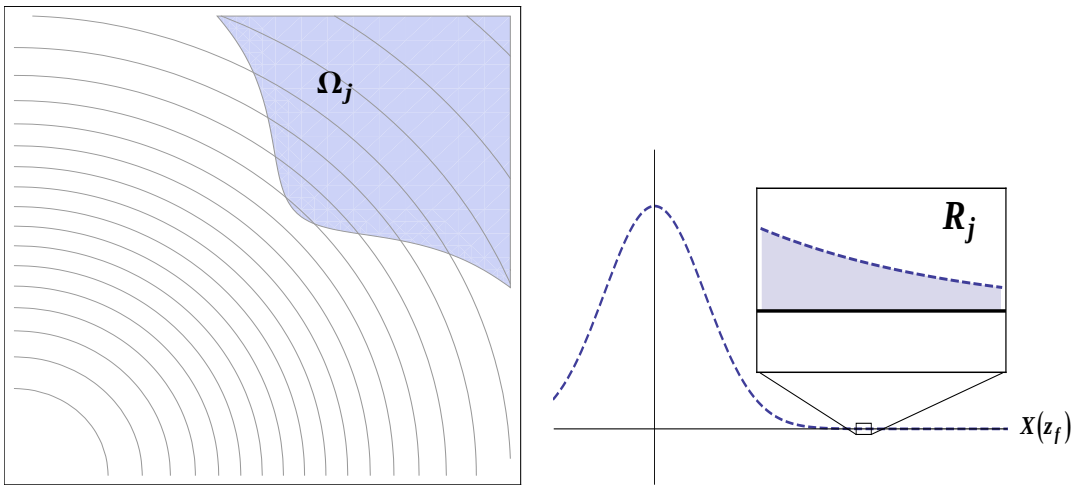
The implementation of the ISMC method introduced above involves numerically solving a discretized version of Equation (2.1), where the initial solution is a discretized soliton with known parameter values. For the results presented below, this was done using a pseudo-spectral method, where the z evolution was calculated using the fourth order Runge-Kutta method in conjunction with an integrating factor. Variations of this method are used in many of the subsequent chapters of this document and it is therefore presented in detail in Appendix B. Since the stochastic forcing is modeled at discrete locations along the fiber, it is included by periodically adding a complex-valued random vector $\boldsymbol{n}_k = \left(n_{1,k}, \ldots, n_{N_p,k}\right)^T$, after evolving the solution for the number of steps representing the non-dimensional amplifier spacing. Here N_p is the number of computational points in the t domain (or equivalently the number of Fourier modes), k indexes the amplification points and $n_{j,k} = x_{j,k} + iy_{j,k}$, where $x_{j,k}$ and $y_{j,k}$ are independent and identically distributed (i.i.d.) normal RVs such that $\mathbb{E}\left[x_{j,k}^2\right] = \mathbb{E}\left[y_{j,k}^2\right] = \sigma_n^2$ with $\sigma_n^2 = \sigma^2/(2\Delta t)$. Note that the noise strength, σ^2, is adjusted so that the power spectral density (PSD) of the discretized noise matches that of the non-truncated idealized white noise. Recalling that the noise vectors are also independently distributed among amplifiers, the entire noise realization (point in sample space), for any particular MC run can be represented by the matrix $\boldsymbol{N} = [\boldsymbol{n}_1 \ldots \boldsymbol{n}_{N_a}]$, which has a distribution that is composed of the product of distributions for the individual noise vectors,

$$p_{\text{tot}}(\boldsymbol{N}) = \prod_{k=1}^{N_a} p(\boldsymbol{n}_k), \text{ where } p(\boldsymbol{n}_k) = \frac{1}{\left(2\pi\sigma_n^2\right)^{N_p}} \exp\left(-\frac{\left(\bar{\boldsymbol{n}}_k^T \boldsymbol{n}_k\right)}{2\sigma_n^2}\right). \tag{2.46}$$

The RVs of interest are the pulse parameters evaluated at the end of the transmission line, i.e., $X(z_{\text{f}}; \boldsymbol{N})$ for $X = A$, T, Ω and Φ[2], which are dependent on the

[2]From this point on, X will be used to denote the four soliton parameters without explicitly indicating "for $X = A$, T, Ω and Φ".

total noise added during propagation. The goal of the ISMC method is to construct the PDFs of these RVs down to the tails, thus approximating the probability of large deviations of a given pulse parameter from its mean value. The mechanics of constructing a PDF using the ISMC method is the same as an ordinary MC method. One partitions the domain of final values for each parameter into non-overlapping intervals or *bins*, represented by R_j for $j = 1, \ldots, N_b$, where N_b is the number of partitions. Each bin R_j is associated with a subset of the sample space, Ω_j, which represents all noise realizations that lead to a final parameter value in the interval R_j, i.e., $\boldsymbol{N} \in \Omega_j \implies X(z_\mathrm{f}; \boldsymbol{N}) \in R_j$. Figure 2.2 shows a 2D schematic of this partitioning, with Figure 2.2(a) showing the level sets of $p(\boldsymbol{N})$ and an example set Ω_j containing all noise realizations leading to a final parameter value in the interval R_j from the example PDF in Figure 2.2(b). To construct an approximation to the



(a) Example Sample Space (b) Example PDF

Figure 2.2 Left: Illustration of an example sample space for the MC method. Right: Example probability density function for the MC method.

total PDF of $X(z_\mathrm{f}; \boldsymbol{N})$, the PDF over each interval is approximated by a constant value, which under the normal MC implementation, is given by the typical unbiased MC estimator given in Equation (2.31). Of course, the standard MC method would

rarely sample points from the set Ω and is thus unable to resolve the tails of the PDF, prompting the use of the ISMC method.

To implement the ISMC method, a biasing distribution (denoted by p^*_{tot}) must be chosen such that sampling under p^*_{tot} increases the frequency at which noise realizations are drawn from set Ω_j, i.e., noise realizations leading to a final parameter value in the interval R_j. As discussed in Section 2.2, this biasing distribution should also reflect the fact that certain points contained in Ω_j contribute much more to the measure of set Ω_j under the original distribution $p_{\text{tot}}(\boldsymbol{N})$. Here, the biasing distribution is simply taken as a mean-shifted version of the original distribution, i.e., $p^*_{\text{tot}}(\boldsymbol{N}^*) = p_{\text{tot}}(\boldsymbol{N}^* - \boldsymbol{B}_j)$, where the mean (labeled $\boldsymbol{B}_j$) is the deterministic point that maximizes the original distribution, subject to the constraint that $\mathbb{E}[X(z_{\text{f}}; \boldsymbol{N} + \boldsymbol{B}_j)] = X_j$, where X_j is the midpoint of interval R_j[3]. The constraint that $\mathbb{E}[X(z_{\text{f}}; \boldsymbol{N} + \boldsymbol{B}_j)] = X_j$, as opposed to $\mathbb{E}[X(z_{\text{f}}; \boldsymbol{N} + \boldsymbol{B}_j)] \in R_j$, is one of convenience since, as will be shown below, this condition is associated with an end point of a boundary value problem (BVP). However, this does restrict the possible location of $\boldsymbol{B}_j$ to a subset, $\hat{\Omega}_j \subset \Omega_j$, of the total noise realizations associated with set R_j. By assuming that the biasing distribution is a mean-shifted version of the original distribution, the only remaining unknown is the biasing point $\boldsymbol{B}_j$. After the biasing point is found, the biasing distribution is given by

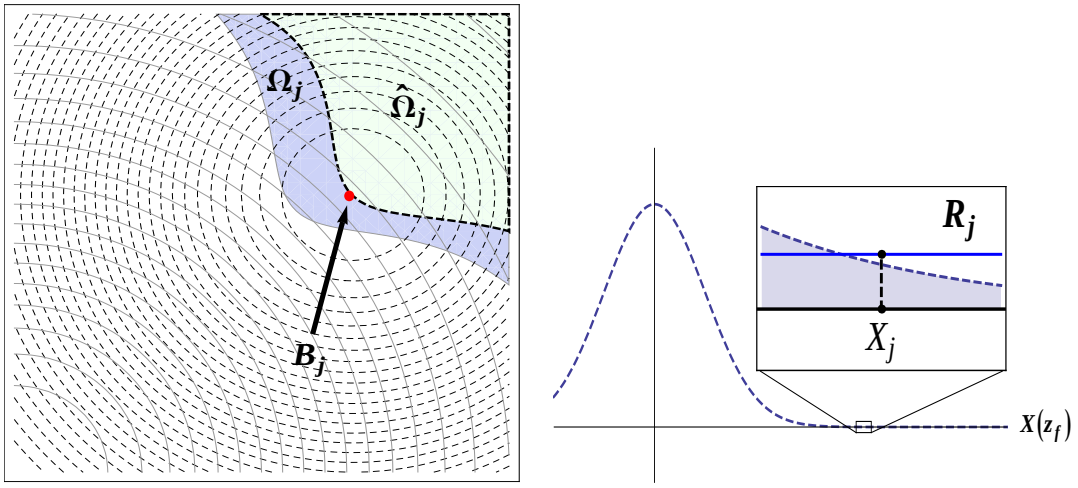
$$p^*_{\text{tot}}(\boldsymbol{N}^*) = p_{\text{tot}}(\boldsymbol{N}^* - \boldsymbol{B}_j) = p_{\text{tot}}(\boldsymbol{N}) \quad \text{with} \quad \boldsymbol{N}^* = \boldsymbol{N} + \boldsymbol{B}_j, \tag{2.47}$$

which gives a likelihood ratio of

$$r(\boldsymbol{N}^*) = \frac{p_{\text{tot}}(\boldsymbol{N} + \boldsymbol{B}_j)}{p_{\text{tot}}(\boldsymbol{N})} = \prod_{k=1}^{N_a} \frac{p(\boldsymbol{n}^*_k)}{p(\boldsymbol{n}_k)}, \tag{2.48}$$

[3]Note that this scheme implicitly assumes that only one maximum $\boldsymbol{B}_j$ exists, which will be shown to be true in the actual implementation.

where $\boldsymbol{n}_k^* = \boldsymbol{n}_k + \boldsymbol{b}_k$ is the kth vector of noise biased around the kth column of $\boldsymbol{B}_j$. Figure 2.3 shows a 2D schematic of the contour lines of an example biasing distribution and the point $\boldsymbol{B}_j$ around which the biasing distribution is centered.



(a) Example Sample Space With Biasing (b) Example PDF With Biasing

Figure 2.3 Left: The contour lines of the original and biasing distributions are give by the solid and dashed curves, respectively. The biasing point is given by the red dot. Right: Corresponding probability density function associated with the ISMC method.

The biasing point $\boldsymbol{B}_j$ is defined with respect to two constraints: (i) it maximizes $p_{\text{tot}}(\boldsymbol{N})$ over the set $\hat{\Omega}_j$ and (ii) it satisfies $\mathbb{E}[X(z_{\text{f}}; \boldsymbol{N} + \boldsymbol{B}_j)] = X_j$. Together, these constraints ensure that, on average, the dynamics of the parameters follow an *optimal path* through four dimensional parameter space, with a particular parameter $X(z)$ starting and ending at the predetermined values $X(0)$ and X_j, respectively. Thus, determining the biasing point $\boldsymbol{B}_j$ is equivalent to determining the biasing vectors $\boldsymbol{b}_k$, which when added to the propagating pulse at each corresponding amplifier (along with the noise realization $\boldsymbol{n}_k$) induce the soliton parameters to change in accordance with the optimal path.

The approach adopted here for finding the biasing point $\boldsymbol{B}_j$ is based on a linear expansion for each biasing vector $\boldsymbol{b}_k$ of the form

$$\boldsymbol{b}_k = \Delta_{A,k}\boldsymbol{b}_{A,k} + \Delta_{\Omega,k}\boldsymbol{b}_{\Omega,k} + \Delta_{T,k}\boldsymbol{b}_{T,k} + \Delta_{\Phi,k}\boldsymbol{b}_{\Phi,k}, \tag{2.49}$$

where the vectors $\boldsymbol{b}_{X,k}$ are derived such that the addition of $\Delta_{X,k}\boldsymbol{b}_{X,k}$ to the propagating pulse at amplifier k produces a $\Delta_{X,k}$ change in parameter X. This approach effectively decomposes the task of finding the vectors $\boldsymbol{b}_k$, and thus the point $\boldsymbol{B}_j$, into two separate problems. The first determines the vectors $\boldsymbol{b}_{X,k}$ that produce unit changes to each of the soliton parameters, while the second determines the weights $\Delta_{X,k}$ for $k = 1, \ldots, N_a$ that drive the parameter $X(z)$ to the targeted final value X_j. Both of these problems are subject to the constraint of maximizing the probability, which can be incorporated by reformulating them in terms of constrained optimizations with use of SPT.

The first optimization problem is to find the most likely biasing vector, which when added to a propagating pulse, produces an instantaneous unit change in a specific soliton parameter. As Equation (2.46) indicates, the noise added at each amplifier consists of a sequence of complex random variables with real and imaginary components that are normally distributed and independent. The likelihood of realizing a particular noise vector can be found by evaluating the Gaussian PDF given in Equation (2.46), and thus, the biasing vectors of maximum likelihood are those that maximize a Gaussian function, or equivalently, those with smallest l^2-norm,

$$\text{argmax}_{\boldsymbol{b}_X}\Big(\exp\Big(-\bar{\boldsymbol{b}}_X^T\boldsymbol{b}_X\Big)\Big) = \text{argmin}_{\boldsymbol{b}_X}\Big(\bar{\boldsymbol{b}}_X^T\boldsymbol{b}_X\Big). \tag{2.50}$$

In the continuous domains of t and z, this corresponds to seeking a biasing function $f_X(z,t)$ with minimum L^2-norm,

$$\|f_X(z,t)\|^2 = \int |f_X(z,t)|^2\, dt. \tag{2.51}$$

Of course, this quantity must be minimized subject to the constraint that the biasing function $f_X(z,t)$ delivers a unit change in parameter X. Recalling that the equations of SPT given in Equation (2.25) have projection terms that represent the (linear) changes in the pulse parameters from the addition of a noise vector to the propagating pulse, this constraint can be approximated by

$$\mathrm{Re}\left[\int \left(\bar{v}_X^\dagger(z,t) - T\bar{v}_\Omega^\dagger(z,t)\delta_{X\Phi}\right) \exp(-i\Theta)\, f_X(z,t)\, dt\right] = 1, \tag{2.52}$$

where $v_X^\dagger$ is the appropriate *adjoint* linear mode corresponding to changes in parameter X, and δ_{kj} is the Kronecker delta. Expressing this constrained minimization problem in Lagrange multiplier form gives the action

$$\begin{aligned} M(f_X(z,t), \bar{f}_X(z,t)) &= \|f_X(z,t)\|^2 \\ &+ \lambda\left[\left(\int \left(\bar{v}_X^\dagger(z,t) - T\bar{v}_\Omega^\dagger(z,t)\delta_{X\Phi}\right) \exp(-i\Theta)\, f_X(z,t)\, dt\right.\right. \\ &\left.\left. + \int \left(v_X^\dagger(z,t) - Tv_\Omega^\dagger(z,t)\delta_{X\Phi}\right) \exp(i\Theta)\, \bar{f}_X(z,t)\, dt\right) - 2\right], \end{aligned} \tag{2.53}$$

which is stationary at the solution

$$f_X(z,t) = \frac{\left(v_X^\dagger(z,t) - Tv_\Omega^\dagger(z,t)\delta_{X\Phi}\right)}{\|v_X^\dagger(z,t) - Tv_\Omega^\dagger(z,t)\delta_{X\Phi}\|^2} \exp(i\Theta)\,. \tag{2.54}$$

It is interesting to note that Equation (2.54) is contrary to intuition, in that it indicates that the most probable way to realize a specific change in a parameter through an additive perturbation occurs by adding a perturbation proportional to the re-normalized *adjoint* mode $v_X^\dagger(z,t)/\|v_X^\dagger(z,t)\|^2$ as opposed to a perturbation proportional to the mode $v_X(z,t)$, which represents the derivative of the pulse with respect to the parameter X and thus produces the desired parameter change without a radiative contribution. However, by simple comparison, it is seen that the L^2-norm of re-normalized *adjoint* modes given by Equation (2.54) are always larger in value than the L^2-norm of corresponding linear mode $v_X(z,t)$. This is a subtle detail that is hard

to determine a priori and is only provided here as a result of using the information found in SPT. Finally, note that since the biasing functions given in Equation (2.54) depend on the adjoint modes $v_X^\dagger(z,t)$, which themselves depend on the value of the soliton parameters (c.f., Equations (2.18) and (2.14)), the calculation of the modes directly depends on the ability to recover accurate estimates for the soliton parameters from a noisy soliton. For solitons of the NLSE, three methods for accomplishing this have been explored in the literature: (i) integral representations for the parameters, i.e., moments, (ii) an iterative projection method [45] and (iii) a solving a numerical version of the Zakharov-Shabat eigenvalue problem [46]. All three of these methods are detailed in Appendix B.

Now that the most likely functional forms (in t) imparting unit changes to soliton parameters are known at any point z, it remains to solve the biasing problem across all amplifiers (in z). In the discrete form, this amounts to determine the weights $\Delta_{X,k}$ that will scale each biasing vector to ensure that the mean evolution path of the parameters coincides with the most likely path between the endpoint constraints associated with a particular biasing distribution. In continuous form, these weights become functions of a continuous variable z, i.e., $\Delta_X(z)$, and thus the continuous representation of the biasing vectors given in Equation (2.49) take the form

$$f(z,t) = \Delta_A(z)f_A(z,t) + \Delta_\Omega(z)f_\Omega(z,t) + \Delta_T(z)f_T(z,t) + \Delta_\Phi(z)f_\Phi(z,t). \tag{2.55}$$

Just as in the above analysis, the likelihood of realizing this biasing vector at any individual fixed point z is given by the Gaussian PDF in Equation (2.46) and thus, the most likely biasing vectors are those with smallest L^2-norm. However, now the biasing in z must also be accounted for, implying that the appropriate quantity to

minimize is the *cumulative* L^2-norm given by,

$$S = \int_0^{z_\mathrm{f}} \|f(z,t)\|^2 \, dz = \int_0^{z_\mathrm{f}} \sum_Y \left(\Delta_Y(z)\right)^2 \|f_Y(z,t)\|^2 + 2\Delta_\Phi(z)\Delta_\Omega(z) \left\langle f_\Phi(z,t), f_\Omega(z,t)\right\rangle \, dz, \tag{2.56}$$

where $\sum_Y$ represents the sum over all four parameters.

The functional in Equation (2.56) must be minimized subject to the constraint $X(z_\mathrm{f}) = X_\mathrm{f}$, where X represents the parameter of interest and X_f is the targeted final parameter value. Before showing how this is done, first note that the four weighting functions $\Delta_Y(z)$ can be related to the optimal path through parameter space by finding the mean evolution of the stochastic ODEs given by SPT in Equations (2.25) after the biasing has been applied. By replacing the mean zero noise $\sigma n(z,t)$ by the biased noise $\sigma n(z,t) + \sum_Y \Delta_Y(z) f_Y(z,t)$, these equations take the form,

$$\begin{aligned} &\dot{A} = n_A(z) + \Delta_A(z), \quad \dot{\Omega} = n_\Omega(z) + \Delta_\Omega(z) + \Delta_\Omega(z)\Delta_\Phi(z)\frac{T\|v_\Omega\|^2}{\|v_\Phi - Tv_\Omega\|^2}, \\ &\dot{T} = \Omega + n_T(z) + \Delta_T(z), \quad \dot{\Phi} = \frac{A^2 - \Omega^2}{2} + n_\Phi(z) + \Delta_\Phi(z) \end{aligned} \tag{2.57}$$

where $n_X(z) = \left\langle \left(v_X^\dagger - T v_\Omega^\dagger \delta_{\Phi X} \right) \exp(i\Theta) \, , \sigma n(z,t) \right\rangle$ and the shorthand notation $\dot{X} = dX/dz$ is used to notate differentiation in z. Taking the expectation of these equations (noting that $\mathbb{E}[n_X(z)] = 0$) gives the deterministic equations relating the weights to the optimal path taken through parameter space,

$$\dot{A}_\mathrm{opt} = \Delta_A(z), \tag{2.58a}$$

$$\dot{\Omega}_\mathrm{opt} = \Delta_\Omega(z) + \Delta_\Omega(z)\Delta_\Phi(z)\frac{T_\mathrm{opt}A_\mathrm{opt}^2}{1 + \frac{\pi^2}{12} + T_\mathrm{opt}^2 A_\mathrm{opt}^2}, \tag{2.58b}$$

$$\dot{T}_\mathrm{opt} = \Omega_\mathrm{opt} + \Delta_T(z), \tag{2.58c}$$

$$\dot{\Phi}_\mathrm{opt} = \frac{A_\mathrm{opt}^2 - \Omega_\mathrm{opt}^2}{2} + \Delta_\Phi(z), \tag{2.58d}$$

Thus, the functional S given in Equation (2.56) is seen to depend on the optimal path each parameter takes through parameter space. Finally, noting that each parameter begins with a deterministic value $X_{\text{opt}}(0) = X_{\text{i}}$, it can be shown that (in general) the minimization of S in Equation (2.56) is equivalent to a four dimensional, second order boundary value problem, some relevant solutions of which are detailed in the following sections.

Finally, it should be noted that, although the ISMC method presented here focuses on the final distribution of individual pulse parameters, this approach is directly applicable for computing the measure of more complex regions of four-dimensional parameter space. Furthermore, even the single-parameter problems discussed here often use optimal parameter paths that involve changes in parameters *other* than the one under consideration, the implications of which will become clear in the following sections.

Biasing the Amplitude. The calculation for the biasing weights is easiest for the amplitude parameter, since by Equation (2.58a), changes to the other parameters have no effect on the amplitude evolution which implies that $\Delta_T(z,t)$, $\Delta_\Omega(z,t)$ and $\Delta_\Phi(z,t)$ can be neglected. This leaves only the biasing vector for the amplitude, $f_A(z,t)$, in the cumulative L^2-norm, which now takes the form

$$S_A(A_{\text{opt}}, \dot{A}_{\text{opt}}) = \int_0^{z_{\text{f}}} (\Delta_A(z))^2 \|f_A(z,t)\|^2 \, dz = \int_0^{z_{\text{f}}} \frac{\dot{A}_{\text{opt}}^2}{2A_{\text{opt}}} \, dz, \tag{2.59}$$

where Equation (2.58a) was used to replace $\Delta_A(z)$ with $\dot{A}_{\text{opt}}$ and A_{opt}. This functional must be minimized subject to the constraints

$$A_{\text{opt}}(0) = A_{\text{i}}, \text{ and } A_{\text{opt}}(z_{\text{f}}) = A_{\text{f}}. \tag{2.60}$$

After taking variations of S_A and applying the boundary conditions, this becomes a second order boundary value problem for the optimal path A_{opt} of the form

$$\ddot{A}_{\text{opt}} = \frac{\dot{A}_{\text{opt}}^2}{2A_{\text{opt}}}, \tag{2.61a}$$

with boundary conditions

$$A_{\text{opt}}(0) = A_{\text{i}} \text{ and } A_{\text{opt}}(z_{\text{f}}) = A_{\text{f}}. \tag{2.61b}$$

This equation has two solutions,

$$A_{\text{opt},\pm}(z) = \left(\frac{z}{z_{\text{f}}} \left[A_{\text{i}}^{\frac{1}{2}} \pm A_{\text{f}}^{\frac{1}{2}} \right] - A_{\text{i}}^{\frac{1}{2}} \right)^2. \tag{2.62}$$

Putting both of these back into S_A shows that

$$A_{\text{opt}}(z) = \left(\frac{z}{z_{\text{f}}} \left[A_{\text{i}}^{\frac{1}{2}} - A_{\text{f}}^{\frac{1}{2}} \right] - A_{\text{i}}^{\frac{1}{2}} \right)^2 \tag{2.63a}$$

is the global minimum path, which gives a derivative, and equivalently a biasing weight function, of

$$\dot{A}_{\text{opt}}(z) = \Delta_A(z) = \frac{2\left[A_{\text{i}}^{\frac{1}{2}} - A_{\text{f}}^{\frac{1}{2}} \right]}{z_{\text{f}}} \left(\frac{z}{z_{\text{f}}} \left[A_{\text{i}}^{\frac{1}{2}} - A_{\text{f}}^{\frac{1}{2}} \right] - A_{\text{i}}^{\frac{1}{2}} \right). \tag{2.63b}$$

The biasing solution given in Equation (2.63) is plotted in Figure 2.4.

Biasing the Frequency. The calculation for $\Delta_\Omega(z)$ is similar to the calculation of $\Delta_A(z)$, however, since changes in the amplitude parameter do affect the L^2-norm of $v_\Omega^\dagger(z,t)$, the dynamics of both $\Delta_\Omega(z)$ and $\Delta_A(z)$ must be included when constructing the cumulative L^2-norm, where as $\Delta_T(z)$ and $\Delta_\Phi(z)$ can be set to zero. This gives

$$\begin{aligned} S_\Omega(A_{\text{opt}}, \dot{A}_{\text{opt}}, \dot{\Omega}_{\text{opt}}) &= \int_0^{z_{\text{f}}} (\Delta_A(z))^2 \left\| f_A(z,t) \right\|^2 + (\Delta_\Omega(z))^2 \left\| f_\Omega(z,t) \right\|^2 dz \\ &= \int_0^{z_{\text{f}}} \frac{\dot{A}_{\text{opt}}^2}{2A_{\text{opt}}} + \frac{3\dot{\Omega}_{\text{opt}}^2}{2A_{\text{opt}}} \, dz, \end{aligned} \tag{2.64}$$

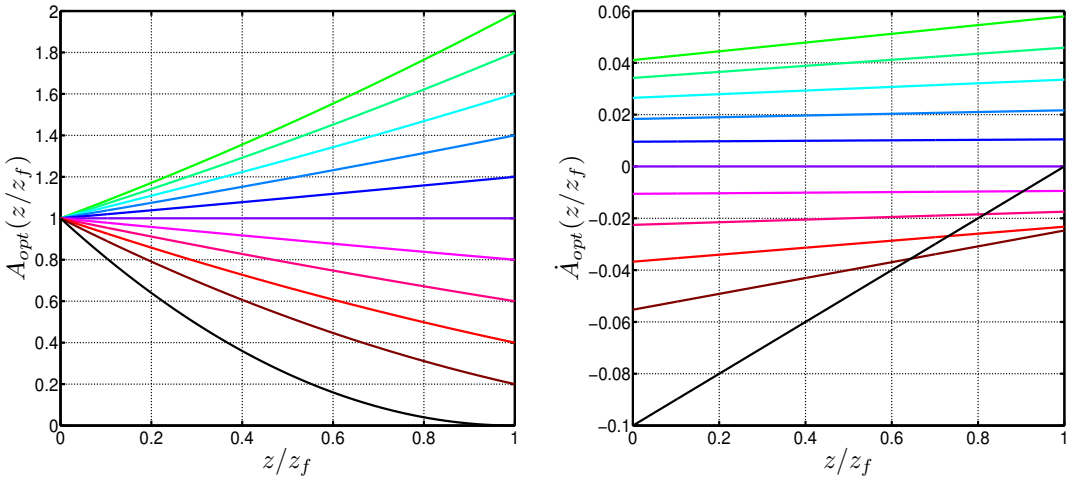


Figure 2.4 Left: The plot of Equation (2.63a) for various targeted final values of amplitude. Right: The corresponding derivatives given by Equation (2.63b) and equivalently the biasing weight functions for the amplitude parameter.

where again the Equations (2.58) were used to replace $\Delta_A(z)$ and $\Delta_\Omega(z)$ with $\dot{A}_{\rm opt}$ and $\dot{\Omega}_{\rm opt}$, respectively. Constraining this minimization are the initial conditions $A_{\rm opt}(0) = A_{\rm i}$ and $\Omega_{\rm opt}(0) = \Omega_{\rm i}$, along with the end point condition $\Omega(z_{\rm f}) = \Omega_{\rm f}$. However, since $A_{\rm opt}(z_{\rm f})$ is not constrained, it will be determined in the course of the minimization. Combining S_Ω and the above constraints into a Lagrange multiplier form and taking variations yields a coupled system of ODEs given by,

$$\ddot{A}_{\rm opt} = \frac{\dot{A}^2_{\rm opt}}{2A_{\rm opt}} - 3\frac{\dot{\Omega}^2_{\rm opt}}{2A_{\rm opt}} \tag{2.65a}$$

and

$$\frac{\dot{\Omega}_{\rm opt}}{A_{\rm opt}} = c_\Omega, \tag{2.65b}$$

with

$$A_{\rm opt}(0) = A_{\rm i}, \quad \dot{A}_{\rm opt}(z_{\rm f}) = 0, \quad \Omega_{\rm opt}(0) = \Omega_{\rm i}, \quad \Omega_{\rm opt}(z_{\rm f}) = \Omega_{\rm f}, \tag{2.66}$$

where c_Ω is a constant of integration. These equations can be decoupled by replacing $\dot{\Omega}_{\text{opt}}$ with $c_\Omega A_{\text{opt}}(z)$, which yields an equation for $A_{\text{opt}}(z)$ that has the solution

$$A_{\text{opt}}(z) = A_{\text{i}} \left[\frac{\cos\left(\frac{\sqrt{3}}{2} c_\Omega (z - z_{\text{f}})\right)}{\cos\left(\frac{\sqrt{3}}{2} c_\Omega z_{\text{f}}\right)} \right]^2, \tag{2.67}$$

and derivative

$$\dot{A}_{\text{opt}}(z) = -A_{\text{i}} \sqrt{3} c_\Omega \frac{\cos\left(\frac{\sqrt{3}}{2} c_\Omega (z - z_{\text{f}})\right) \sin\left(\frac{\sqrt{3}}{2} c_\Omega (z - z_{\text{f}})\right)}{\left(\cos\left(\frac{\sqrt{3}}{2} c_\Omega z_{\text{f}}\right)\right)^2}, \tag{2.68}$$

This also gives

$$\dot{\Omega}_{\text{opt}} = A_{\text{i}} c_\Omega \left[\frac{\cos\left(\frac{\sqrt{3}}{2} c_\Omega (z - z_{\text{f}})\right)}{\cos\left(\frac{\sqrt{3}}{2} c_\Omega z_{\text{f}}\right)} \right]^2, \tag{2.69}$$

which after integrating gives,

$$\Omega_{\text{opt}}(z) = \Omega_{\text{i}} + \frac{A_{\text{i}} \left[\sqrt{3} c_\Omega z + \sin\left(\sqrt{3} c_\Omega (z - z_{\text{f}})\right) + \sin\left(\sqrt{3} c_\Omega z_{\text{f}}\right)\right]}{2\sqrt{3} \left(\cos\left(\frac{\sqrt{3}}{2} c_\Omega z_{\text{f}}\right)\right)^2}. \tag{2.70}$$

The final boundary condition $\Omega_{\text{opt}}(z_{\text{f}}) = \Omega_{\text{f}}$ determines the value of c_Ω and thus completes the solution which is plotted in Figure 2.5.

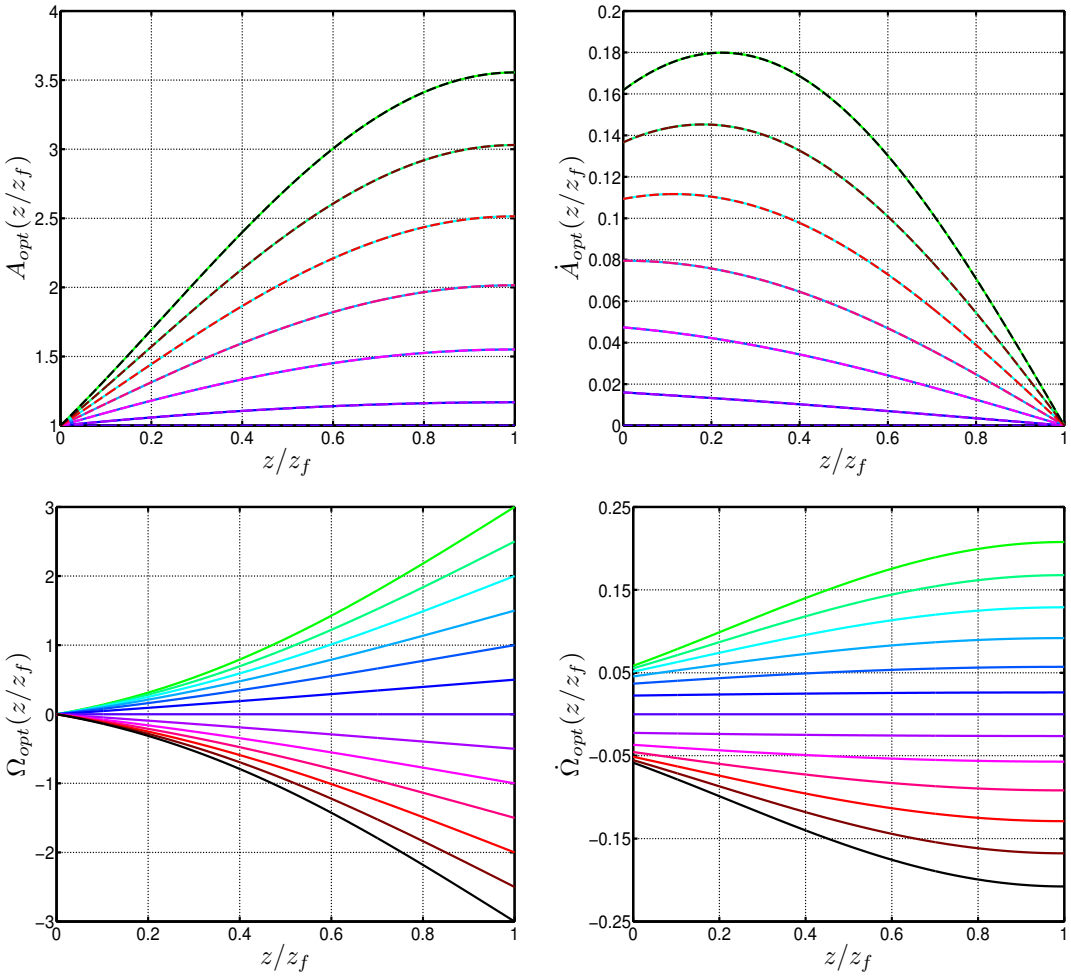


Figure 2.5 Top Left: Optimal amplitude parameter paths given in Equation (2.67) for various targeted final values of the frequency parameter. Top Right: Derivatives of the optimal paths for the amplitude parameter given in Equation (2.68) for various targeted final values of the frequency parameter. Bottom Left: Optimal frequency parameter paths given in Equation (2.69) for various targeted final values. Bottom Right: Derivatives of the optimal paths for the frequency parameter (and equivalently the biasing paths) given in Equation (2.70) for various targeted final values of the frequency parameter.

Biasing the Timing. Because the optimal path of the time parameter depends directly on the optimal path of the frequency parameter, as seen in Equations (2.58c) and (2.58b), one must include both Δ_T and Δ_Ω when constructing the cumulative L^2-norm associated with biasing the time parameter. In addition, since the L^2-norm of both f_T and f_Ω depend on A_{opt}, non-zero biasing weights for the amplitude should

be included as well, leading to a complicated three-dimensional BVP which must be solved numerically. However, if it is assumed that amplitude changes are small enough to be neglected, i.e., $A_{\text{opt}}(z) = A_{\text{i}}$, an analytical solution can be found. Using this approach, the cumulative L^2-norm takes the form

$$\begin{aligned} S_T(\dot{T}_{\text{opt}}, \Omega_{\text{opt}}, \dot{\Omega}_{\text{opt}}) &= \int_0^{z_{\text{f}}} (\Delta_T)^2 \left\| f_T(z,t) \right\|^2 + (\Delta_\Omega)^2 \left\| f_\Omega(z,t) \right\|^2 dz \\ &= \int_0^{z_{\text{f}}} \frac{6A_{\text{i}}^3}{\pi^2} \left(\dot{T}_{\text{opt}} - \Omega_{\text{opt}} \right)^2 + \frac{3}{2A_{\text{i}}} \dot{\Omega}_{\text{opt}}^2 \, dz, \end{aligned} \tag{2.71}$$

where $\Delta_T(z)$ and $\Delta_\Omega(z)$ were replaced with $\left(\dot{T}_{\text{opt}} - \Omega_{\text{opt}} \right)$ and $\dot{\Omega}_{\text{opt}}$, respectively, through the use of Equations (2.58c) and (2.58b). This functional needs to minimized under the constraints $\Omega_{\text{opt}}(0) = \Omega_{\text{i}}$, $T_{\text{opt}}(0) = T_{\text{i}}$ and $T_{\text{opt}}(z_{\text{f}}) = T_{\text{f}}$, where it is noted that since $\Omega_{\text{opt}}(z_{\text{f}})$ is unconstrained, it will be determined through the minimization process.

Taking variations of equation (2.71) yields the system,

$$\ddot{\Omega}_{\text{opt}}(z) = -c_T \frac{A_{\text{i}}}{3}, \tag{2.72a}$$

and

$$\dot{T}_{\text{opt}}(z) = \Omega_{\text{opt}}(z) + c_T \frac{\pi^2}{12A_{\text{i}}^3}, \tag{2.72b}$$

with boundary conditions $\Omega_{\text{opt}}(0) = \Omega_{\text{i}}$, $\dot{\Omega}_{\text{opt}}(z_{\text{f}}) = 0$, $T_{\text{opt}}(0) = T_{\text{i}}$ and $T_{\text{opt}}(z_{\text{f}}) = T_{\text{f}}$, and where c_T represents a constant of integration. These equations are easily solved, giving solutions of the form,

$$\Omega_{\text{opt}}(z) = \Omega_{\text{i}} + c_T \frac{A_{\text{i}}}{6} \left(2z_{\text{f}} - z \right) z, \tag{2.73a}$$

$$\dot{\Omega}_{\text{opt}}(z) = c_T \frac{A_{\text{i}}}{3} \left(z_{\text{f}} - z \right), \tag{2.73b}$$

$$T_{\text{opt}}(z) = T_{\text{i}} + \Omega_{\text{i}} z + c_T \left[\frac{A_{\text{i}}}{18} \left(3z_{\text{f}} - z \right) z^2 + \frac{\pi^2}{12A_{\text{i}}^3} z \right], \tag{2.73c}$$

and

$$\dot{T}_{\mathrm{opt}}(z) = \Omega_{\mathrm{opt}}(z) + c_T \frac{\pi^2}{12 A_{\mathrm{i}}^3}, \tag{2.73d}$$

where c_T is found to be,

$$c_T = \frac{(T_{\mathrm{f}} - T_{\mathrm{i}}) - \Omega_{\mathrm{i}} z_{\mathrm{f}}}{\frac{\pi^2}{12 A_{\mathrm{i}}^3} z_{\mathrm{f}} + \frac{A_{\mathrm{i}}}{9} z_{\mathrm{f}}^3}. \tag{2.74}$$

These solutions are plotted in Figure 2.6.

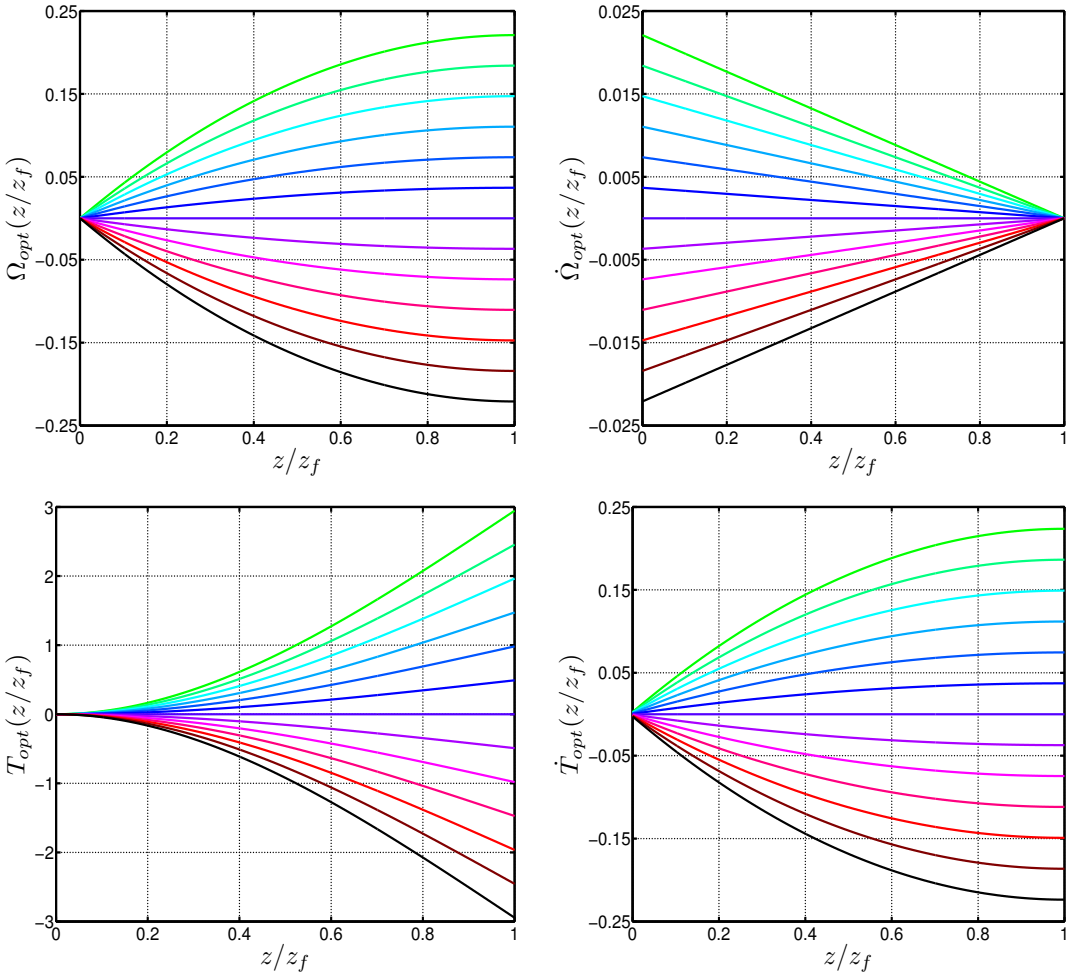


Figure 2.6 Top Left: Optimal frequency parameter paths given by Equation (2.73a) for various targeted final values of the time parameter. Top Right: Derivatives of the optimal paths for the frequency parameter given in Equation (2.73b) for various targeted final values of the time parameter. Bottom Left: Optimal paths for the time parameter given by Equation (2.73c) for various targeted final values. Bottom Right: Derivatives of the optimal paths for the time parameter given by Equation (2.73d) for various targeted final values of the time parameter. In all plots, $A_{\mathrm{i}} = 1$.

Biasing the Phase. The phase biasing is the most difficult to determine, since Equation (2.58d) indicates that all four biasing vectors need to be considered in the formation of the cumulative L^2-norm, resulting in a four-dimensional, second-order BVP for the optimal path, with a solution only attainable through the use of numerical methods.

However, if the initial values of both the timing and frequency parameters are zero, these parameters can be assumed to have a negligible effect on the evolution of the phase parameter. In contrast, the amplitude parameter starts out as an $O(1)$ quantity and therefore imparts a significant contribution to the phase evolution. With this approximation, the cumulative L^2-norm takes the form

$$\begin{aligned} S_\Phi(A_{\text{opt}}, \dot{A}_{\text{opt}}, \dot{\Phi}_{\text{opt}}) &= \int_0^{z_\text{f}} \Delta_A(z)\|f_A(z,t)\|^2 + \Delta_\Phi(z)\|f_\Phi(z,t)\|^2\, dz \\ &= \int_0^{z_\text{f}} \frac{\dot{A}_{\text{opt}}^2}{2A_{\text{opt}}} + \frac{18}{12+\pi^2} A_{\text{opt}} \left(\dot{\Phi}_{\text{opt}} - \frac{A_{\text{opt}}^2}{2}\right)^2 dz, \end{aligned} \tag{2.75}$$

under the constraints $A_{\text{opt}}(0) = A_\text{i}$, $\Phi_{\text{opt}}(0) = \Phi_\text{i}$ and $\Phi_{\text{opt}}(z_\text{f}) = \Phi_\text{f}$, with $A_{\text{opt}}(z_\text{f})$ determined through the minimization process. Taking variations gives the following two-dimensional BVP,

$$\ddot{A}_{\text{opt}} = \frac{\dot{A}_{\text{opt}}^2}{2A_{\text{opt}}} + \frac{12+\pi^2}{72} c_\Phi^2 \frac{1}{A_{\text{opt}}} - c_\Phi A_{\text{opt}}^2, \tag{2.76a}$$

and

$$\dot{\Phi}_{\text{opt}} = \frac{A_{\text{opt}}^2}{2} + c_\Phi \frac{12+\pi^2}{36} \frac{1}{A_{\text{opt}}}, \tag{2.76b}$$

where c_Φ is a constant of integration and $A_{\text{opt}}(0) = A_\text{i}$, $\dot{A}_{\text{opt}}(z_\text{f}) = 0$, $\Phi_{\text{opt}}(0) = \Phi_\text{i}$, and $\Phi_{\text{opt}}(z_\text{f}) = \Phi_\text{f}$.

This system can be solved by integrating Equation (2.76b) and using the boundary conditions on Φ_{opt} to get c_Φ in terms of integrals involving A_{opt},

$$c_\Phi = \frac{\Phi_\text{f} - \frac{1}{2}\int_0^z A_{\text{opt}}^2(z')\, dz'}{\frac{12+\pi^2}{36}\int_0^z \frac{1}{A_{\text{opt}}(z')}\, dz'}. \tag{2.77}$$

Substituting this back into Equation (2.76a), gives an integro-differential equation for the optimal amplitude path which can be solved using a modified shooting method, the results of which are presented in Figure 2.7

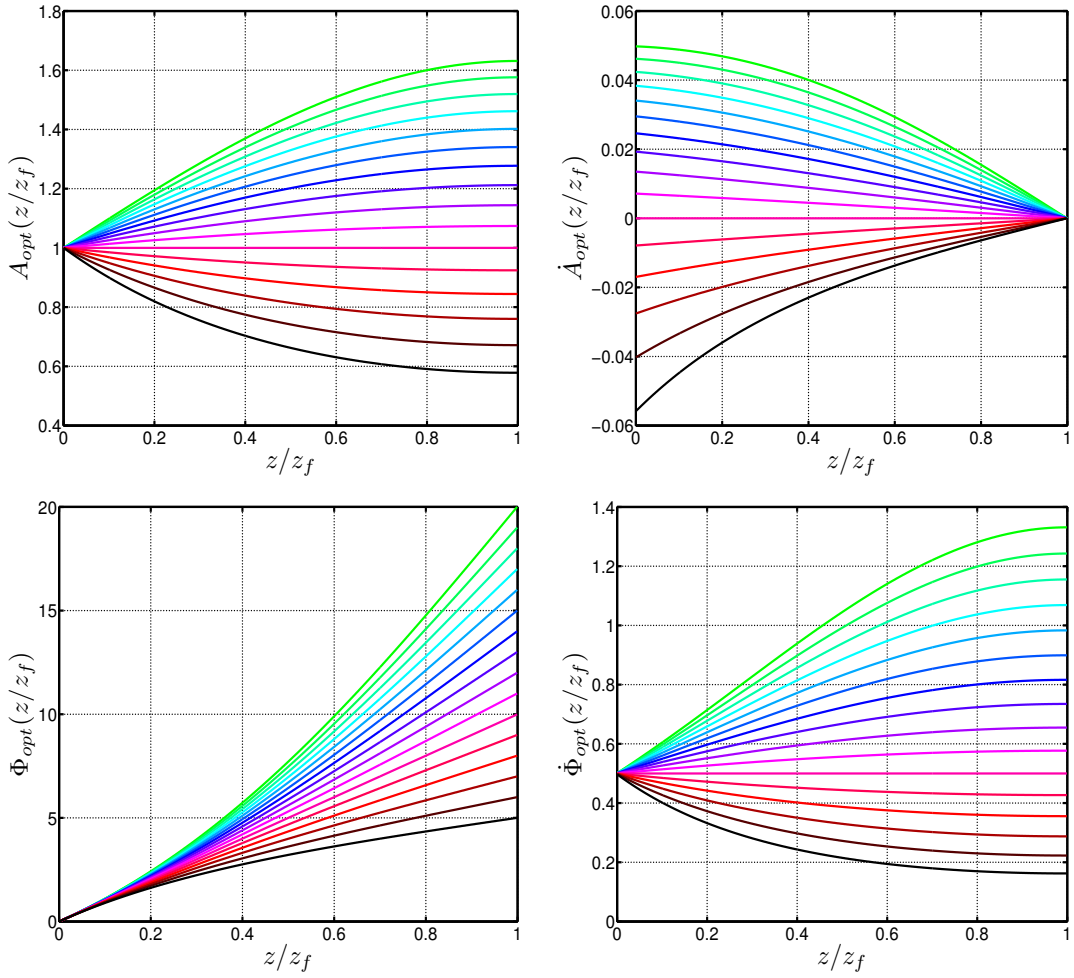


Figure 2.7 Top Left: Optimal amplitude parameter paths given by Equation (2.76a) for various targeted final values of the phase parameter. Top Right: Derivatives of the optimal paths for the amplitude parameter for various targeted final values of the phase parameter. Bottom Left: Optimal phase parameter paths given by Equation (2.76b) for various final targeted values of the phase parameter. Bottom Right: Derivatives of the optimal paths for the phase parameter given by Equation (2.76b).

2.4 Results and Discussion

This section shows the results of applying the ISMC method to construct the PDFs of the amplitude, frequency, time and phase parameters of a soliton propagated under the stochastic NLSE in (2.1). As stated previously, the numerical integration of this equation was performed using a pseudospectral method, where evolutionary

stepping in z used a fourth-order integrating factor Runge-Kutta scheme [47, 48] which is outlined in Appendix B. All simulations use solitons with a full width at half maximum (FWHM) of 17.6 ps, which were represented numerically using 128 computational point (modes) on an computational domain that was taken to be 704 ps (or 40 dimensionless time units) wide. The group velocity dispersion and nonlinear parameters of the fiber were taken to be -0.2 ps^2/km and 2.0 1/(km W), respectively. The pulses were transmitted over a distance 10,000 km with amplifiers spaced 50 km apart, translating to a dimensionless step size of $\Delta z = 0.01$, a dimensionless amplifier spacing of $z_a = 0.1$ and total numerical transmission length of 20 non-dimensional units. The fiber loss was taken as 0.2 dB/km (gain of $G = 10$), which when combined with an assumed spontaneous emission factor of 1.4, yields a dimensionless noise strength of $\sigma^2 = 6.3 \times 10^{-5}$. The exact meaning of these parameters is explained in more detail in Appendix A and should be compared with the typical fiber values presented in Table A.1.

Figures 2.8, 2.9, 2.10 and 2.11 displayed below each consist of three plots, the first of which compares the PDFs of a given parameter plotted on a log scale of base 10. In this plot, the curves represent: (black) the ISMC method applied to the stochastic NLSE in (2.1), (green) the standard MC method applied to the stochastic NLSE in (2.1), (blue) the standard MC method applied to the SODEs given in Equation (2.26) and (red) an analytical calculation for the solution of the SODEs in Equation (2.26) linearized around the initial conditions of $A = 1$ and $\Omega = T = \Phi = 0$. The second plot in each figure shows the coefficient of variation (COV) for all the MC curves presented in the top plot, while the last plot in each figure shows number of hits (NOH) each bin received.

Figures 2.8, 2.9 and 2.10 show the results of applying the ISMC method for construction of the PDFs for the amplitude, frequency and time parameters, respectively. The distributions for the frequency and time parameters are in relatively good

agreement with the analytically calculated Gaussian curves obtained from linearizing the SPT equations. This indicates that the true PDFs of these parameters only differ from their Gaussian predictions by a slight change in variance attributed to the amplitude dependence of the variances in fluctuations of the frequency and time parameters, which is given in Equations (2.29) and ignored by the linearization that led to the Gaussian approximations. By contrast, this dependence leads to a significant deviation of the amplitude PDF from its Gaussian approximation, indicating that fluctuations in the amplitude contribute significantly to the high order moments of the amplitude distribution.

Finally, the PDF of the *unwrapped* phase (i.e., extended via Riemannian sheets to the real line) is displayed in Figure 2.11, which shows a large disagreement with the corresponding Gaussian prediction from the linearized SPT equations. This disagreement is expected since the dynamics of all four parameters contribute to the final value of the phase, which is again ignored by the linearization that leads to a Gaussian distribution. More important, however, is the fact that it also disagrees with MC simulations of the full nonlinear SPT equations. Further investigations of this phenomenon, which is shown in Figure 2.12, indicates that the magnitude of this disagreement, as quantified by the translation in the peak of the PDF, grows as the noise bandwidth increases, suggesting that the dispersive radiation that is neglected by first-order SPT plays an important role in the evolution of the phase parameter. Moreover, these plots also indicate that the dispersive radiation has a particularly important effect on the *mean* evolution of the phase.

2.5 Summary

The first section of this chapter began by presenting a stochastically forced version of the nonlinear Schrödinger equation as a model for electromagnetic pulse propagation in nonlinear optical fiber in the presence of noise generated by pulse amplification.

After assuming pulses in the form of the well-known four-parameter family of solitons, it was shown that the first order effect of adding noise to this system is to induce the soliton parameters to undergo a random walk from their initial values. This walk is quantified by soliton perturbation theory, which provides a way of describing the stochastic evolution of the pulse parameters by reducing their dynamics to a four-dimensional SODE. Section 2.2 presented the general ISMC method including the aggregation technique of multiple importance sampling, which enables the simultaneous use of multiple biasing distributions to construct different regions of the PDF in an efficient manner. Section 2.3 described the implementation of the ISMC method to estimate the probability of rare error events in soliton-based communication systems, including the biasing scheme, which is based on the results of SPT. Finally, the results of this method were presented in Section 2.4, in the form of PDF plots for the four soliton parameters.

The results of the ISMC method presented here indicate that, although sufficient in describing the perturbed evolution of a soliton's amplitude, frequency and timing, first-order SPT cannot correctly capture the evolution of the soliton's phase in the presence of dispersive radiation. Furthermore, this inadequacy is exacerbated in numerical studies involving idealized white noise, since as shown above, the evolution of the phase's mean value is dependent on the bandwidth of noise feeding the radiation, a quantity that is often dictated by the bandwidth of the simulation itself. It should also be noted that this radiation-driven phase rotation is not unique to the integrable form of the NLSE, and in fact has been commented on in past investigations focusing on more complicated models of stochastic pulse propagation in optical fiber [49], the most notable being stochastic versions of the dispersion managed nonlinear Schrödinger equation (DMNLSE) [50, 51, 52] and the complex Ginzburg-Landau equation [53, 54, 55]. The next chapter is devoted to understanding the interaction between dispersive radiation and the evolution of the soliton's phase

parameter, leading to the derivation of an improved low-order reduction of the soliton dynamics in the presence of radiation. This scheme is used to build an improved importance sampling method for NLSE in Chapter 4, and is extended to the case of DMNLSE in Chapter 5.

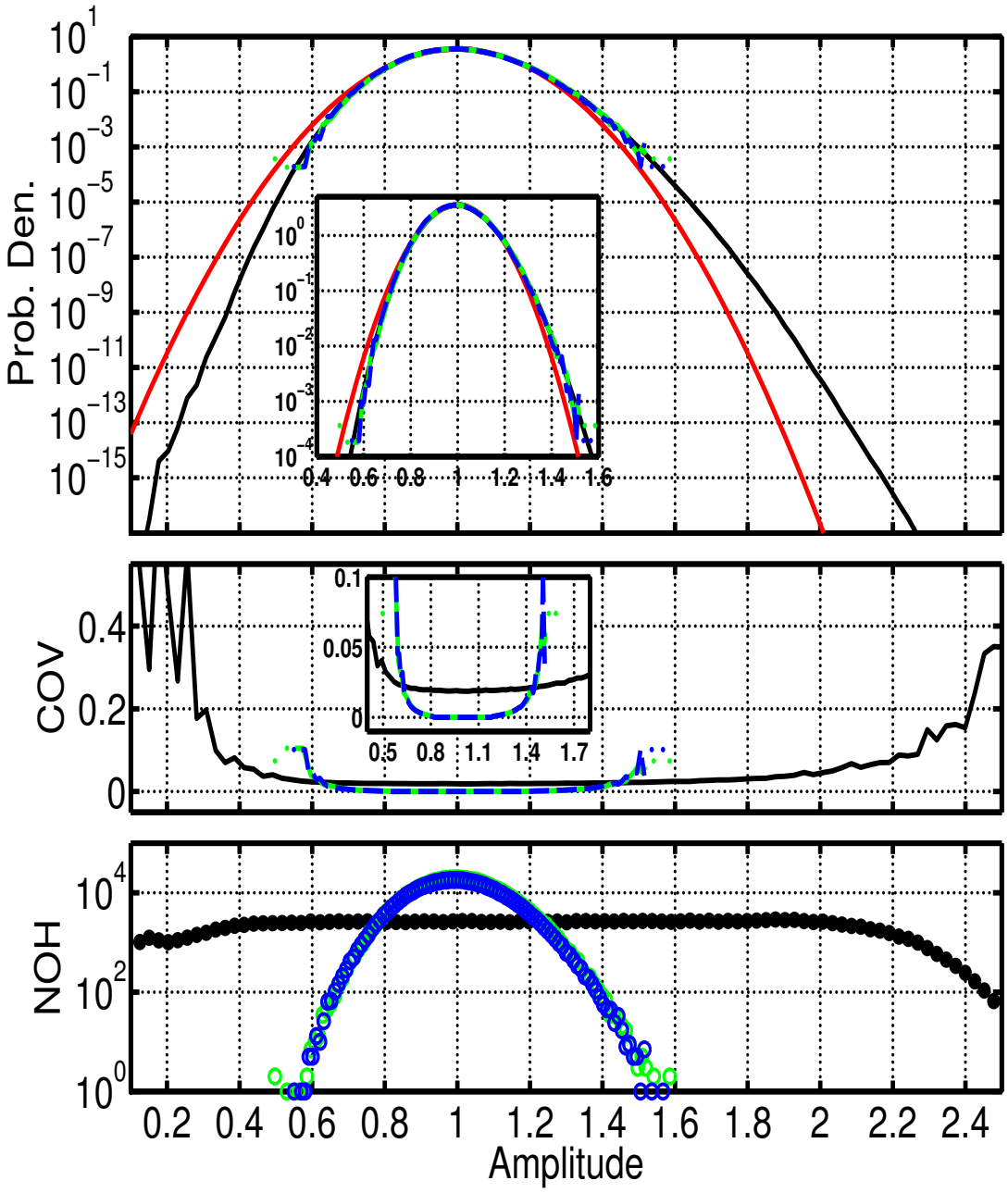


Figure 2.8 Top: The black curve gives the PDF of the amplitude parameter using 2×10^5 ISMC runs of the stochastically forced NLSE in (2.1). For comparison, the dashed blue and solid green curves represent the results of 5×10^5 MC runs of the SODEs given in (2.26) and the stochastically forced NLSE in (2.1), respectively. In addition, the solid red curve is the analytical solution of Equations (2.26) linearized around the initial conditions $A = 1$ and $\Omega = T = \Phi = 0$. All plots are on a Log scale of base 10. Middle: Each curve represents the COV for corresponding curve in top figure. Bottom: Each sequence of markers represents the NOH each bin received in the construction of the PDFs in the top figure. All plots are on a Log scale of base 10.

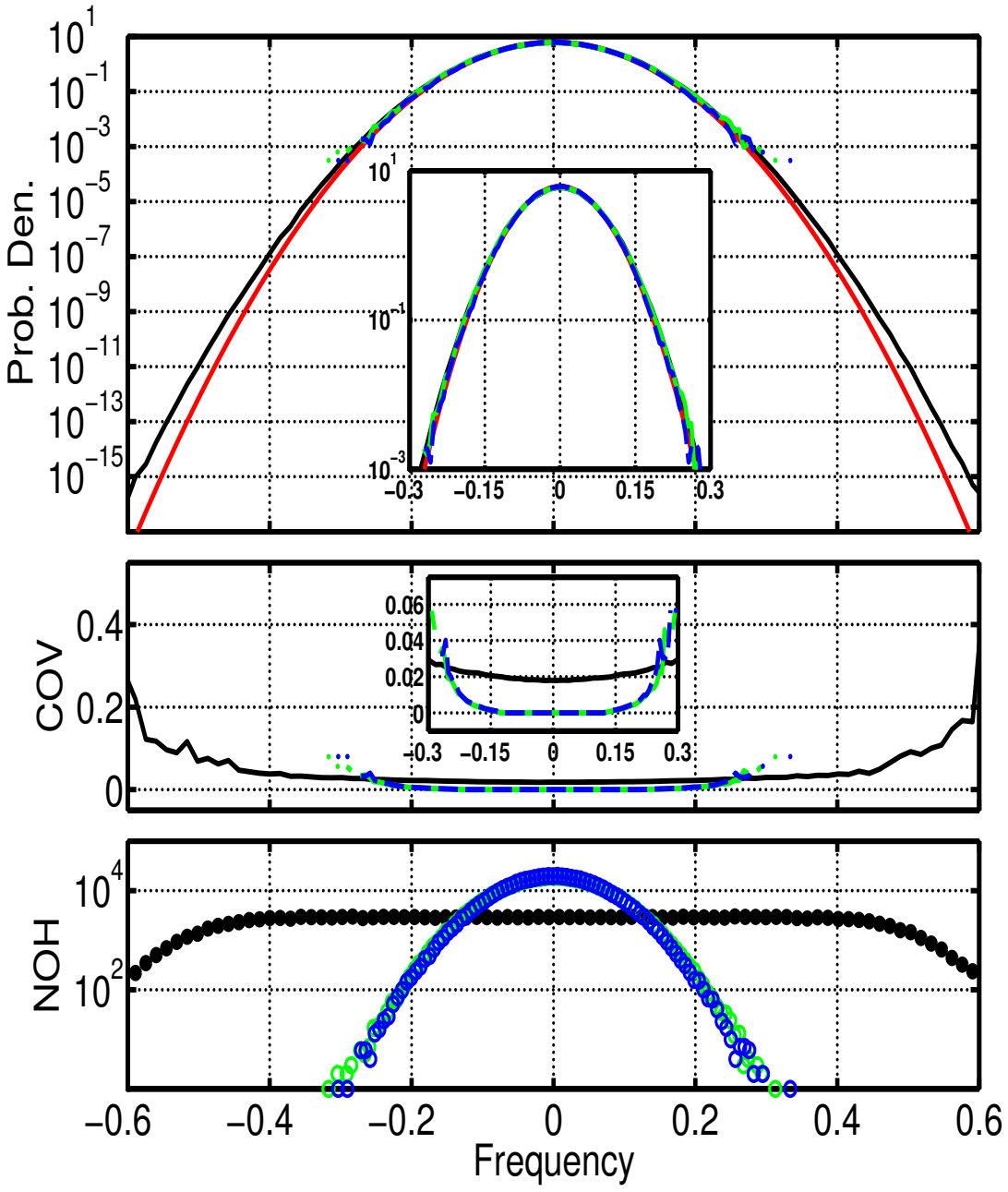


Figure 2.9 Top: The black curve gives the PDF of the frequency parameter using 2×10^5 ISMC runs of the stochastically forced NLSE in (2.1). For comparison, the dashed blue and solid green curves represent the results of 5×10^5 MC runs of the SODEs given in (2.26) and the stochastically forced NLSE in (2.1), respectively. In addition, the solid red curve is the analytical solution of Equations (2.26) linearized around the initial conditions $A = 1$ and $\Omega = T = \Phi = 0$. All plots are on a Log scale of base 10. Middle: Each curve represents the COV for corresponding curve in top figure. Bottom: Each sequence of markers represents the NOH each bin received in the construction of the PDFs in the top figure. All plots are on a Log scale of base 10.

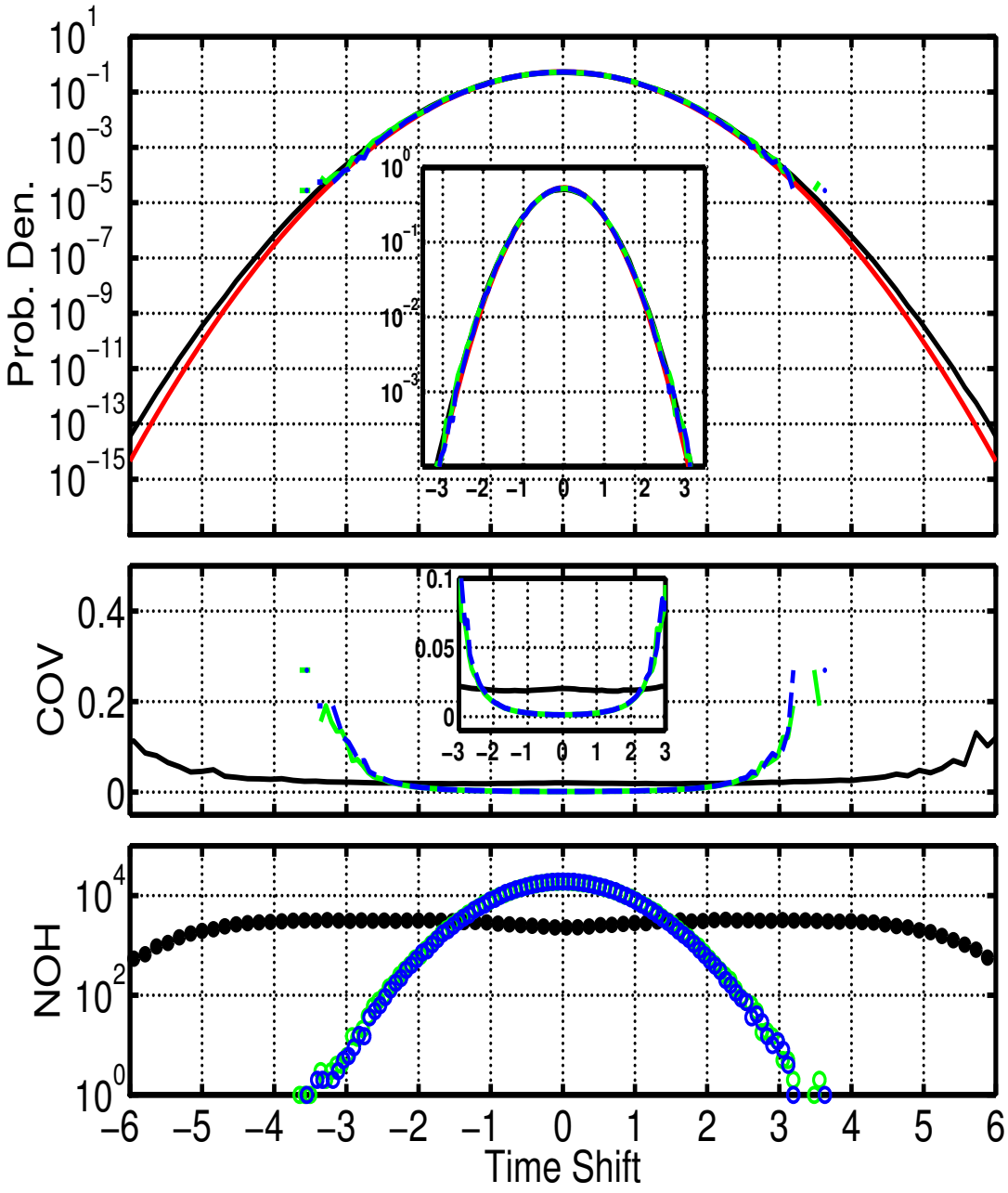


Figure 2.10 Top: The black curve gives the PDF of the time shift parameter using 2×10^5 ISMC runs of the stochastically forced NLSE in (2.1). For comparison, the dashed blue and solid green curves represent the results of 5×10^5 MC runs of the SODEs given in (2.26) and the stochastically forced NLSE in (2.1), respectively. In addition, the solid red curve is the analytical solution of Equations (2.26) linearized around the initial conditions $A = 1$ and $\Omega = T = \Phi = 0$. All plots are on a Log scale of base 10. Middle: Each curve represents the COV for corresponding curve in top figure. Bottom: Each sequence of markers represents the NOH each bin received in the construction of the PDFs in the top figure. All plots are on a Log scale of base 10.

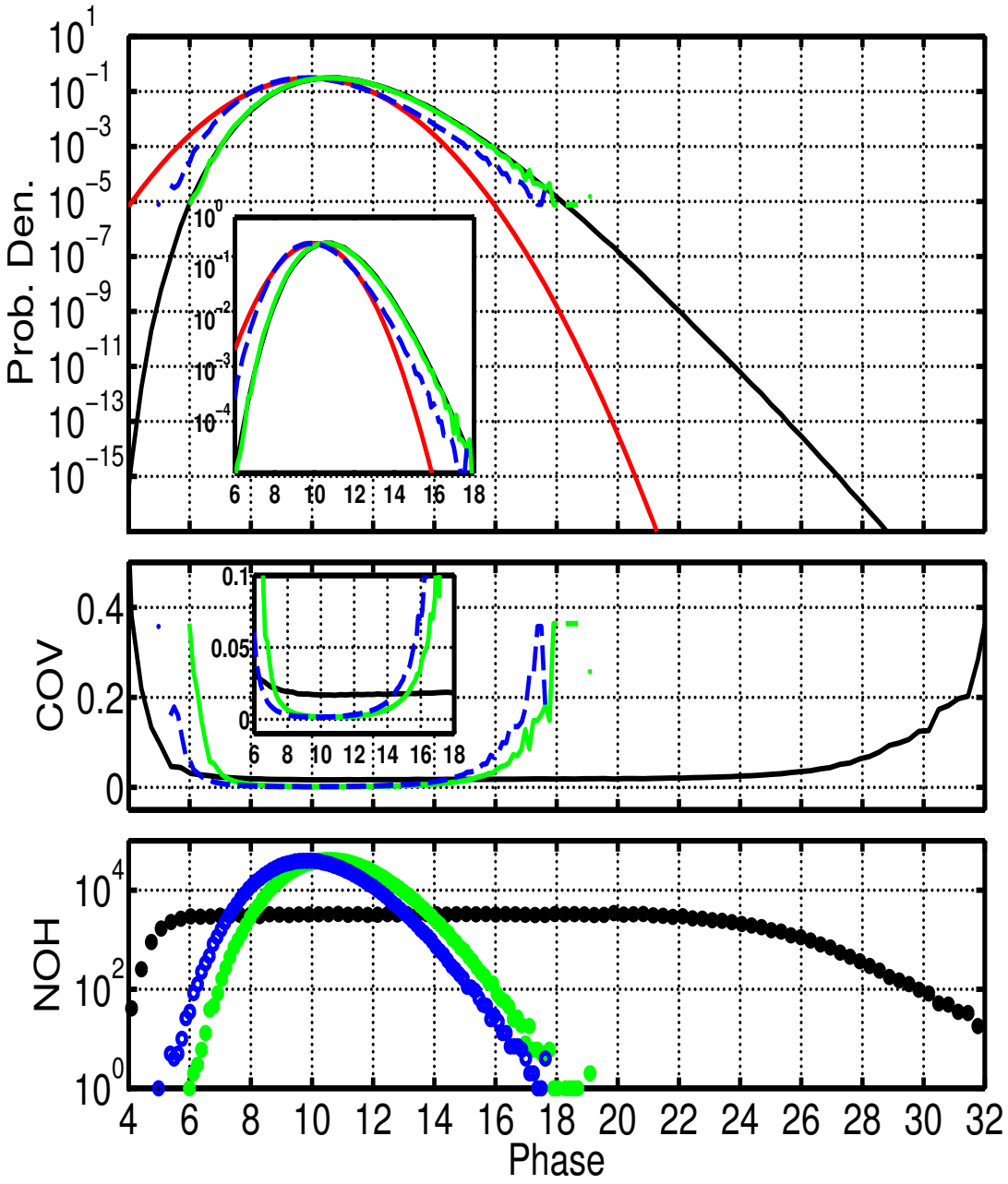


Figure 2.11 Top: The black curve gives the PDF of the phase parameter using 2×10^5 ISMC runs of the stochastically forced NLSE in (2.1). For comparison, the dashed blue and solid green curves represent the results of 5×10^5 MC runs of the SODEs given in (2.26) and the stochastically forced NLSE in (2.1), respectively. In addition, the solid red curve is the analytical solution of Equations (2.26) linearized around the initial conditions $A = 1$ and $\Omega = T = \Phi = 0$. All plots are on a Log scale of base 10. Middle: Each curve represents the COV for corresponding curve in top figure. Bottom: Each sequence of markers represents the NOH each bin received in the construction of the PDFs in the top figure. All plots are on a Log scale of base 10.

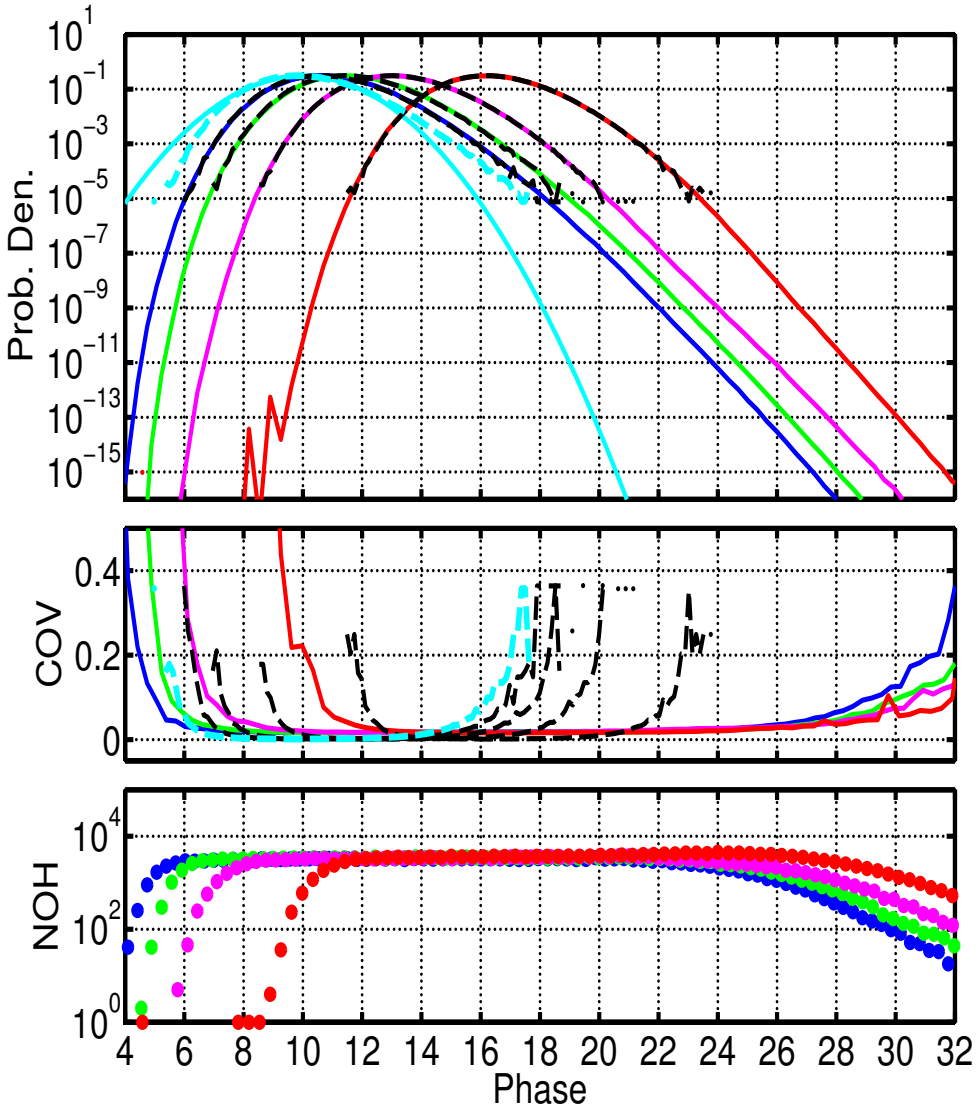


Figure 2.12 Top: Comparison of PDFs for the phase parameter while varying the number of simulation modes. The dashed and solid cyan curves are respectively, the results of 5×10^5 MC runs of the nonlinear system of Equations in (2.26) and an analytical solution to these equations linearized around the initial conditions $A = 1$ and $\Omega = T = \Phi = 0$. The blue, green, magenta and red curves are the PDFs generated with 2×10^5 runs of the ISMC method, using 128, 256, 512 and 1024 simulation modes, respectively. Each colored curve is accompanied by a black dashed line representing the results of 1×10^6 standard MC runs of the stochastically forced NLSE in (2.1). All plots are on a Log scale of base 10. Middle: Each colored curve represents the COV for the corresponding PDF in the top figure, while the black dashed curve represents the COV for the 5×10^5 MC runs of the stochastically forced NLSE in (2.1). Bottom: Each sequence of colored markers correspond to the NOH each bin received under the ISMC runs that produced to the PDF plots in the top figure.

CHAPTER 3

RADIATION-INDUCED PHASE DRIFT

As the preceding chapter has shown, dispersive radiation, which appears as second order term in a perturbation expansion, has the ability to contribute a first-order effect to the evolution of a soliton's phase parameter. Although this was first realized in numerical simulations of the NLSE [23, 56], similar effects have been reported in related equations [51], most notably in the dispersion managed nonlinear Schrödinger equation (DMNLSE), suggesting that this phenomenon has origins in the cubic nonlinear structure common to both equations. In this chapter, the evolution of a NLSE soliton in the presence of radiation is examined with the aim of understanding this phase drift in detail and developing a more accurate reduction method, capable of capturing the correct evolution of the soliton's phase parameter that can be used to build an improved ISMC scheme.

The first part of this chapter is devoted to understanding the evolution of radiation in a nonlinear medium and the interactions between radiation and soliton solutions of the NLSE. The main difficulty this problem presents is in the lack of functional form for the radiation. Indeed, a complete representation of the radiation's evolution can only be found by appealing to the integrability of the NLSE [46] and the inverse scattering transform (IST) [57]. As one might expect, however, the solutions resulting from the IST are too cumbersome to be used directly in computations. Here, this theory will only be used as a source of insight into the evolution of dispersive radiation, from which an appropriate evolutionary approximation for the radiation can be constructed. This approximation is used in the second part of this chapter to construct a reduction method capable of producing a set of SODEs for the evolution of the pulse and the radiation, along with the dominant interaction that occurs between the two.

Although the initial approach is based on extending the multiple scales approximation of SPT in Section 2.1.1, a simpler method is found in a variational approach, which can be extended to non-integrable (and even non-hamiltonian) evolution equations, such as the DMNLSE [58, 59, 60].

The construction of an improved reduction method has two important implications. Broadly, it demonstrates an interesting problem in the area of asymptotic methods, where the approximating expansion becomes disordered in certain parameter regimes, with second order terms (radiation) contributing at first order. More narrowly, it provides an improved method for a biasing distribution needed in the application of importance sampling. In this sense, the work presented here "closes the loop" in a problem where an initially crude approximation was used to bias simulations of a high-dimensional model, which provided additional information on the dynamics to feed back into an improved low-dimensional reduction and therefore improved biasing. This potential for iterative improvements in ISMC used for rare event computations is exciting and novel, and is expected to have relevance to other contexts where rare events are important. Applying the information from this chapter to the application of the ISMC method for rare event detection is considered in the next chapter.

3.1 Extending SPT to Second Order

The most direct approach for modeling the radiation-induced phase drift is to extend the SPT derived in Section 2.1.1 to second order. In Section 2.1.1, it was assumed that the stochastic forcing manifest as adiabatic motion of the four soliton parameters, which motivated the introduction of the short length scale $z_1 = \sigma z$. Extending this to the next order requires the introduction of an additional length scale $z_2 = \sigma^2 z$, which in turn introduces an additional derivative in the NLSE, now given as

$$i\frac{\partial u}{\partial z} + i\sigma\frac{\partial u}{\partial z_1} + i\sigma^2\frac{\partial u}{\partial z_2} + \frac{1}{2}\frac{\partial^2 u}{\partial t^2} + |u|^2 u = i\sigma n(t, z, z_1, \dots), \tag{3.1}$$

and prompts the inclusion of an additional term in the power series expansion for the solution,

$$u = \left[v_0(t,z,z_1,z_2) + \sigma v_1(t,z,z_1,z_2) + \sigma^2 v_2(t,z,z_1,z_2)\right] \exp(i\Theta(t,z,z_1,z_2))\,. \quad (3.2)$$

The $O(1)$ and $O(\sigma)$ equations remain unchanged, so the first order solution remains a soliton with parameters that now depend on both z_1 and z_2,

$$\begin{aligned} &v_0(t,z,z_1,z_2)\exp(i\Theta(t,z,z_1,z_2)) \\ &\qquad = u_{\rm sol}(t,z;A(z_1,z_2),\Omega(z_1,z_2),T_0(z_1,z_2),\Phi_0(z_1,z_2)) \\ &\qquad = u_0(t,z,z_1,z_2)\exp(i\Theta(t,z,z_1,z_2))\,, \end{aligned} \quad (3.3)$$

and the second order equation still governs the evolution of the radiation,

$$\begin{aligned} L_{\rm nls}(v_1\exp(i\Theta)\,;u_{\rm sol}) &= in(t,z) \\ &\quad -i\left[\frac{\partial u_{\rm sol}}{\partial A}\frac{dA}{dz_1} + \frac{\partial u_{\rm sol}}{\partial T}\frac{\partial T}{\partial z_1} + \frac{\partial u_{\rm sol}}{\partial \Omega}\frac{d\Omega}{dz_1} + \frac{\partial u_{\rm sol}}{\partial \Phi}\frac{\partial \Phi}{\partial z_1}\right]. \end{aligned} \quad (3.4)$$

The first order evolution of each parameter is still found by appealing to the Fredholm orthogonality requirement. Recalling that $L(\cdot;u_{\rm sol})$ is the nonlinear Schrödinger operator linearized around the soliton solution, it can be written using the alternative representation

$$\begin{aligned} \left(\frac{\partial}{\partial z} + \Omega\frac{\partial}{\partial t}\right)v_1 - L(v_1;u_0) &= n(t,z)\exp(-i\Theta(t,z,z_1)) \\ &\quad - \left[\frac{\partial u_0}{\partial A}\frac{dA}{dz_1} + \frac{\partial u_0}{\partial T}\frac{\partial T}{\partial z_1} + iu_0 t\frac{d\Omega}{dz_1} + iu_0\frac{\partial \Phi}{\partial z_1}\right], \end{aligned} \quad (3.5a)$$

where

$$L(v_1;u_0) = \frac{i}{2}\frac{\partial^2 v_1}{\partial t^2} - \frac{i}{2}A^2 v_1 + 2i|u_0|^2 v_1 + i(u_0)^2\bar{v}_1. \quad (3.5b)$$

Once the Fredholm orthogonality requirement is satisfied, the evolution of the radiation follows

$$\left(\frac{\partial}{\partial z} + \Omega\frac{\partial}{\partial t}\right)v_1 - L(v_1;u_0) = n_{\rm r}(t,z), \quad (3.6)$$

where $n_{\mathrm{r}}(t,z)$ is remaining noise after the contributions to changes in the four soliton parameters are projected out. Normally, this equation is solved before continuing to the next order, however, this would require transforming the radiation into a basis of eigenfunctions for the $L_{\mathrm{nls}}(\cdot;u_{\mathrm{sol}})$ (or equivalently $L(\cdot;u_0)$), alternatively known as radiation modes, which correspond to non-zero eigenvalues. Finding these eigenfunctions is a nontrivial task, which is primarily complicated by the dependence of linearized operator $L(v_1;u_0)$ on the pulse u_0. Indeed, in the region $|t|\gg 1$, these terms can be neglected, implying that the radiation modes consist of *dressed* Fourier modes [61, 62, 63] which asymptote to the standard Fourier modes far from the position of the soliton.

Formally continuing with the expansion, the $O(\sigma^2)$ equation takes the form

$$\begin{aligned} L_{\mathrm{nls}}(v_2\exp(i\Theta)\,;u_{\mathrm{sol}}) = -i&\left[\frac{\partial u_{\mathrm{sol}}}{\partial A}\frac{dA}{dz_2}+\frac{\partial u_{\mathrm{sol}}}{\partial T}\frac{\partial T}{\partial z_2}+\frac{\partial u_{\mathrm{sol}}}{\partial \Omega}\frac{d\Omega}{dz_2}+\frac{\partial u_{\mathrm{sol}}}{\partial \Phi}\frac{\partial \Phi}{\partial z_2}\right] \\ &-\left[i\left(\frac{\partial v_1}{\partial z_1}+iv_1\frac{\partial \Theta}{\partial z_1}\right)+2|v_1|^2u_0+\bar{u}_0v_1^2\right]\exp(i\Theta)\,, \end{aligned} \tag{3.7}$$

or equivalently

$$\begin{aligned} \left(\frac{\partial}{\partial z}+\Omega\frac{\partial}{\partial t}\right)v_2 - L(v_2;u_0) = -&\left[v_A\frac{dA}{dz_2}+v_T\frac{\partial T}{\partial z_2}+(v_\Omega+Tv_\Phi)\frac{d\Omega}{dz_2}+v_\Phi\frac{\partial \Phi}{\partial z_2}\right] \\ &-\left[\left(\frac{\partial v_1}{\partial z_1}+iv_1\frac{\partial \Theta}{\partial z_1}\right)-2v_\Phi|v_1|^2-v_\Phi v_1^2\right], \end{aligned} \tag{3.8}$$

where L is defined in Equation (3.5b) and the normalizable soliton eigenfunctions have been inserted where appropriate. Using the same orthogonality requirement as before allows the projection of the right-hand side of Equation (3.8) against the adjoint eigenfunctions, where it is again noted that the radiation, represented as v_1, remains orthogonal to the basis of adjoint modes throughout its evolution, thus giving

$$\frac{dA}{dz_2}=0,\quad \frac{\partial T}{\partial z_2}=0,\quad \frac{d\Omega}{dz_2}=0, \tag{3.9}$$

and

$$\frac{\partial \Phi}{\partial z_2} = \left\langle v_\Phi^\dagger, \left(2|v_1|^2 + v_1^2\right) v_\Phi \right\rangle = \left\langle v_\Phi^\dagger, 2|v_1|^2 v_\Phi \right\rangle + \left\langle v_\Phi^\dagger, v_1^2 v_\Phi \right\rangle . \tag{3.10}$$

Since the form of v_1 is still unknown at this point, progress can only be made by making assumptions as to how the radiation evolves. However, it is clear from Equation (3.10) that the term $\left\langle v_\Phi^\dagger, 2|v_1|^2 v_\Phi \right\rangle$ represents a significant contribution to the radiation drift since it contains the radiation power which has a nonzero mean.

3.2 Variational Formulation

The NLSE belongs to a class of evolution equations whose solutions can be recast as stationary points of a functional, referred to as the *action*, which for the NLSE presented here, takes the form of a double integral over a *Lagrangian density* [64]. By constraining the functional form of the solution to a parametrized family, integration over the transverse variable(s) yields an effective Lagrangian for the parameters, whose evolution is then dictated by the usual Euler-Lagrange equations [58]. With suitable modifications, one can also treat terms in the evolution equation that are not strictly variational in nature [60]. As is shown below, this approach results in the derivation of SODEs identical to the first order SPT approximation given in Section 2.1.1 for the NLSE with a hyperbolic secant (sech) pulse ansatz. In addition, when a functional form for the radiation is included in the ansatz, this method results in an augmented set of SODEs that approximate the second order effects of radiation on the phase evolution of the soliton. It should be noted however, that unlike the asymptotic approach of SPT, quantifying the accuracy of these variational approximations is difficult [65], and the ability to do so depends in large part on the form of approximating ansatz. Thus, these approximations are commonly validated through numerical comparisons with the PDE model.

3.2.1 Derivation of SODEs with Radiation Terms

Recall from Chapter 2 that the model for soliton propagation in a noisy environment is given as

$$i\frac{\partial u}{\partial z} + \frac{1}{2}\frac{\partial^2 u}{\partial t^2} + |u|^2 u = i\sigma n(z,t), \tag{3.11}$$

where $n(z,t)$ represents the ASE noise from amplification. Before beginning the variational procedure, it is advantageous to explicitly write the solution to Equation (3.11) as the sum of two parts, $u = u_{\mathrm{s}} + u_{\mathrm{r}}$, where u_{s} and u_{r} represent the solitonic and radiative components of the solution, respectively. Making this substitution into Equation (3.11) gives

$$\begin{aligned} i\frac{\partial u_{\mathrm{s}}}{\partial z} + i\frac{\partial u_{\mathrm{r}}}{\partial z} + \frac{1}{2}\frac{\partial^2 u_{\mathrm{s}}}{\partial t^2} + \frac{1}{2}\frac{\partial^2 u_{\mathrm{r}}}{\partial t^2} + |u_{\mathrm{s}}|^2 u_{\mathrm{s}} + 2|u_{\mathrm{r}}|^2 u_{\mathrm{s}} \\ + u_{\mathrm{s}}^2 \bar{u}_{\mathrm{r}} + u_{\mathrm{r}}^2 \bar{u}_{\mathrm{s}} + |u_{\mathrm{r}}|^2 u_{\mathrm{r}} + 2|u_{\mathrm{s}}|^2 u_{\mathrm{r}} = i\sigma n(z,t). \end{aligned} \tag{3.12}$$

As seen above, the cubic nonlinear term in Equation (3.11) produces six new nonlinear terms in the equation above, each representing a different nonlinear effect. The terms $|u_{\mathrm{s}}|^2 u_{\mathrm{s}}$ and $|u_{\mathrm{r}}|^2 u_{\mathrm{r}}$ represent self-phase modulation [1] (SPM) for the pulse and radiation, respectively, where the rate of phase rotation across each quantity is dependent on instantaneous intensity. The terms $2|u_{\mathrm{r}}|^2 u_{\mathrm{s}}$ and $2|u_{\mathrm{s}}|^2 u_{\mathrm{r}}$ represent cross-phase modulation [1] (CPM). Like SPM, these terms cause a nonlinear phase rotation in both the pulse and radiation. Unlike SPM however, this is due to the presence of the intensity of the opposing quantity. Finally, $u_{\mathrm{s}}^2 \bar{u}_{\mathrm{r}}$ and $u_{\mathrm{r}}^2 \bar{u}_{\mathrm{s}}$ represent mixing terms that shift energy among the various frequency components that satisfy a resonance condition involving the wave number [1].

The CPM terms are independent of radiation phase and therefore generically contribute to the phase rotation of the signal, however, the mixing terms are sensitive to the random phase of the radiation, and can therefore be neglected. In fact, comparing these terms to those of Equation (3.10), the radiation dependent rotation

of the soliton's phase is seen to originate from the CPM term $2|u_{\rm r}|^2 u_{\rm s}$, i.e., the CPM of the pulse from the presence of radiation. Also neglected is the SPM for the radiation, since it contributes at $O(u_{\rm r}^3)$, corresponding to $O(\sigma^3)$ in the power series expansion. With these terms excluded, the above equation can be separated into two coupled equations [1], one for the soliton and one for the radiation,

$$i\frac{\partial u_{\rm s}}{\partial z} + \frac{1}{2}\frac{\partial^2 u_{\rm s}}{\partial t^2} + |u_{\rm s}|^2 u_{\rm s} + 2|u_{\rm r}|^2 u_{\rm s} = i\sigma n_{\rm s}(z,t), \tag{3.13a}$$

$$i\frac{\partial u_{\rm r}}{\partial z} + \frac{1}{2}\frac{\partial^2 u_{\rm r}}{\partial t^2} + 2|u_{\rm s}|^2 u_{\rm r} = i\sigma n_{\rm r}(z,t). \tag{3.13b}$$

Notice that the noise term is also separated into a portion that drives the soliton equation, $n_{\rm s}(z,t)$, and portion that drives the radiation, $n_{\rm r}(z,t)$. The justification for this lies in the fact that to $O(\sigma)$, the only portion of the noise that results in changes to the parameters is that which projects onto the eigenbasis formed by the discrete eigenfunctions associated with the four soliton parameters. Because of this, the noise can be decomposed as

$$\sigma n(z,t) = \sigma n_{\rm s}(z,t) + \sigma n_{\rm r}(z,t). \tag{3.14a}$$

where

$$\begin{aligned} n_{\rm s}(z,t) = &\left\langle v_A^\dagger(z,t)\exp(i\Theta)\,, n(z,t)\right\rangle v_A(z,t)\exp(i\Theta) \\ &+ \left\langle v_T^\dagger(z,t)\exp(i\Theta)\,, n(z,t)\right\rangle v_T(z,t)\exp(i\Theta) \\ &+ \left\langle v_\Omega^\dagger\exp(i\Theta)\,, n(z,t)\right\rangle v_\Omega\exp(i\Theta) \\ &+ \left\langle \left(v_\Phi^\dagger - Tv_\Omega^\dagger\right)\exp(i\Theta)\,, n(z,t)\right\rangle (v_\Phi - Tv_\Omega)\exp(i\Theta)\,, \end{aligned} \tag{3.14b}$$

and

$$n_{\rm r}(z,t) = n(z,t) - n_{\rm s}(z,t). \tag{3.14c}$$

The Equations in (3.13) are together referred to as the coupled NLSE, which models the separate evolution of pulse and radiation, which are coupled through nonlinear interaction terms. Just as in the NLSE, the envelopes u_{s} and u_{r} are complex-valued, so the complete system must also include the complex conjugates of Equations (3.13a) and (3.13b).

As discussed above, the system of evolution equations given in (3.13) can be represented as a variational equation, where the solution of the system corresponds to a point in function space that extremizes an associated functional, referred to as the action, taking the form

$$J(u_{\mathrm{s}}, \bar{u}_{\mathrm{s}}, u_{\mathrm{r}}, \bar{u}_{\mathrm{r}}) = \int_0^{z_f} \int L_{\mathrm{den}}(u_{\mathrm{s}}, \bar{u}_{\mathrm{s}}, u_{\mathrm{r}}, \bar{u}_{\mathrm{r}})\, dtdz, \tag{3.15}$$

where L_{den} is the Lagrangian density given by

$$\begin{aligned} L_{\mathrm{den}} = \mathrm{Re}\left[i\bar{u}_{\mathrm{s}}\frac{\partial u_{\mathrm{s}}}{\partial z}\right] + \mathrm{Re}\left[i\bar{u}_{\mathrm{r}}\frac{\partial u_{\mathrm{r}}}{\partial z}\right] - \frac{1}{2}\left[\left|\frac{\partial u_{\mathrm{s}}}{\partial t}\right|^2 + \left|\frac{\partial u_{\mathrm{r}}}{\partial t}\right|^2\right] + \frac{1}{2}\left|u_{\mathrm{s}}\right|^4 \\ + 2\left|u_{\mathrm{s}}\right|^2\left|u_{\mathrm{r}}\right|^2 - \mathrm{Re}\left[\bar{u}_{\mathrm{s}}(in_{\mathrm{s}}(z,t))\right] - \mathrm{Re}\left[\bar{u}_{\mathrm{r}}(in_{\mathrm{r}}(z,t))\right]. \end{aligned} \tag{3.16}$$

To see more clearly how this formulation is related to the Equations (3.13), one can calculate the functional derivatives $\delta J/\delta\bar{u}_{\mathrm{k}}$ and $\delta J/\delta u_{\mathrm{k}}$, for $k = s$ and r, showing that

$$\frac{\delta J}{\delta\bar{u}_{\mathrm{k}}} = 0 \quad \text{and} \quad \frac{\delta J}{\delta u_{\mathrm{k}}} = 0 \tag{3.17}$$

are equivalent to an application of the Euler-Lagrange equations to L_{den}, i.e.,

$$\frac{\partial L_{\mathrm{den}}}{\partial\bar{u}_{\mathrm{k}}} = \frac{d}{dz}\frac{\partial L_{\mathrm{den}}}{\partial\left(\partial_z\bar{u}_{\mathrm{k}}\right)} + \frac{d}{dt}\frac{\partial L_{\mathrm{den}}}{\partial\left(\partial_t\bar{u}_{\mathrm{k}}\right)} \tag{3.18a}$$

and

$$\frac{\partial L_{\mathrm{den}}}{\partial u_{\mathrm{k}}} = \frac{d}{dz}\frac{\partial L_{\mathrm{den}}}{\partial\left(\partial_z u_{\mathrm{k}}\right)} + \frac{d}{dt}\frac{\partial L_{\mathrm{den}}}{\partial\left(\partial_t u_{\mathrm{k}}\right)}, \tag{3.18b}$$

where $\partial_j = \frac{\partial}{\partial j}$ for $j = z$ and t. Inserting the definition of L_{den} given in Equation (3.16) into the Euler-Lagrange equations of (3.18) shows that for $k = s$ they are equivalent to Equation (3.13a) and its complex conjugate, while for $k = r$ they are equivalent to Equation (3.13b) and its complex conjugate. Thus, the stationary points of $J(u_{\text{s}}, \bar{u}_{\text{s}}, u_{\text{r}}, \bar{u}_{\text{r}})$ given in Equation (3.15), correspond to solutions of the coupled NLSE.

Now that Equation (3.13) is in a variational form, an approximation can be made by restricting the infinite-dimensional general solutions $u_{\text{s}}(t, z)$ and $u_{\text{r}}(t, z)$ to fixed functional forms with a finite number of parameters that account for the z evolution, i.e., $u_{\text{s}} = f_{\text{s}}(t, \boldsymbol{p}_{\text{s}}(z))$ and $u_{\text{r}} = f_{\text{r}}(t, \boldsymbol{p}_{\text{r}}(z))$, where f_{s} and f_{r} are known functions with $\boldsymbol{p}_{\text{s}}(z)$ and $\boldsymbol{p}_{\text{r}}(z)$ representing vectors consisting of z dependent parameters for the soliton and radiation, respectively. Under this assumption, the t integral in Equation (3.15) can be calculated, leading to a reduced action of the form

$$J(\boldsymbol{p}_{\text{s}}, \dot{\boldsymbol{p}}_{\text{s}}, \boldsymbol{p}_{\text{r}}, \dot{\boldsymbol{p}}_{\text{r}}) = \int_0^{z_f} L_{\text{avg}} \left(\boldsymbol{p}_{\text{s}}(z), \dot{\boldsymbol{p}}_{\text{s}}(z), \boldsymbol{p}_{\text{r}}(z), \dot{\boldsymbol{p}}_{\text{r}}(z)\right) \, dz, \tag{3.19}$$

where L_{avg} is referred to as the averaged Lagrangian and is given by

$$L_{\text{avg}}(\boldsymbol{p}_{\text{s}}(z), \dot{\boldsymbol{p}}_{\text{s}}(z), \boldsymbol{p}_{\text{r}}(z), \dot{\boldsymbol{p}}_{\text{r}}(z)) = \int L_{\text{den}} \left(f_{\text{s}}(t, \boldsymbol{p}_{\text{s}}), f_{\text{r}}(t, \boldsymbol{p}_{\text{r}})\right) \, dt, \tag{3.20}$$

with the notation $\dot{x} = dx/dz$. The corresponding stationary points for this action are given by the Euler-Lagrange equations applied to each parameter,

$$\begin{aligned}
\frac{\partial L_{\text{avg}}}{\partial p_{\text{s},1}} &= \frac{d}{dz}\frac{\partial L_{\text{avg}}}{\partial \dot{p}_{\text{s},1}}, \qquad & \frac{\partial L_{\text{avg}}}{\partial p_{\text{r},1}} &= \frac{d}{dz}\frac{\partial L_{\text{avg}}}{\partial \dot{p}_{\text{r},1}}, \\
\frac{\partial L_{\text{avg}}}{\partial p_{\text{s},2}} &= \frac{d}{dz}\frac{\partial L_{\text{avg}}}{\partial \dot{p}_{\text{s},2}}, \qquad & \frac{\partial L_{\text{avg}}}{\partial p_{\text{r},2}} &= \frac{d}{dz}\frac{\partial L_{\text{avg}}}{\partial \dot{p}_{\text{r},2}}, \\
&\vdots & & \\
\frac{\partial L_{\text{avg}}}{\partial p_{\text{s},N_{\text{s}}}} &= \frac{d}{dz}\frac{\partial L_{\text{avg}}}{\partial \dot{p}_{\text{s},N_{\text{s}}}}, \qquad & \frac{\partial L_{\text{avg}}}{\partial p_{\text{r},N_{\text{r}}}} &= \frac{d}{dz}\frac{\partial L_{\text{avg}}}{\partial \dot{p}_{\text{r},N_{\text{r}}}},
\end{aligned} \tag{3.21}$$

where N_s and N_r are the number of parameters contained in p_s and p_r, respectively. Equations (3.21) can now be solved for each parameter $\dot{p}_{\mathrm{s},k}$ and $\dot{p}_{\mathrm{r},j}$, $k = 1, \ldots, N_\mathrm{s}$ and $j = 1, \ldots, N_\mathrm{r}$, giving first order SODEs for their evolution.

For the case of NLSE, the obvious choice for the pulse ansatz is the NLSE soliton, with the four free parameters discussed previously,

$$f_\mathrm{s}\left(t, A(z), T(z), \Omega(z), \Phi(z)\right) = A(z)\,\mathrm{sech}(A(z)(t - T(z)))\,\exp(\Omega(z)t + \Phi(z))\,. \quad (3.22)$$

The choice of radiation ansatz is much more difficult, primarily because it represents the evolution of noise which lacks a simple functional form. However, the second order SPT calculation in Section 3.1 implies that there exists a basis of eigenfunctions (radiation modes) capable of representing the radiation to arbitrary accuracy. This basis is given by the eigenfunctions of the linearized NLSE or equivalently by the squared Jost functions from inverse scattering theory [66, 62, 61], which are known as "dressed" modes, in that they consist of the classical Fourier modes with local variations around the position of the soliton to account for its presence. In the absence of a soliton, these modes collapse back to the standard Fourier modes, suggesting that they can be reasonably approximated as such. With this in mind, the radiation is approximated by a windowed Fourier decomposition,

$$f_\mathrm{r}(t, \alpha_1, \alpha_2, \ldots, \alpha_N) = \sum_{n=-N/2}^{N/2-1} \alpha_n(z)\exp(-i\omega_n t)\;\mathrm{H}(w - |t|)\,, \quad (3.23)$$

where the parameters $\alpha_n(z)$ are the Fourier coefficients of each mode. The decomposition has been truncated at N modes, corresponding either to a deliberate choice of radiation bandwidth or to the bandwidth limitation imposed by a computational method. For simplicity, the frequencies are taken to match the numerical frequencies of $\omega_n = n\pi/w$. With these two functional forms, the Lagrangian density and its

average are found to be

$$\begin{aligned}
L_{\text{den}} = & -A^2 \left[\operatorname{sech}(A[t-T])\right]^2 \left[\dot{\Omega}t + \dot{\Phi}\right] \\
& - \frac{A^4}{2} \left[\operatorname{sech}(A[t-T])\, \tanh(A(t-T))\right]^2 \\
& - \frac{A^2\Omega^2}{2} \left[\operatorname{sech}(A[t-T])\right]^2 + \frac{A^4}{2} \left[\operatorname{sech}(A[t-T])\right]^4 \\
& - \mathrm{H}(w-|t|) \sum_{j=-N/2}^{N/2-1} \mathrm{Re}\left[i\alpha_j \dot{\bar{\alpha}}_j\right] - \frac{1}{2}\mathrm{H}(w-|t|) \sum_{j=-N/2}^{N/2-1} \omega_j^2 \left|\alpha_j\right|^2 \\
& + 2A^2 \left[\operatorname{sech}(A[t-T])\right]^2 \sum_{j=-N/2}^{N/2-1} \left|\alpha_j\right|^2 + 2\mathrm{Re}\left[\bar{u}_{\text{s}}(in_{\text{s}})\right] \\
& + 2 \sum_{j=-N/2}^{N/2-1} \mathrm{Re}\left[\bar{\alpha}_j \exp(i\omega_j t)\, \mathrm{H}(w-|t|)\, (in_{\text{r}})\right]
\end{aligned} \tag{3.24}$$

and

$$\begin{aligned}
L_{\text{avg}} = & -2AT\dot{\Omega} - 2A\dot{\Phi} + \frac{A^3}{3} - A\Omega^2 - w \sum_{j=-N/2}^{N/2-1} \left[2i\mathrm{Re}\left[\alpha_j \dot{\bar{\alpha}}_j\right] + \omega_j^2 \left|\alpha_j\right|^2\right] \\
& + 4A \sum_{j=-N/2}^{N/2-1} \left|\alpha_j\right|^2 - 2\mathrm{Re}\left[\int \bar{u}_{\text{s}}(in_{\text{s}})\right] - 2 \sum_{j=-N/2}^{N/2-1} \mathrm{Re}\left[\int_{-w}^{w} \bar{\alpha}_j \exp(i\omega_j)\, (in_{\text{r}})\, dt\right].
\end{aligned} \tag{3.25}$$

Note that since the window $[-w, w]$ is much wider than the soliton width, the integrals involving both the soliton and radiation can be well approximated by integrating over the entire real line.

Applying the Euler-Lagrange equations to Equation (3.25) results in the following set of SODEs:

$$\dot{A} = \sigma\mathrm{Re}\left[\int i\frac{\partial \bar{u}_{\text{s}}}{\partial \Phi}\, n_{\text{s}}(z,t)\, dt\right] = \sigma\mathrm{Re}\left[\int \bar{v}_A^{\dagger}\, n(z,t)\, dt\right], \tag{3.26a}$$

$$\dot{\Omega} = \sigma\mathrm{Re}\left[\int (-i)\frac{1}{A}\frac{\partial \bar{u}_{\text{s}}}{\partial T}\, n_{\text{s}}(z,t)\, dt\right] = \sigma\mathrm{Re}\left[\int \bar{v}_{\Omega}^{\dagger}\, n(z,t)\, dt\right], \tag{3.26b}$$

$$\begin{aligned}
\dot{T} &= \Omega + \sigma\mathrm{Re}\left[\int \left[i\frac{1}{A}\frac{\partial \bar{u}_{\text{s}}}{\partial \Omega} - i\frac{T}{A}\frac{\partial \bar{u}_{\text{s}}}{\partial \Phi}\right] n_{\text{s}}(z,t)\, dt\right] \\
&= \Omega + \sigma\mathrm{Re}\left[\int \bar{v}_T^{\dagger}\, n(z,t)\, dt\right],
\end{aligned} \tag{3.26c}$$

$$\begin{aligned}\dot{\Phi} &= \frac{A^2-\Omega^2}{2} + 2\sum_{j=-N/2}^{N/2-1}|\alpha_j|^2 + \sigma\mathrm{Re}\left[\int\left[i\frac{T}{A}\frac{\partial\bar{u}_\mathrm{s}}{\partial T} - i\frac{\partial\bar{u}_\mathrm{s}}{\partial A}\right]\, n_\mathrm{s}(z,t)\,dt\right] \\ &= \frac{A^2-\Omega^2}{2} + 2\sum_{j=-N/2}^{N/2-1}|\alpha_j|^2 + \sigma\mathrm{Re}\left[\int\left(\bar{v}_\Phi^\dagger - T\bar{v}_\Omega^\dagger\right)\, n(z,t)\,dt\right],\end{aligned} \tag{3.26d}$$

where the radiation modes evolve according to,

$$\begin{aligned}\dot{\alpha}_j &= i\left(\frac{2A}{w} - \frac{\omega_j^2}{2}\right)\alpha_j + \frac{\sigma}{2w}\int_{-w}^{w}\exp(i\omega_j t)\; n_\mathrm{r}(z,t)\,dt \\ &\approx i\left(\frac{2A}{w} - \frac{\omega_j^2}{2}\right)\alpha_j + \frac{\sigma}{2w}\int_{-w}^{w}\exp(i\omega_j t)\; n(z,t)\,dt,\end{aligned} \tag{3.26e}$$

and

$$\begin{aligned}\dot{\bar{\alpha}}_j &= -i\left(\frac{2A}{w} - \frac{\omega_j^2}{2}\right)\bar{\alpha}_j + \frac{\sigma}{2w}\int_{-w}^{w}\exp(-i\omega_j t)\; \bar{n}_r(z,t)\,dt \\ &\approx -i\left(\frac{2A}{w} - \frac{\omega_j^2}{2}\right)\bar{\alpha}_j + \frac{\sigma}{2w}\int_{-w}^{w}\exp(-i\omega_j t)\; \bar{n}(z,t)\,dt.\end{aligned} \tag{3.26f}$$

Notice that the approximations in Equations (3.26e) and (3.26f) are due to the replacement of n_r with the full noise n. This results in the entire noise contributing to growth in the radiation, whereas the appropriate contribution should only be from the portion of the noise that is orthogonal to the eigenbasis formed from the discrete eigenfunctions associated with changes in the soliton parameters. In practice, however, the noise power contributing to perturbations in the soliton parameters is small, since it only incorporates four of the many modes comprising the total noise power, and thus this approximation is reasonable.

3.2.2 Discrete Noise Equations

The SODEs in (3.26) are almost identical to those derived through SPT, with the only exceptions being the phase Equation (3.26d), which includes a term representing the power in the radiation, and the evolution of the radiation modes themselves given by Equations (3.26e) and (3.26f). Using the definition of noise that corresponds to

discrete amplification, i.e.,

$$n(z,t) = \sum_{k=1}^{N_\mathrm{a}} n_k(t)\,\delta(z - kz_\mathrm{a})\,, \tag{3.27a}$$

with

$$\mathbb{E}[n_k(t)] = 0 \quad \text{and} \quad \mathbb{E}[n_k(t)\bar{n}_j(t')] = \delta(t-t')\delta_{kj}, \tag{3.27b}$$

the equations in (3.26) take the form

$$\frac{dA}{dz} = \sum_{k=1}^{N_\mathrm{a}} \sigma\mathrm{Re}\left[\int \bar{v}_A^\dagger \exp(-i\Theta)\, n_k(t)\, dt\right]\, \delta(z - kz_\mathrm{a}), \tag{3.28a}$$

$$\frac{dT}{dz} = \Omega + \sum_{k=1}^{N_\mathrm{a}} \sigma\mathrm{Re}\left[\int \bar{v}_T^\dagger \exp(-i\Theta)\, n_k(t)\, dt\right]\, \delta(z - kz_\mathrm{a}), \tag{3.28b}$$

$$\frac{d\Omega}{dz} = \sum_{k=1}^{N_\mathrm{a}} \sigma\mathrm{Re}\left[\int \bar{v}_\Omega^\dagger \exp(-i\Theta)\, n_k(t)\, dt\right]\, \delta(z - kz_\mathrm{a}), \tag{3.28c}$$

$$\begin{aligned}\frac{d\Phi}{dz} = \frac{A^2 - \Omega^2}{2} + 2P(\boldsymbol{\alpha}, \bar{\boldsymbol{\alpha}}) \\ + \sum_{k=1}^{N_\mathrm{a}} \sigma\mathrm{Re}\left[\int \left(\bar{v}_\Phi^\dagger - T\bar{v}_\Omega^\dagger\right) \exp(-i\Theta)\, n_k(t)\, dt\right]\, \delta(z - kz_\mathrm{a}),\end{aligned} \tag{3.28d}$$

where the radiation power is represented as $P(\boldsymbol{\alpha}, \bar{\boldsymbol{\alpha}})$. This term takes the place of the sum over the squared coefficients of the radiation modes, the evolution of which are given by the $2N$ Equations (3.26e) and (3.26f). However, as Equation (3.28d) indicates, only the radiation's *power* contributes to the evolution of the soliton's phase, which is independent of the individual phases of the radiation coefficients. Because of this, the phase rotation terms in Equations (3.26e) and (3.26f) can be neglected, leaving only the contributions of the stochastic terms, i.e.,

$$\dot{\alpha}_j = \frac{\sigma}{2w} \sum_{k=1}^{N_\mathrm{a}} \int_{-w}^{w} \exp(i\omega_j t)\; n_k(t)\, dt\, \delta(z - kz_\mathrm{a}) \tag{3.28e}$$

and

$$\dot{\bar{\alpha}}_j = \frac{\sigma}{2w} \sum_{k=1}^{N_\mathrm{a}} \int_{-w}^{w} \exp(-i\omega_j t)\ \bar{n}_k(t)\, dt\, \delta(z - kz_\mathrm{a}). \tag{3.28f}$$

In this form, these equation can be formally integrated, to give

$$\alpha_j(z) = \alpha_j(0) + \sum_{k=1}^{N_\mathrm{a}} \Delta\alpha_{j,k}\ \mathrm{H}(z - kz_\mathrm{a}) \tag{3.29a}$$

and

$$\bar{\alpha}_j(z) = \bar{\alpha}_j(0) + \sum_{k=1}^{N_\mathrm{a}} \Delta\bar{\alpha}_{j,k}\ \mathrm{H}(z - kz_\mathrm{a})\,, \tag{3.29b}$$

where the complex valued stochastic jump terms, $\Delta\alpha_{j,k}$ and $\Delta\bar{\alpha}_{j,k}$, are given by

$$\Delta\alpha_{j,k} = \frac{\sigma}{2w} \int_{-w}^{w} \exp(i\omega_j t)\ n_k(t)\, dt, \tag{3.30a}$$

and

$$\Delta\bar{\alpha}_{j,k} = \frac{\sigma}{2w} \int_{-w}^{w} \exp(-i\omega_j t)\ \bar{n}_k(t)\, dt. \tag{3.30b}$$

Using Equations (3.29a) and (3.29b), the radiation power term takes the form

$$\begin{aligned} P(\boldsymbol{\alpha}, \bar{\boldsymbol{\alpha}}) &= \sum_{j=-N/2}^{N/2-1} |\alpha_j(z)|^2 \\ &= \sum_{j=-N/2}^{N/2-1} |\alpha_j(0)|^2 + \sum_{j=-N/2}^{N/2-1} \sum_{k=1}^{N_\mathrm{a}} 2\mathrm{Re}\left[\alpha_j(0)\Delta\bar{\alpha}_{j,k}\right]\ \mathrm{H}(z - kz_\mathrm{a}) \\ &\quad + \sum_{j=-N/2}^{N/2-1} \sum_{k_1=1}^{N_\mathrm{a}} \sum_{k_2=1}^{N_\mathrm{a}} \Delta\alpha_{j,k_1}\Delta\bar{\alpha}_{j,k_2}\ \mathrm{H}(z - k_1 z_\mathrm{a})\ \mathrm{H}(z - k_2 z_\mathrm{a})\,, \end{aligned} \tag{3.31}$$

which can conveniently be written as

$$P(z) = P(0) + \Delta P(z), \tag{3.32a}$$

where

$$P(0) = \sum_{j=-N/2}^{N/2-1} |\alpha_j(0)|^2 \tag{3.32b}$$

and

$$\begin{aligned} \Delta P(z) = & \sum_{j=-N/2}^{N/2-1} \sum_{k=1}^{N_\mathrm{a}} 2\mathrm{Re}\left[\alpha_j(0)\Delta\bar{\alpha}_{j,k}\right] \ \mathrm{H}(z - kz_\mathrm{a}) \\ & + \sum_{j=-N/2}^{N/2-1} \sum_{k_1=1}^{N_\mathrm{a}} \sum_{k_2=1}^{N_\mathrm{a}} \Delta\alpha_{j,k_1}\Delta\bar{\alpha}_{j,k_2} \ \mathrm{H}(z - k_1 z_\mathrm{a}) \ \mathrm{H}(z - k_2 z_\mathrm{a}) \,. \end{aligned} \tag{3.32c}$$

The SODEs for the soliton parameters can also be integrated giving

$$A(z) = A(0) + \sum_{k=1}^{N_\mathrm{a}} \Delta A_k \ \mathrm{H}(z - kz_\mathrm{a}) \,, \tag{3.33a}$$

$$\begin{aligned} T(z) = T(0) + \Omega(0)z & + \sum_{k=1}^{N_\mathrm{a}} \Delta\Omega_k \left(z - kz_\mathrm{a}\right) \mathrm{H}(z - kz_\mathrm{a}) \\ & + \sum_{k=1}^{N_\mathrm{a}} \Delta T_k \ \mathrm{H}(z - kz_\mathrm{a}) \,, \end{aligned} \tag{3.33b}$$

$$\Omega(z) = \Omega(0) + \sum_{k=1}^{N_\mathrm{a}} \Delta\Omega_k \ \mathrm{H}(z - kz_\mathrm{a}) \,, \tag{3.33c}$$

$$
\begin{aligned}
\Phi(z) = \Phi(0) &+ \frac{A(0)^2}{2} z + \sum_{k=1}^{N_\mathrm{a}} A(0)\Delta A_k \left(z - kz_\mathrm{a}\right) \mathrm{H}(z - kz_\mathrm{a}) \\
&+ \frac{1}{2} \sum_{k=1}^{N_\mathrm{a}} (\Delta A_k)^2 \left(z - kz_\mathrm{a}\right) \mathrm{H}(z - kz_\mathrm{a}) \\
&+ \frac{1}{2} \sum_{k=1}^{N_\mathrm{a}} \sum_{j=1}^{k-1} \Delta A_k \Delta A_j \left(z - kz_\mathrm{a}\right) \mathrm{H}(z - kz_\mathrm{a}) \\
&+ \frac{1}{2} \sum_{k=1}^{N_\mathrm{a}} \sum_{j=k+1}^{N_\mathrm{a}} \Delta A_k \Delta A_j \left(z - jz_\mathrm{a}\right) \mathrm{H}(z - jz_\mathrm{a}) \\
&- \frac{\Omega(0)^2}{2} z - \sum_{k=1}^{N_\mathrm{a}} \Omega(0)\Delta\Omega_k \left(z - kz_\mathrm{a}\right) \mathrm{H}(z - kz_\mathrm{a}) \\
&- \frac{1}{2} \sum_{k=1}^{N_\mathrm{a}} (\Delta\Omega_k)^2 \left(z - kz_\mathrm{a}\right) \mathrm{H}(z - jz_\mathrm{a}) \\
&- \frac{1}{2} \sum_{k=1}^{N_\mathrm{a}} \sum_{j=1}^{k-1} \Delta\Omega_k \Delta\Omega_j \left(z - kz_\mathrm{a}\right) \mathrm{H}(z - kz_\mathrm{a}) \\
&- \frac{1}{2} \sum_{k=1}^{N_\mathrm{a}} \sum_{j=k+1}^{N_\mathrm{a}} \Delta\Omega_k \Delta\Omega_j \left(z - jz_\mathrm{a}\right) \mathrm{H}(z - jz_\mathrm{a}) \\
&+ 2P(0)z + 2I_{\Delta P}(z) + \sum_{k=1}^{N_\mathrm{a}} \Delta\Phi_k \, \mathrm{H}(z - kz_\mathrm{a}) \,,
\end{aligned}
\tag{3.33d}
$$

where

$$
\begin{aligned}
I_{\Delta P}(z) &= \int_0^z \Delta P(z') \, dz' \\
&= \sum_{j=-N/2}^{N/2-1} \sum_{k=1}^{N_\mathrm{a}} 2\mathrm{Re}\left[\alpha_j(0)\Delta\bar{\alpha}_{j,k}\right] \left(z - kz_\mathrm{a}\right) \mathrm{H}(z - kz_\mathrm{a}) \\
&\quad + \sum_{j=-N/2}^{N/2-1} \sum_{k=1}^{N_\mathrm{a}} \left|\Delta\alpha_{j,k}\right|^2 \left(z - kz_\mathrm{a}\right) \mathrm{H}(z - kz_\mathrm{a}) \\
&\quad + \sum_{j=-N/2}^{N/2-1} \sum_{k_1=1}^{N_\mathrm{a}} \sum_{k_2=1}^{k_1-1} \Delta\alpha_{j,k_1}\Delta\bar{\alpha}_{j,k_2} \left(z - k_1 z_\mathrm{a}\right) \mathrm{H}(z - k_1 z_\mathrm{a}) \\
&\quad + \sum_{j=-N/2}^{N/2-1} \sum_{k_1=1}^{N_\mathrm{a}} \sum_{k_2=k_1+1}^{N_\mathrm{a}} \Delta\alpha_{j,k_1}\Delta\bar{\alpha}_{j,k_2} \left(z - k_2 z_\mathrm{a}\right) \mathrm{H}(z - k_2 z_\mathrm{a}) \,.
\end{aligned}
\tag{3.33e}
$$

The stochastic jump perturbations of the soliton parameters are equivalent to those given in Chapter 2, i.e.,

$$
\begin{aligned}
\Delta X_k = \sigma\mathrm{Re}\Bigg[\int \Big(\bar{v}_X^\dagger(kz_\mathrm{a}, t) - T(kz_\mathrm{a})\bar{v}_\Omega^\dagger(kz_\mathrm{a}, t)\delta_{X,\Phi}\Big) \times \\
\exp(-i\Theta(kz_\mathrm{a})) \, n_k(t) \, dt\Bigg] \,,
\end{aligned}
\tag{3.34}
$$

for $X = A, \Omega, T$ and Φ, where it is recalled that the functions $v_X^\dagger$ are the adjoint eigenfunctions, associated with the adjoint operator $L^\dagger$, which are defined in Equations (2.18). All of the stochastic jump terms above are mean zero and, assuming that the parameter values immediately after the kth amplifier are known, the stochastic jumps at amplifier $k+1$ have variances given by

$$\begin{aligned}
\mathbb{E}\big[(\Delta A_{k+1})^2\big] &= \frac{\sigma^2}{2}\|v_A^\dagger(kz_{\mathrm{a}})\|^2 = \sigma^2 A_k, \\
\mathbb{E}\big[(\Delta T_{k+1})^2\big] &= \frac{\sigma^2}{2}\|v_T^\dagger(kz_{\mathrm{a}})\|^2 = \frac{\sigma^2\pi^2}{12A_k^3}, \\
\mathbb{E}\big[(\Delta \Omega_{k+1})^2\big] &= \frac{\sigma^2}{2}\|v_\Omega^\dagger(kz_{\mathrm{a}})\|^2 = \frac{\sigma^2 A_k}{3}, \\
\mathbb{E}\big[(\Delta \Phi_{k+1})^2\big] &= \frac{\sigma^2}{2}\|v_\Phi^\dagger(kz_{\mathrm{a}}) - T_k v_\Omega^\dagger(kz_{\mathrm{a}})\|^2 \\
&= \sigma^2\frac{12\left(1+\left(T_k+\Omega_k z_{\mathrm{a}}\right)^2 A_k^2\right)+\pi^2}{36A_k}, \\
\mathbb{E}[\Delta\alpha_{j,k+1}\Delta\bar{\alpha}_{j,k+1}] &= \frac{\sigma^2}{2w},
\end{aligned} \tag{3.35}$$

where it is noted that

$$\mathbb{E}\big[(\Delta\alpha_{j,k+1})^2\big] = \mathbb{E}\big[(\Delta\bar{\alpha}_{j,k+1})^2\big] = 0. \tag{3.36}$$

3.3 Linearized Equations and Parameter Statistics

Note that, although presented in closed form, the parameter statistics at a given amplifier as given in Equations (3.35) are still dependent on the value of the parameters evaluated at all prior amplification points. Thus, to make quantitative measures of these statistics, one must still numerically solve the SODEs in (3.28). However, approximate values for these statistics can be found if the equations are first linearized around the initial parameter values, which for all simulations here are taken to be $A(0) = 1$, $\Omega(0) = T(0) = \Phi(0) = 0$ and $\alpha_j(0) = 0$ for $j = N/2, \ldots, N/2-1$. Doing this gives random jumps that are independent of prior changes in the parameters,

which implies that parameter variances are now independent of amplifier index, i.e.,

$$\begin{aligned} \mathbb{E}\big[(\Delta A_k)^2\big] &= \sigma^2, \quad \mathbb{E}\big[(\Delta T_k)^2\big] = \frac{\sigma^2\pi^2}{12}, \\ \mathbb{E}\big[(\Delta \Omega_k)^2\big] &= \frac{\sigma^2}{3}, \quad \mathbb{E}\big[(\Delta \Phi_k)^2\big] = \sigma^2\frac{12+\pi^2}{36}, \end{aligned} \tag{3.37}$$

where as before

$$\mathbb{E}[\Delta\alpha_{j,k}\Delta\bar{\alpha}_{j,k}] = \frac{\sigma^2}{2w}. \tag{3.38}$$

After neglecting higher order terms, the solutions in (3.33) become

$$A(z) = 1 + \sum_{k=1}^{N_\mathrm{a}} \Delta A_k \ \mathrm{H}(z - kz_\mathrm{a}) \,, \tag{3.39a}$$

$$T(z) = \sum_{k=1}^{N_\mathrm{a}} \Delta\Omega_k \, (z - kz_\mathrm{a}) \ \mathrm{H}(z - kz_\mathrm{a}) + \sum_{k=1}^{N_\mathrm{a}} \Delta T_k \ \mathrm{H}(z - kz_\mathrm{a}) \,, \tag{3.39b}$$

$$\Omega(z) = \sum_{k=1}^{N_\mathrm{a}} \Delta\Omega_k \ \mathrm{H}(z - kz_\mathrm{a}) \,, \tag{3.39c}$$

$$\Phi(z) = \frac{1}{2}z + \sum_{k=1}^{N_\mathrm{a}} \Delta A_k \, (z - kz_\mathrm{a}) \, \mathrm{H}(z - kz_\mathrm{a}) + 2\tilde{I}_{\Delta P}(z) + \sum_{k=1}^{N_\mathrm{a}} \Delta\Phi_k \ \mathrm{H}(z - kz_\mathrm{a}) \,, \tag{3.39d}$$

where

$$\tilde{I}_{\Delta P}(z) = \sum_{j=-N/2}^{N/2-1} \sum_{k=1}^{N_\mathrm{a}} |\Delta\alpha_{j,k}|^2 \ (z - kz_\mathrm{a}) \, \mathrm{H}(z - kz_\mathrm{a}) \,. \tag{3.39e}$$

It is important to note that, although the term $\tilde{I}_{\Delta P}(z)$ appears to be $O(\sigma^2)$, the aggregation over all radiation modes reduces the effective order of this term to $O(\sigma)$.

Using the approximate solutions given in Equations (3.39) evaluated at the end of the transmission line, $z = N_\mathrm{a} z_\mathrm{a}$, the PDF's of the parameters A, Ω, and T are calculated as

$$A(z_f) \sim \mathcal{N}\big(1, \sigma^2 N_\mathrm{a}\big) \,, \tag{3.40a}$$

$$T(z_f) \sim \mathcal{N}\left(0, \sigma^2 \left(z_{\mathrm{a}}^2 N_{\mathrm{a}}(N_{\mathrm{a}}-1)\frac{2N_{\mathrm{a}}-1}{18} + \frac{\pi^2}{12}\right)\right), \tag{3.40b}$$

and

$$\Omega(z_f) \sim \mathcal{N}\left(0, \sigma^2 \frac{N_{\mathrm{a}}}{3}\right), \tag{3.40c}$$

where $\mathcal{N}(\mu, \sigma^2)$ represents the normal distribution with mean μ and variance σ^2. The distribution of the phase, however, is now altered by the radiation term,

$$\tilde{I}_{\Delta P}(z_f) = \sum_{k=1}^{N_{\mathrm{a}}} \sum_{j=-N/2}^{N/2-1} |\Delta\alpha_{j,k}|^2 \; z_{\mathrm{a}}(N_{\mathrm{a}}-k), \tag{3.41}$$

which follows a χ^2 distribution and thus contributes to both the mean and variance of final phase values. To simplify this term, first notice that since $\Delta\alpha_{j,k}$ and $\Delta\bar{\alpha}_{j,k}$ are normally distributed with mean zero and a cross variance of $\sigma^2/2w$ (recall Equation (3.38)), the inner summation in (3.41) can be written as

$$\sum_{j=-N/2}^{N/2-1} |\Delta\alpha_{j,k}|^2 = \frac{\sigma^2}{2w}\mathcal{X}_k, \tag{3.42}$$

where $\mathcal{X}_k$ is a χ^2_N-distributed random variable with N degrees of freedom, indexed by the amplifier index k. In addition, since N is assumed to be large, the central limit theorem allows $\mathcal{X}_k$ to be well approximated by a normal distribution, i.e.,

$$N \to \infty \implies \chi^2_N \to \mathcal{N}(N, 2N)\,. \tag{3.43}$$

Using this, the contribution from radiation can be written as a normal RV with

$$\mathbb{E}\left[\tilde{I}_{\Delta P}(z_f)\right] \approx \frac{\sigma^2}{2w} z_{\mathrm{a}} N \frac{N_{\mathrm{a}}(N_{\mathrm{a}}-1)}{2}, \tag{3.44a}$$

and

$$\mathbb{V}\left[\tilde{I}_{\Delta P}(z_f)\right] \approx \left(\frac{\sigma^2}{2w} z_{\mathrm{a}}\right)^2 2N \frac{N_{\mathrm{a}}(N_{\mathrm{a}}-1)(2N_{\mathrm{a}}-1)}{6}. \tag{3.44b}$$

Thus,

$$\Phi(z_f) \sim \mathcal{N}(\mu_\Phi, \sigma_\Phi)\,, \tag{3.45a}$$

where

$$\mu_\Phi = \frac{N_\mathrm{a} z_\mathrm{a}}{2} + \frac{\sigma^2}{2w} z_\mathrm{a} N N_\mathrm{a}(N_\mathrm{a} - 1), \tag{3.45b}$$

and

$$\sigma_\Phi = \sigma^2 z_\mathrm{a}^2 \frac{(N_\mathrm{a}^2 - N_\mathrm{a})(2N_\mathrm{a} - 1)}{6} \left(2N\frac{\sigma^2}{w^2} + 1\right) + \sigma^2 N_\mathrm{a} \frac{12 + \pi^2}{36}. \tag{3.45c}$$

From these results, the contributions to the mean and variance of the soliton's phase originating from direct phase perturbations, amplitude perturbations and radiation perturbations can be individually identified. These are given in Table 3.1, along with with their quantitative approximations, which were calculated from the parameter values used in the simulations presented in Chapter 2.

3.4 Results and Discussion

To verify the analysis above, 10^6 standard MC simulations were conducted using the stochastic NLSE in (3.11) and the SODEs in (3.28) using the same parameter values as in the simulations of Chapter 2. The results these simulations are displayed in Figure 3.1, where the solid curves and markers correspond to the results of the stochastic NLSE and the SODEs, respectively. The top left and top right plots compare the mean and variance of the phase parameter, respectively, which are plotted as functions of the normalized transmission length using 128, 256, 512 and 1024 simulation modes. Alternatively, the bottom left and bottom right plots compare the mean and variance of the phase parameter, respectively, which are plotted as functions of the number of radiation modes at $z = 5$, 10, 15 and 20. In all of these plots, the mean and variance

Table 3.1 Approximate Contributions to the Nonlinear Soliton Phase Rotation

Source:	Phase	Amplitude	Radiation
Mean Contribution	N/A	$\frac{N_{\mathrm{a}} z_{\mathrm{a}}}{2}$	$\frac{\sigma^2}{2w} z_{\mathrm{a}} N (N_{\mathrm{a}}^2 - N_{\mathrm{a}})$
Value (128 Modes)	0	10	0.8
Value (256 Modes)	0	10	1.6
Value (512 Modes)	0	10	3.2
Value (1024 Modes)	0	10	6.4
Variance Contribution	$\sigma^2 N_{\mathrm{a}} \frac{12+\pi^2}{36}$	$\sigma^2 z_{\mathrm{a}}^2 \frac{(N_{\mathrm{a}}^2 - N_{\mathrm{a}})(2N_{\mathrm{a}}-1)}{6}$	$\sigma^4 z_{\mathrm{a}}^2 N \frac{(N_{\mathrm{a}}^2 - N_{\mathrm{a}})(2N_{\mathrm{a}}-1)}{3w^2}$
Value (128 Modes)	7.5×10^{-3}	1.7	6.7×10^{-5}
Value (256 Modes)	7.5×10^{-3}	1.7	1.3×10^{-4}
Value (512 Modes)	7.5×10^{-3}	1.7	2.7×10^{-4}
Value (1024 Modes)	7.5×10^{-3}	1.7	5.4×10^{-4}

calculations from the linearized SODEs of SPT are plotted using the black dashed line. The results of this figure confirm the predictions of the linearized theory above, showing that the effects of radiation are primarily seen in the mean evolution of the phase, while only making a small contribution to variance.

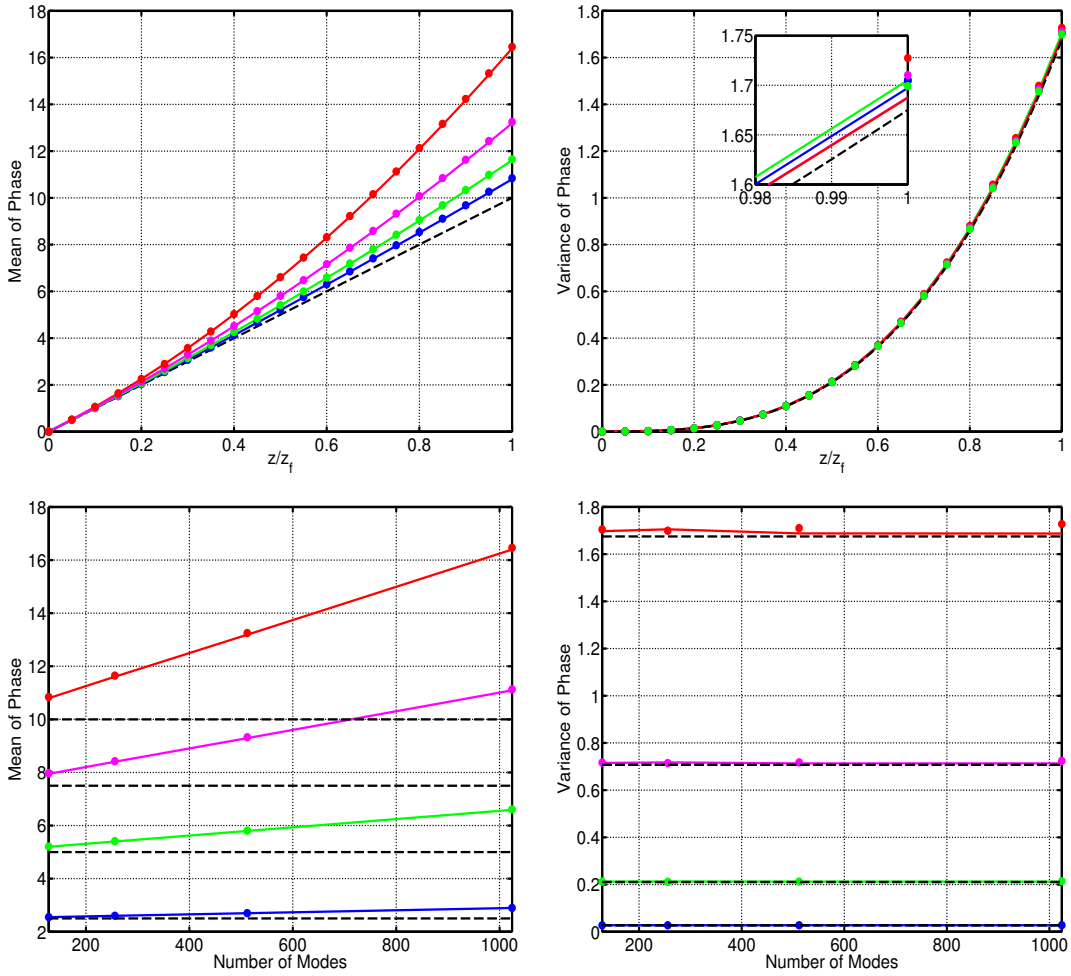


Figure 3.1 Top: The blue, green, magenta and red lines are the mean (Left) and variance (Right) of the phase parameter as a function of the normalized transmission length as calculated from 1×10^6 MC runs of the stochastic NLSE in (3.11) (solid curves) and the SODEs in (3.28) (markers), using 128, 256, 512 and 1024 simulation modes, respectively. Bottom: The blue, green, magenta and red lines are the mean (Left) and variance (Right) of the phase parameter as a function of the number of simulation modes as calculated from 1×10^6 MC runs of the stochastic NLSE in (3.11) (solid curves) and the SODEs in (3.28) (markers), using $z = 5$, 10, 15 and 20, respectively. In all plots, the circle markers are the corresponding solutions of the calculated from 1×10^6 MC runs.

3.5 Summary

This chapter began by formally extending the linear SPT of Chapter 2 to second order, giving the first analytical indication that dispersive radiation can make an effective first order contribution to the nonlinear evolution of a soliton's phase. Although the

initial calculation was based on a formal application of the method of multiple scales, it can be tied explicitly to an eigenfunction expansion of the radiation in terms of the squared eigenfunctions from inverse scattering theory [61, 62, 63]. Motivated by the results of this formal calculation, Section 3.2 introduced a less complex variational reformulation, leading to a modified low-dimensional reduction with additional terms that correctly account for the presence of dispersive radiation. From linear analysis of these equations, the individual contributions to both the mean and variance of soliton's phase evolution, as originating from direct perturbations to the phase and indirect perturbations in the amplitude and radiation which integrate to phase changes, were identified. The final section presented numerical verification that this new reduced system correctly approximates the full evolution of the soliton's phase in the presence of radiation and thus provides an improved basis around which a more efficient IS scheme can be constructed, which is the topic investigated in the next chapter.

CHAPTER 4

IMPROVEMENTS TO IMPORTANCE SAMPLING FOR NLSE

In this chapter, the ISMC method presented in Chapter 2 for the investigation of large phase deviations is modified to account for the presence of dispersive radiation. The approach taken here is similar to the original approach outlined in [23] and reviewed in Chapter 2, however, under the newly derived ODEs of Chapter 3, the evolution of the phase includes an additional term that accounts for the impact of dispersive radiation. This adds an additional layer of complexity in the calculation of the correct biasing distributions, which now must include changes to all four soliton parameters *and* to the dispersive radiation.

4.1 Evolution of the Phase Parameter with Radiation

The biasing for the phase parameter is now calculated using the modified ODEs for the reduced soliton dynamics derived in Chapter 3, which under a general noise term, take the form

$$\frac{dA}{dz} = \sigma \mathrm{Re}\left[\int \bar{v}_A^\dagger(z,t)\exp(-i\Phi(z))\, n(z,t)\, dt\right], \tag{4.1a}$$

$$\frac{d\Omega}{dz} = \sigma \mathrm{Re}\left[\int \bar{v}_\Omega^\dagger(z,t)\exp(-i\Phi(z))\, n(z,t)\, dt\right], \tag{4.1b}$$

$$\frac{dT}{dz} = \Omega + \sigma \mathrm{Re}\left[\int \bar{v}_T^\dagger(z,t)\exp(-i\Phi(z))\, n(z,t)\, dt\right], \tag{4.1c}$$

$$\begin{aligned}\frac{d\Phi}{dz} &= \frac{1}{2}(A^2-\Omega^2) + 2\sum_n |\alpha_n|^2 \\ &\quad + \sigma \mathrm{Re}\left[\int \left[\bar{v}_\Phi^\dagger(z,t) - T\bar{v}_\Omega^\dagger(z,t)\right] \exp(-i\Phi(z))\, n(z,t)\, dt\right]\end{aligned} \tag{4.1d}$$

$$\frac{d\alpha_k}{dz} = i\left(\frac{2A^2}{w} - \frac{1}{2}\omega_k^2\right)\alpha_k + \frac{\sigma}{2w}\int \exp(i\omega_k t)\ \mathrm{H}(w-|t|)\, n(z,t)\, dt, \tag{4.1e}$$

and

$$\frac{d\bar{\alpha}_k}{dz} = -i\left(\frac{2A^2}{w} - \frac{1}{2}\omega_k^2\right)\bar{\alpha}_k + \frac{\sigma}{2w}\int \exp(-i\omega_k t)\ \mathrm{H}(w-|t|)\, \bar{n}(z,t)\, dt. \tag{4.1f}$$

Recall that these equations were derived under the assumption that the radiation is well represented by a sum over Fourier coefficients $\alpha_n(z)$, i.e.,

$$R(z,t) = \sum_{n=-N/2}^{N/2-1} \alpha_n(z) \exp(-i\omega_n t)\ \mathrm{H}(w-|t|)\,, \tag{4.2}$$

each of which evolves according to the Equations (4.1e) and (4.1f). These equations indicate that the radiation evolution consists of two terms; (i) a phase rotation, reflecting the effects of dispersion and cross phase modulation, and (ii) a stochastic perturbation in the form of a projection of the noise onto the same Fourier basis used to represent the radiation in Equation (4.2).

It is important to note that, as previously discussed, the radiation *cannot* be completely represented by the Fourier expansion in Equation (4.2), since by definition, the radiation must be orthogonal to the basis of discrete eigenfunctions, i.e., those modes forming the projection operator for capturing leading order perturbations to the soliton parameters, which is is clearly not the case under the representation in Equation (4.2). This fact is explicit in the construction of Equations (4.1e) and (4.1f), which originally used the noise term n_{r} to indicate that the radiation was only forced by the portion of noise orthogonal to the four discrete eigenfunctions. In addition, the correct expansion would be in *dressed* Fourier modes, which consist of standard Fourier modes with additional local variations that account for presence of the soliton [61, 62, 63].

It is also important to note that the radiation term in Equation (4.1d), is in the form of a sum over squared Fourier coefficients, implying that each radiation mode

contributes equally to the total mean phase drift of the soliton. This is counter intuitive to the expectation that the effects of radiation should saturate as higher order modes are included. However, in the course of reducing Equations (4.1) to their present form, all dependencies on the modal index are lost, resulting in the radiation power term in Equation (4.1d). This is consistent with the assumption that the soliton evolves in a *radiation bath*, i.e., periodic domain, with radiation continuously entering and leaving the computational window as the total background radiation power increases linearly with the number of amplification points that the pulse passes. If this assumption was relaxed to allow the radiation to leave the computational window, one would need to rederive ODEs for each radiation mode that account for the loss in radiation power, much like what was done by Kath and Smyth [67]. Finally, it should be pointed out that this model depends critically on the uniformity of the noise over the entire domain, whether it be periodic or the real line. Dispersive radiation generated by a localized disturbance, such as that considered by Kath and Smyth, would not build up a mean background level in the same way.

4.1.1 Further Simplifications

The reduced system in (4.1) has four equations for the evolution of the soliton parameters, and $2N$ equations for the evolution of the real and imaginary parts of the radiation. However, as previously discussed, the term that couples the radiation to the evolution of the soliton parameters is a sum over the *squared modulus* of the radiation coefficients, i.e., the *radiation's power*, that appears in the Equation (4.1d). This suggest that, rather than accounting for $2N$ individual radiation modes, a more economical way of representing the radiation is to use a single parameter for the radiation's power,

$$P(z) = \frac{1}{2w} \int |R|^2 \, dt = \sum_{n=-N/2}^{N/2-1} |\alpha_n(z)|^2, \tag{4.3}$$

which takes place of the sum in Equation (4.1d) over the $2N$ radiation coefficients. In addition, the soliton's phase evolution is independent of the phase of the radiation, implying that the phase rotation terms in Equations (4.1e) and (4.1f) can be neglected. Using this, the evolution of the parameter representing the radiation's power in Equation (4.3) is found to be

$$\frac{dP}{dz} = \frac{1}{2w}\int \frac{\partial R}{\partial z}\bar{R} + R\frac{\partial \bar{R}}{\partial z}\,dt = \frac{1}{w}\mathrm{Re}\left[\int \bar{R}(z,t)n(z,t)\,dt\right]. \tag{4.4}$$

Recalling is the variance in the discretized noise and approximating the noise as continuous, i.e.,

$$\mathbb{E}[n(z,t)\bar{n}(z',t')] = \delta(t-t')\delta(z-z'), \tag{4.5}$$

The mean of $P(z)$ is calculated as

$$\mathbb{E}[P(z)] = \sigma_{\mathrm{n}}^2 z, \tag{4.6}$$

where is it recalled that $\sigma_{\mathrm{n}}^2 = \sigma^2 N/2w = \sigma^2/\Delta t$ is the modified noise strength. From this, it is immediately apparent that

$$\mathbb{E}\left[\frac{dP}{dz}\right] = \frac{2}{2w}\mathbb{E}\left[\mathrm{Re}\left[\int \bar{R}(z,t)n(z,t)\,dt\right]\right] = \sigma_{\mathrm{n}}^2. \tag{4.7}$$

Note that both equations (4.6) and (4.7) have alternative forms under the discrete noise representation, that are given in Equations (3.35).

4.2 Biasing the Radiation Power Parameter

Since the radiation power is now represented as a single parameter, the radiation can be biased in much the same way as each soliton parameter was biased in the original approach. Thus, the radiation biasing will be separated into two distinct steps, the first of which involves finding the optimal biasing mode which, when added

to the radiation, produces a (linear) unit change in the radiation power parameter P. The second step is to determine the weighting functions for the four soliton parameters *and* the new parameter P, such that the evolution of these parameters through parameter space follows an optimal path, which is determined through a constrained optimization problem that includes changes in the radiation power.

4.2.1 Biasing Mode for Radiation Power Parameter

Equation (4.4) implies that the instantaneous form of the radiation plays the role of an "eigenfunction" for the radiation power. Although this is not surprising, it provides an easy way of incorporating the radiation in the biasing scheme without dealing with each radiation mode separately. Just like the four soliton parameter modes, the likelihood of realizing a mode that results in a shift in the radiation power is given by the multivariate Gaussian distribution in Section 2.3.1. Likewise, the most probable biasing mode for the radiation is equivalent to the one with the smallest L^2-norm. Combining this with the constraint of imparting a unit change in the radiation power gives a Lagrange multiplier problem of the form

$$\begin{aligned} M_P(f_P, \bar{f}_P) &= \int |f_P|^2 dt + \lambda \left[\frac{dP}{dz} - 1 \right] \\ &= \int |f_P|^2 dt + \lambda \left[\int \frac{\bar{R}(z,t')}{2w} f_P + \frac{R(z,t')}{2w} \bar{f}_P \, dt' - 1 \right], \end{aligned} \tag{4.8}$$

where Equation (4.4) was used in expressing the constraint. Taking variations of Equation (4.8) over f_P and $\bar{f}_P$ gives a solution of

$$f_P(z,t) = \frac{R(z,t)}{2P(z)}, \tag{4.9}$$

which, as expected, is seen to be a normalized form of the radiation.

4.2.2 Biasing Path for Radiation Power Parameter

By including a weighted version of the biasing radiation mode in Equation (4.9), the continuous representation of the biasing vector takes the form

$$\begin{aligned} f(z,t) = \Delta_A(z) f_A(z,t) + \Delta_\Omega(z) f_\Omega(z,t) \\ + \Delta_T(z) f_T(z,t) + \Delta_\Phi(z) f_\Phi(z,t) + \Delta_P(z) f_P(z,t), \end{aligned} \tag{4.10}$$

where as previous calculated

$$f_X(z,t) = \frac{\left(v_X^\dagger(z,t) - T v_\Omega^\dagger(z,t)\delta_{X\Phi}\right)}{\|v_X^\dagger(z,t) - T v_\Omega^\dagger(z,t)\delta_{X\Phi}\|^2} \exp(i\Theta)\,. \tag{4.11}$$

The quantity that must be minimized is still given by the cumulative L^2-norm, i.e.,

$$S_\Phi = \int_0^{z_\mathrm{f}} \|f(z,t)\|^2\, dz, \tag{4.12}$$

which now includes the weighted biasing mode for the radiation parameter. However, the inclusion of this mode only effects the evolution of the phase, and thus, any assumptions made in the original phase biasing problem of Section 2.3.1 are still valid. In particular, the assumption that both the timing and frequency parameters, i.e., T and Ω, are limited in their capacity to change the phase parameter still holds, implying that $\Delta_\Omega(z)$ and $\Delta_T(z)$ can be neglected and $T_\mathrm{opt}(z) = \Omega_\mathrm{opt}(z) = 0$, since both have initial values of zero in all simulations. Using this assumption, the cumulative L^2-norm takes the form

$$\begin{aligned} S_\Phi = \int_0^{z_\mathrm{f}} (\Delta_A(z))^2 \|f_A(z,t)\|^2 + (\Delta_\Phi(z))^2 \|f_\Phi(z,t)\|^2 \\ + (\Delta_P(z))^2 \|f_P(z,t)\|^2\, dz. \end{aligned} \tag{4.13}$$

After applying the biasing, the mean zero noise $\sigma n(z,t)$ takes the form $\sigma n(z,t) + f(z,t)$, where $f(z,t)$ is the biasing vector given in (4.10). Thus, under the biased

noise, the evolution equations take the form

$$\begin{aligned}
\dot{A} &= n_A(z) + \Delta_A(z), \\
\dot{\Omega} &= n_\Omega(z) + \Delta_\Omega(z) + \Delta_\Omega(z)\Delta_\Phi(z)\frac{T\|v_\Omega\|^2}{\|v_\Phi - Tv_\Omega\|^2} \\
\dot{T} &= \Omega + n_T(z) + \Delta_T(z), \\
\dot{\Phi} &= \frac{A^2 - \Omega^2}{2} + 2P + n_\Phi(z) + \Delta_\Phi(z), \\
\dot{P} &= n_P(z) + \Delta_P(z),
\end{aligned} \tag{4.14}$$

where $n_X(z) = \left\langle \left(v_X^\dagger - Tv_\Omega^\dagger \delta_{\Phi X} \right) \exp(i\Theta), n(z,t) \right\rangle$ for $X = A$, Ω, T and Φ, and $n_P(z) = \frac{1}{w}\langle R(z,t), n(z,t)\rangle$. The relations between biasing weights and the optimal parameter paths are given by taking the expectation of these equations, which for the parameters A, Ω and T, gives the same relations obtained in Chapter 2, i.e.,

$$\dot{A}_{\text{opt}} = \Delta_A(z) \tag{4.15a}$$

$$\dot{\Omega}_{\text{opt}} = \Delta_\Omega(z) + \Delta_\Omega(z)\Delta_\Phi(z)\frac{T_{\text{opt}}A_{\text{opt}}^2}{1 + \frac{\pi^2}{12} + T_{\text{opt}}^2 A_{\text{opt}}^2} \tag{4.15b}$$

and

$$\dot{T}_{\text{opt}} = \Omega_{\text{opt}} + \Delta_T(z). \tag{4.15c}$$

Of course, as previously noted, the assumption that $\Delta_T(z) = \Delta_\Omega(z) = 0$ yield trivial optimal paths for both the timing and frequency parameters, i.e., $\dot{T}_{\text{opt}}(z) = T_{\text{opt}}(z) = 0$ and $\dot{\Omega}_{\text{opt}}(z) = \Omega_{\text{opt}}(z) = 0$. The equation for $\dot{\Phi}_{\text{opt}}(z)$ is given by

$$\dot{\Phi}_{\text{opt}} = \frac{A_{\text{opt}}^2 - \Omega_{\text{opt}}^2}{2} + 2P_{\text{opt}} + \Delta_\Phi(z), \tag{4.15d}$$

and now includes a term for the optimal path of the radiation power, P_{opt}, which is given by the equation

$$\dot{P}_{\text{opt}} = \sigma_{\text{n}}^2 + \Delta_P(z). \tag{4.15e}$$

Note that unlike the soliton parameters, P_{opt} has a nonzero contribution from the expectation of the noise term driving the radiation, the value of which was given in Equation (4.6). Including these path constraints in the cumulative L^2-norm gives,

$$S_\Phi = \int_0^{z_\text{f}} \frac{\dot{A}_{\text{opt}}^2}{2A_{\text{opt}}} + \left(\dot{\Phi}_{\text{opt}} - \frac{1}{2}A_{\text{opt}}^2 - 2P_{\text{opt}}\right)^2 \frac{18A_{\text{opt}}}{12+\pi^2} + \frac{w\left(\dot{P}_{\text{opt}} - \sigma_\text{n}^2\right)^2}{2P_{\text{opt}}}\, dz \tag{4.16}$$

with initial value constraints

$$A_{\text{opt}}(0) = A_\text{i}, \quad \Phi_{\text{opt}}(0) = \Phi_\text{i}, \quad P_{\text{opt}}(0) = P_\text{i}, \tag{4.17}$$

and final value constraint

$$\Phi_{\text{opt}}(z_\text{f}) = \Phi_\text{f}. \tag{4.18}$$

The values A_i and Φ_i are, respectfully, the amplitude and phase of the launched soliton, the value P_i is the initial power in the radiation, and Φ_f is the targeted final phase value. Note that $A_{\text{opt}}(z_\text{f})$ and $P_{\text{opt}}(z_\text{f})$ are both unconstrained and must be determined through the optimization process. Because of this, the targeted final phase value is best enforced through the use of a Lagrangian multiplier, giving

$$S_\Phi = \int_0^{z_\text{f}} F(A_{\text{opt}}, \dot{A}_{\text{opt}}, \dot{\Phi}_{\text{opt}}, P_{\text{opt}}, \dot{P}_{\text{opt}})\, dz - \lambda\left(\Phi_{\text{opt}}(z_\text{f}) - \Phi_\text{f}\right), \tag{4.19a}$$

with

$$F = \frac{\dot{A}_{\text{opt}}^2}{2A_{\text{opt}}} + \left(\dot{\Phi}_{\text{opt}} - \frac{1}{2}A_{\text{opt}}^2 - 2P_{\text{opt}}\right)^2 \frac{18A_{\text{opt}}}{12+\pi^2} + \frac{w\left(\dot{P}_{\text{opt}} - \sigma_\text{n}^2\right)^2}{2P_{\text{opt}}}, \tag{4.19b}$$

where the initial value constraints can be applied externally.

Taking variations with respect to each parameter gives the following BVP:

$$\left(\dot{\Phi}_{\text{opt}} - \frac{1}{2}A_{\text{opt}}^2 - 2P_{\text{opt}}\right) = \lambda\frac{12+\pi^2}{36A_{\text{opt}}}, \quad \Phi_{\text{opt}}(0) = \Phi_\text{i}, \quad \Phi_{\text{opt}}(z_\text{f}) = \Phi_\text{f} \tag{4.20a}$$

$$\ddot{A}_{\text{opt}} = \frac{\dot{A}_{\text{opt}}^2}{2A_{\text{opt}}} + \lambda^2\frac{12+\pi^2}{72A_{\text{opt}}} - \lambda A_{\text{opt}}^2, \quad A_{\text{opt}}(0) = A_\text{i}, \quad \dot{A}_{\text{opt}}(z_\text{f}) = 0 \tag{4.20b}$$

$$\frac{d}{dz}\left(\frac{\dot{P}_{\text{opt}}-\sigma_{\text{n}}^2}{P_{\text{opt}}}\right)+\frac{1}{2}\left(\frac{\dot{P}_{\text{opt}}-\sigma_{\text{n}}^2}{P_{\text{opt}}}\right)^2=-\frac{2\lambda}{w},\quad P_{\text{opt}}(0)=P_{\text{i}},\quad \dot{P}_{\text{opt}}(z_{\text{f}})=\sigma_{\text{n}}^2 \tag{4.20c}$$

The first two equations are similar to those derived in Section 2.3.1, the only difference being that the optimal phase path now depends on the optimal path of the radiation power, which satisfies an additional second order equation, bring the total number of degrees of freedom to five. This severly complicates any attempt at finding a numerical solution via a shooting method approach, as was done in Section 2.3.1. Fortunately, the equation for the optimal path of the radiation power, i.e., Equation (4.20c), can be solved exactly, as detailed in the next section, which reduces the dimensionality of the system back to three. Thus, finding the biasing distribution with radiation has the same computational cost as the original biasing problems of Chapter 2.

4.2.3 Solution for Optimal Path of Radiation Power Parameter

First, note that $\lambda = 0$ corresponds to the case of no biasing, since in this case, Equation (4.20b) is solved by the constant solution $A_{\text{opt}}(z) = A_{\text{i}}$, Equation (4.20c) is solved by $P_{\text{opt}}(z) = P_{\text{i}}+\sigma_{\text{n}}^2 z$, which only includes the linear evolution from the non-zero mean noise contribution, and Equation (4.20a) can be integrated to give $\Phi_{\text{opt}}(z) = \Phi_{\text{i}} + \frac{A_{\text{i}}^2}{2}z + \sigma_{\text{n}}^2 z^2 + 2P_{\text{i}}z$. Of course, for this to be an actual solution, the final value of the phase must match the targeted value, i.e., $\Phi_{\text{opt}}(z_{\text{f}}) = \Phi_{\text{i}}+\frac{A_{\text{i}}^2}{2}z_{\text{f}}+\sigma_{\text{n}}^2 z_{\text{f}}^2+2P_{\text{i}}z_{\text{f}} = \Phi_{\text{f}}$. Assuming this is not the case, Equation (4.20c) must be solved for $\lambda > 0$ and $\lambda < 0$.

Assuming that $\lambda > 0$, the substitutions

$$z=\sqrt{\frac{w}{\lambda}}x \quad \text{and} \quad \frac{\dot{P}_{\text{opt}}-\sigma_{\text{n}}^2}{P_{\text{opt}}}=-2\sqrt{\frac{\lambda}{w}}\hat{P} \tag{4.21}$$

transform equation (4.20c) to the form

$$\frac{d\hat{P}}{dx}=1+\hat{P}^2, \tag{4.22}$$

which is identified as a trigonometric identity for $\hat{P}(x) = \tan(x)$. Transforming back to the original variables gives,

$$\frac{\dot{P}_{\text{opt}} - \sigma_{\text{n}}^2}{P_{\text{opt}}} = -2\sqrt{\frac{\lambda}{w}} \tan\left(\sqrt{\frac{\lambda}{w}}(z - z_{\text{f}})\right), \tag{4.23}$$

where the constant from integrating (4.22) was used to satisfy the end point constraint $\dot{P}_{\text{opt}}(z_{\text{f}}) = \sigma_{\text{n}}^2$, leaving an initial value problem of the form

$$\dot{P}_{\text{opt}} + 2\sqrt{\frac{\lambda}{w}} \tan\left(\sqrt{\frac{\lambda}{w}}(z - z_{\text{f}})\right) P_{\text{opt}} = \sigma_{\text{n}}^2, \quad P_{\text{opt}}(0) = P_{\text{i}}. \tag{4.24}$$

This equation is easily integrated, giving a solution of

$$\begin{aligned} P_{\text{opt}}(z) = {} & P_{\text{i}} \left[\frac{\cos\left(\sqrt{\frac{\lambda}{w}}(z - z_{\text{f}})\right)}{\cos\left(\sqrt{\frac{\lambda}{w}} z_{\text{f}}\right)}\right]^2 \\ & + \sigma_{\text{n}}^2 \sqrt{\frac{w}{\lambda}} \left(\cos\left(\sqrt{\frac{\lambda}{w}}(z - z_{\text{f}})\right)\right)^2 \left[\tan\left(\sqrt{\frac{\lambda}{w}}(z - z_{\text{f}})\right) + \tan\left(\sqrt{\frac{\lambda}{w}} z_{\text{f}}\right)\right]. \end{aligned} \tag{4.25}$$

Finally, the weighting function, i.e., $\Delta_P(z) = \dot{P}_{\text{opt}}(z) - \sigma_{\text{n}}^2$, is given by

$$\begin{aligned} \Delta_P(z) = {} & -\cos\left(\sqrt{\frac{\lambda}{w}}(z - z_{\text{f}})\right) \sin\left(\sqrt{\frac{\lambda}{w}}(z - z_{\text{f}})\right) \times \\ & \left[P_{\text{i}} \sqrt{\frac{\lambda}{w}} \frac{2}{\left(\cos\left(\sqrt{\frac{\lambda}{w}} z_{\text{f}}\right)\right)^2} + 2\sigma_{\text{n}}^2 \left[\tan\left(\sqrt{\frac{\lambda}{w}}(z - z_{\text{f}})\right) + \tan\left(\sqrt{\frac{\lambda}{w}} z_{\text{f}}\right)\right]\right]. \end{aligned} \tag{4.26}$$

The case of $\lambda < 0$ similar to the $\lambda > 0$ case above, where the initial transformations

$$z = \sqrt{\frac{w}{\lambda}} x \quad \text{and} \quad \frac{\dot{P}_{\text{opt}} - \sigma_{\text{n}}^2}{P_{\text{opt}}} = 2\sqrt{\frac{\lambda}{w}} \hat{P} \tag{4.27}$$

inserted into equation (4.20c) yields

$$\frac{d\hat{P}}{dx} = 1 - \hat{P}^2, \tag{4.28}$$

which is solved by $\hat{P}(x) = \tanh(x)$, giving a similar initial value problem of

$$\dot{P}_{\text{opt}} - 2\sqrt{\frac{\lambda}{w}}\tanh\left(\sqrt{\frac{\lambda}{w}}(z - z_{\text{f}})\right) P_{\text{opt}} = \sigma_{\text{n}}^2, \quad P_{\text{opt}}(0) = P_{\text{i}}. \tag{4.29}$$

This equation has the solution

$$\begin{aligned} P_{\text{opt}}(z) = P_{\text{i}} \left[\frac{\cosh\left(\sqrt{\frac{\lambda}{w}}(z - z_{\text{f}})\right)}{\cosh\left(\sqrt{\frac{\lambda}{w}}z_{\text{f}}\right)}\right]^2 + \sigma_{\text{n}}^2\sqrt{\frac{w}{\lambda}}\left(\cosh\left(\sqrt{\frac{\lambda}{w}}(z - z_{\text{f}})\right)\right)^2 \times \\ \left[\tanh\left(\sqrt{\frac{\lambda}{w}}(z - z_{\text{f}})\right) + \tanh\left(\sqrt{\frac{\lambda}{w}}z_{\text{f}}\right)\right], \end{aligned} \tag{4.30}$$

which implies a biasing function of

$$\begin{aligned} \Delta_P(z) = \cosh\left(\sqrt{\frac{\lambda}{w}}(z - z_{\text{f}})\right) \sinh\left(\sqrt{\frac{\lambda}{w}}(z - z_{\text{f}})\right) \\ \left[P_{\text{i}}\sqrt{\frac{\lambda}{w}}\frac{2}{\left(\cosh\left(\sqrt{\frac{\lambda}{w}}z_{\text{f}}\right)\right)^2} + 2\sigma_{\text{n}}^2\left[\tanh\left(\sqrt{\frac{\lambda}{w}}(z - z_{\text{f}})\right) + \tanh\left(\sqrt{\frac{\lambda}{w}}z_{\text{f}}\right)\right]\right]. \end{aligned} \tag{4.31}$$

4.2.4 The Optimal Biasing Solution

Since the solution for the optimal path of radiation power is known exactly, the remaining two equations can be solved by the same shooting method that was implemented in Section 2.3.1. Following this, equation (4.20a) is integrated to given λ in terms of integrals of A_{opt} and P_{opt} and the boundary conditions for the phase, i.e.,

$$\lambda(A_{\text{opt}}, z) = \frac{(\Phi_{\text{f}} - \Phi_{\text{i}}) - \frac{1}{2}\int_0^{z_{\text{f}}} A_{\text{opt}}^2\, dz - 2\int_0^{z_{\text{f}}} P_{\text{opt}}\, dz}{\frac{12+\pi^2}{36}\int_0^{z_{\text{f}}} \frac{1}{A_{\text{opt}}}\, dz}. \tag{4.32}$$

Putting this back into equation (4.20b) gives the integro-differential equation

$$\ddot{A}_{\text{opt}} = \frac{\dot{A}_{\text{opt}}^2}{2A_{\text{opt}}} + (\lambda(A_{\text{opt}}, z))^2\frac{12 + \pi^2}{72A_{\text{opt}}} - \lambda(A_{\text{opt}}, z)A_{\text{opt}}^2, \tag{4.33a}$$

with boundary conditions

$$A_{\text{opt}}(0) = A_{\text{i}} \quad \text{and} \quad \dot{A}_{\text{opt}}(z_{\text{f}}) = 0. \tag{4.33b}$$

Figures 4.1 (128 modes), 4.2 (256 modes), 4.3 (512 modes) and 4.4 (1024 modes) plot the optimal paths for the amplitude, phase and radiation power parameters for various final targeted phase values. These plots show that, as expected, the amount of biasing the radiation requires increases with the number of modes included in the simulations. At the same time, the biasing of the amplitude and phase parameters are decrease as the number of radiation modes are increased, indicating that the radiation can make significant contributions to the evolution of the phase parameter.

4.3 Results and Discussion

This section shows the results of using the improved low-dimensional reduction given in (4.1) to guide the application of the ISMC method. Figures 4.5, 4.6, 4.7 and 4.8 display the modified PDFs resulting from the use of the biasing functions displayed in Figures 4.1-4.4. As expected, these plots show that including radiation in the biasing scheme results in as lower COV for the same number MC runs. This improvement becomes more pronounced as the number of radiation modes increase. In particular, Figure 4.8 shows an approximate 50% improvement in the coefficient of variation for around the portion of the PDF corresponding to the mean phase value. In addition, this improvement extends far down into the left tails of the distribution, indicating that radiation plays an important role in the production of large phase deviations. Unfortunately, increasing the radiation to drive the phase to higher values, also makes it much more difficult to accurately resolve the parameters of the soliton at those amplifiers toward the end of the transmission line. This is particularly evident in the right half of the COV curve in Figure 4.8, which represents both a large number of radiation modes in the simulations and high final phase values.

4.4 Summary

This chapter used the improved low-dimensional reduction, derived in Chapter 3 and given in Equations (4.1), to derived new biasing paths for the parameters representing the phase, amplitude *and* power in the radiation. Using these biasing paths, an improved ISMC method is constructed, which unlike the original implementation, also includes biasing in the radiation to drive the phase parameter. With the successful implementation of the improved ISMC method above, the next chapter extends this approach to more realistic evolution equations.

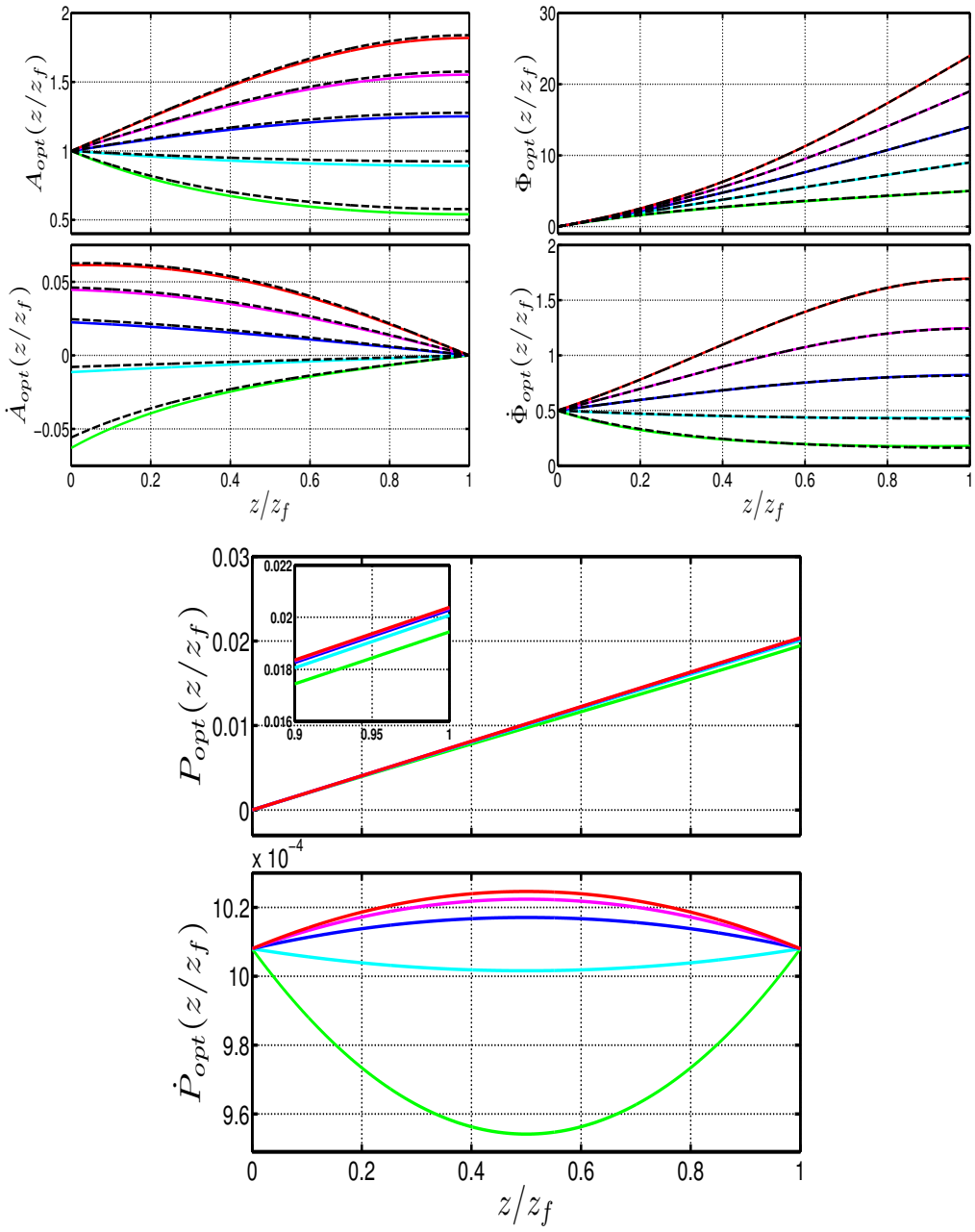


Figure 4.1 Top: The biasing paths (Top) and derivatives (Bottom) for the amplitude (Left) and phase (Right) parameters using **128** radiation modes. In all plots the green, cyan, blue, magenta and red curves represent final phase targets of 5.0, 9.0, 14.0, 19.0 and 24.0, respectively, which cover the entire range of values used in the implementation of the ISMC method. In addition, the black dashed curves are the corresponding biasing paths of Chapter 2, which do not include radiation. Bottom: The biasing paths (Top) and derivatives (Bottom) for the radiation power parameter using the same final phase targets and color scheme as the top four figures.

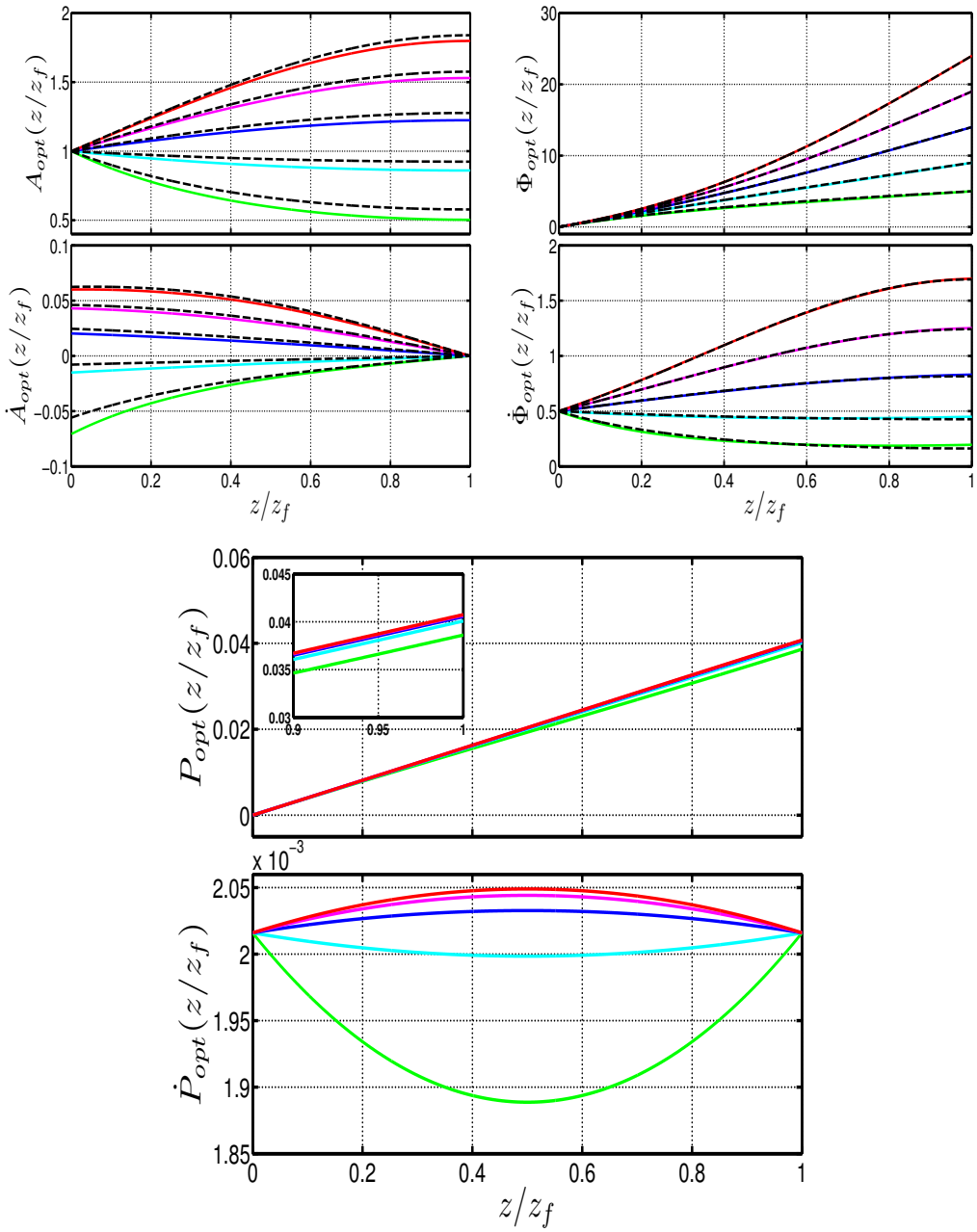


Figure 4.2 Top: The biasing paths (Top) and derivatives (Bottom) for the amplitude (Left) and phase (Right) parameters using **256** radiation modes. In all plots the green, cyan, blue, magenta and red curves represent final phase targets of 5.0, 9.0, 14.0, 19.0 and 24.0, respectively, which cover the entire range of values used in the implementation of the ISMC method. In addition, the black dashed curves are the corresponding biasing paths of Chapter 2, which do not include radiation. Bottom: The biasing paths (Top) and derivatives (Bottom) for the radiation power parameter using the same final phase targets and color scheme as the top four figures.

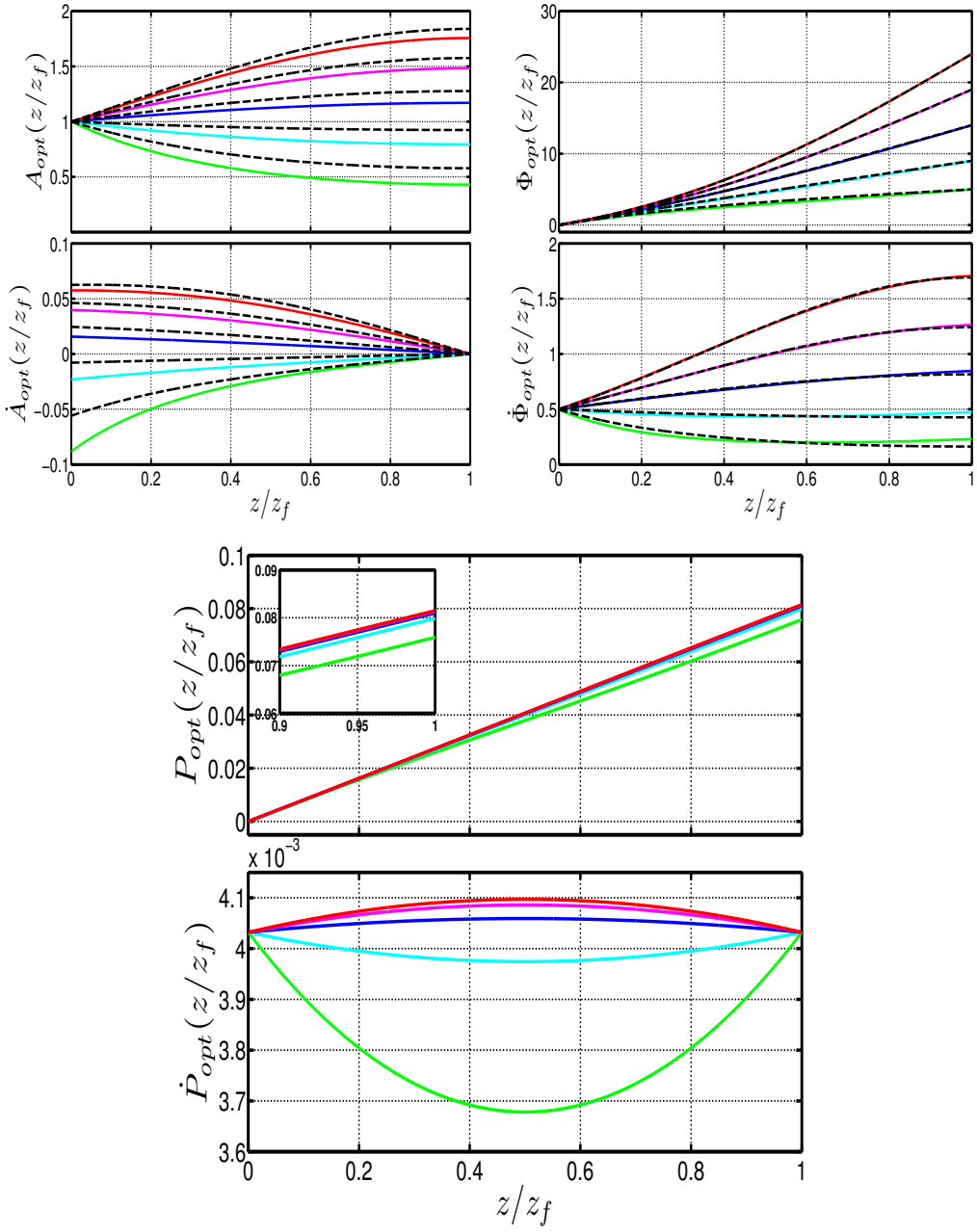


Figure 4.3 Top: The biasing paths (Top) and derivatives (Bottom) for the amplitude (Left) and phase (Right) parameters using **512** radiation modes. In all plots the green, cyan, blue, magenta and red curves represent final phase targets of 5.0, 9.0, 14.0, 19.0 and 24.0, respectively, which cover the entire range of values used in the implementation of the ISMC method. In addition, the black dashed curves are the corresponding biasing paths of Chapter 2, which do not include radiation. Bottom: The biasing paths (Top) and derivatives (Bottom) for the radiation power parameter using the same final phase targets and color scheme as the top four figures.

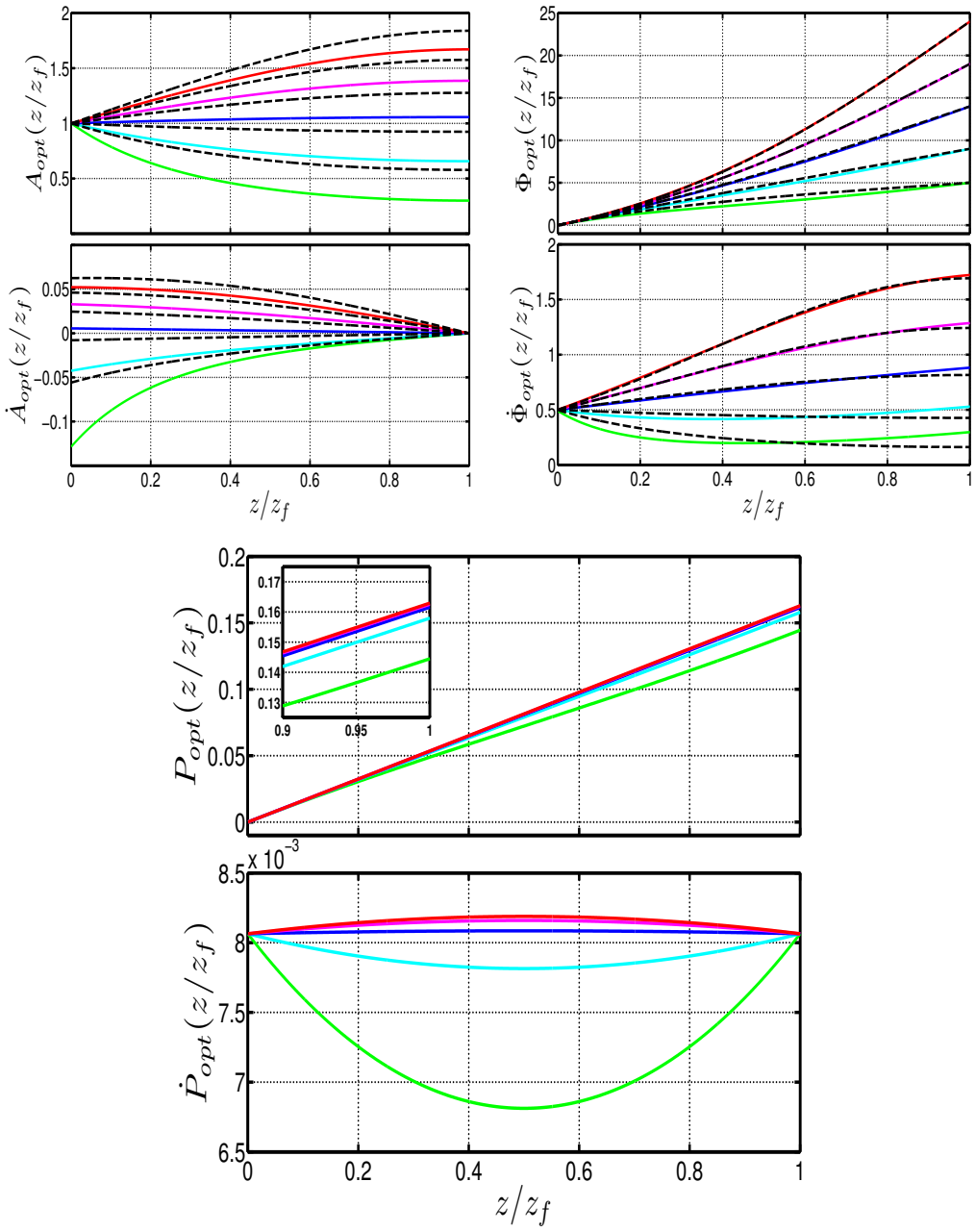


Figure 4.4 Top: The biasing paths (Top) and derivatives (Bottom) for the amplitude (Left) and phase (Right) parameters using **1024** radiation modes. In all plots the green, cyan, blue, magenta and red curves represent final phase targets of 5.0, 9.0, 14.0, 19.0 and 24.0, respectively, which cover the entire range of values used in the implementation of the ISMC method. In addition, the black dashed curves are the corresponding biasing paths of Chapter 2, which do not include radiation. Bottom: The biasing paths (Top) and derivatives (Bottom) for the radiation power parameter using the same final phase targets and color scheme as the top four figures.

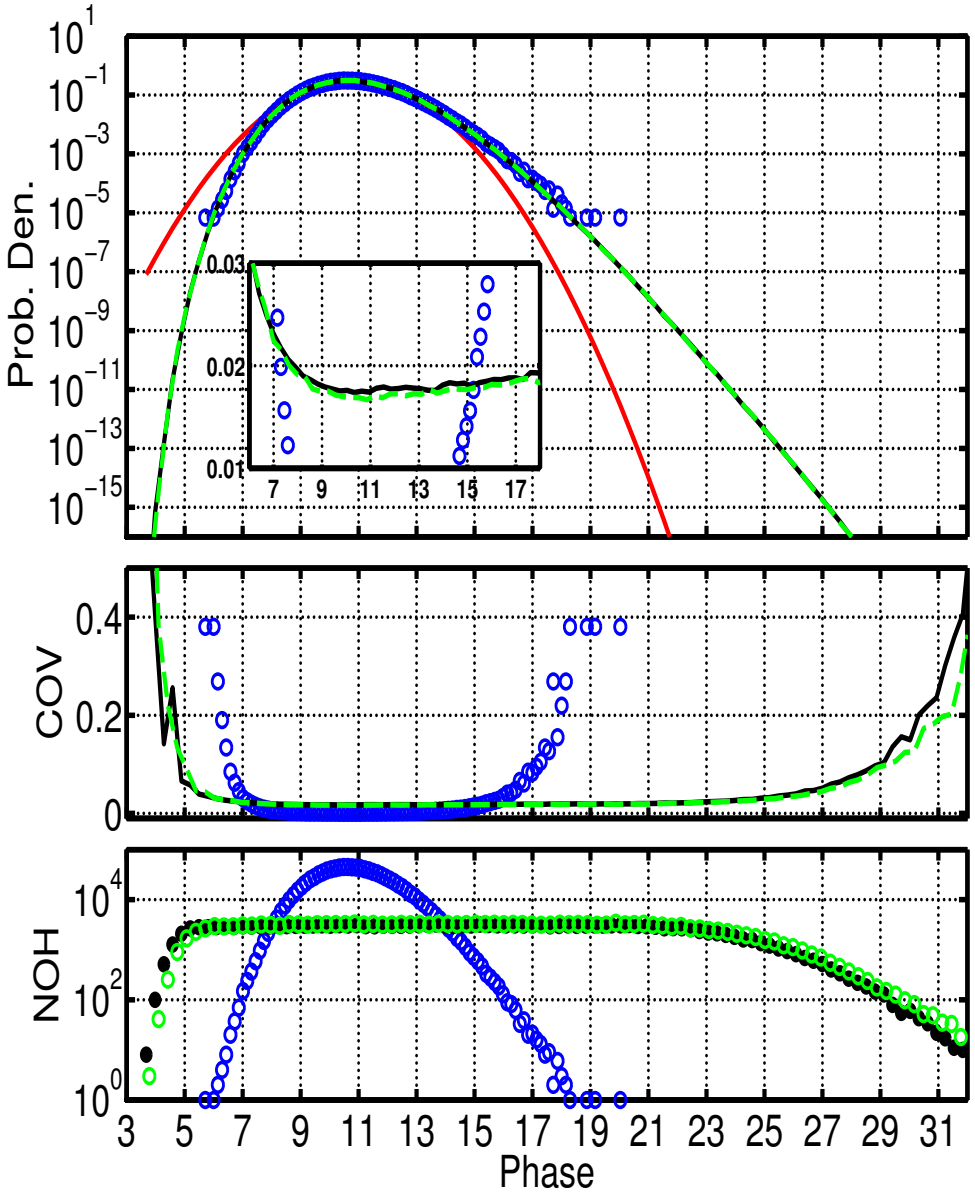


Figure 4.5 Top: The black curve gives the PDF of the phase parameter using 2×10^5 ISMC simulations of the stochastically forced NLSE in (3.11) with **128** simulation modes. This is compared to the green curve representing the results of 2×10^5 ISMC runs using **128** modes, but guided by the SODEs of SPT. For comparison, the blue markers represent the results of 1×10^6 MC runs of the SODEs given in Equation (3.28), while the red curve is the analytical solution of Equation (3.28) linearized around the initial conditions $A = 1$ and $\Omega = T = \Phi = 0$. All plots are on a Log scale of base 10. Middle: Each colored curve is the represents the COV for the corresponding PDF in the top figure. Bottom: Each sequence of colored markers correspond to the NOH each bin received under the MC simulations that produced to the PDF plots in the top figure.

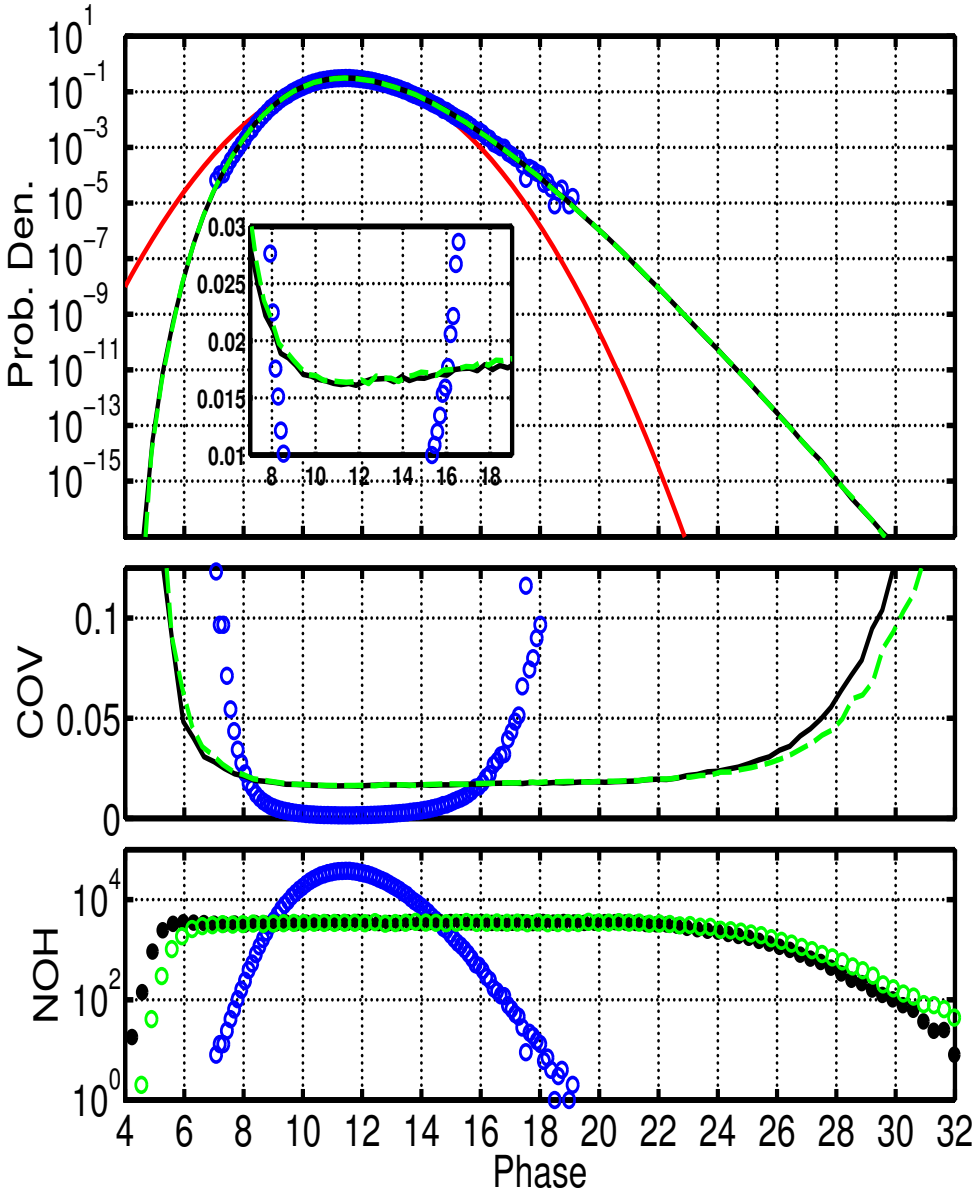


Figure 4.6 Top: The black curve gives the PDF of the phase parameter using 2×10^5 ISMC simulations of the stochastically forced NLSE in (3.11) with **256** simulation modes. This is compared to the green curve representing the results of 2×10^5 ISMC runs using **256** modes, but guided by the SODEs of SPT. For comparison, the blue markers represent the results of 1×10^6 MC runs of the SODEs given in Equation (3.28), while the red curve is the analytical solution of Equation (3.28) linearized around the initial conditions $A = 1$ and $\Omega = T = \Phi = 0$. All plots are on a Log scale of base 10. Middle: Each colored curve is the represents the COV for the corresponding PDF in the top figure. Bottom: Each sequence of colored markers correspond to the NOH each bin received under the MC simulations that produced to the PDF plots in the top figure.

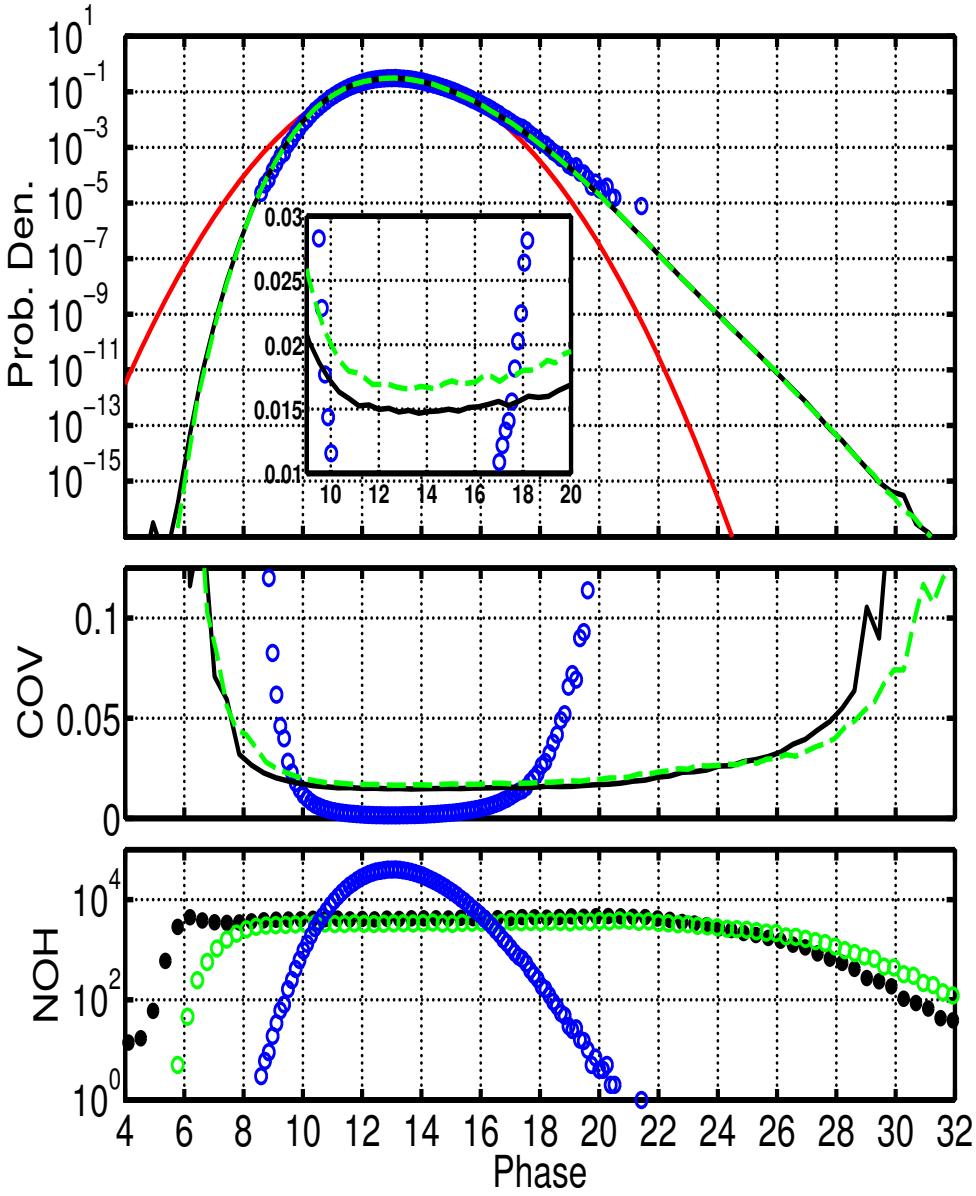


Figure 4.7 Top: The black curve gives the PDF of the phase parameter using 2×10^5 ISMC simulations of the stochastically forced NLSE in (3.11) with **512** simulation modes. This is compared to the green curve representing the results of 2×10^5 ISMC runs using **512** modes, but guided by the SODEs of SPT. For comparison, the blue markers represent the results of 1×10^6 MC runs of the SODEs given in Equation (3.28), while the red curve is the analytical solution of Equation (3.28) linearized around the initial conditions $A = 1$ and $\Omega = T = \Phi = 0$. All plots are on a Log scale of base 10. Middle: Each colored curve is the represents the COV for the corresponding PDF in the top figure. Bottom: Each sequence of colored markers correspond to the NOH each bin received under the MC simulations that produced to the PDF plots in the top figure.

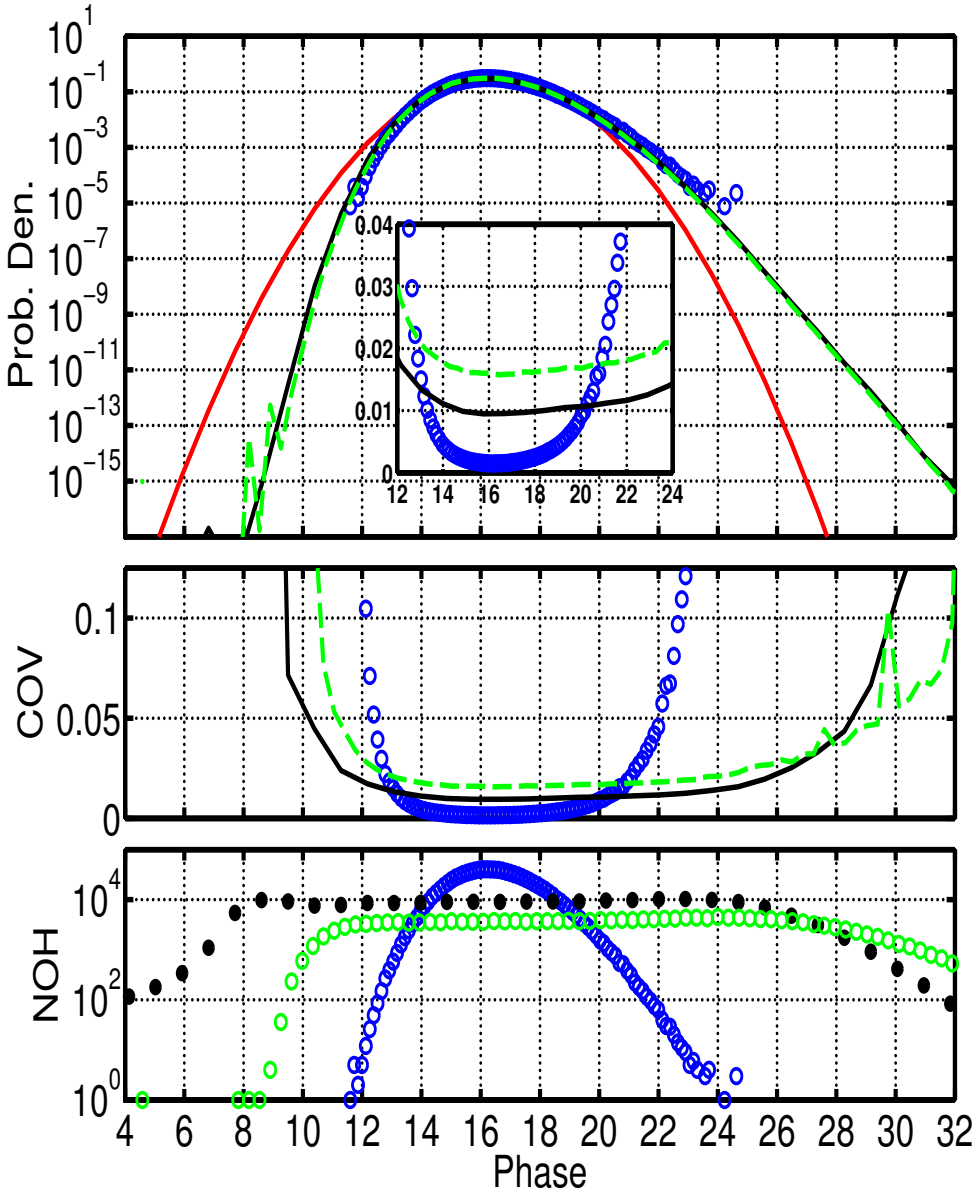


Figure 4.8 Top: The black curve gives the PDF of the phase parameter using 2×10^5 ISMC simulations of the stochastically forced NLSE in (3.11) with **1024** simulation modes. This is compared to the green curve representing the results of 2×10^5 ISMC runs using **1024** modes, but guided by the SODEs of SPT. For comparison, the blue markers represent the results of 1×10^6 MC runs of the SODEs given in Equation (3.28), while the red curve is the analytical solution of Equation (3.28) linearized around the initial conditions $A = 1$ and $\Omega = T = \Phi = 0$. All plots are on a Log scale of base 10. Middle: Each colored curve is the represents the COV for the corresponding PDF in the top figure. Bottom: Each sequence of colored markers correspond to the NOH each bin received under the MC simulations that produced to the PDF plots in the top figure.

CHAPTER 5

VARIATIONAL APPROXIMATION FOR THE NLSE WITH DISPERSION MANAGEMENT AND DMNLSE

Prior to 1995, models for electromagnetic signal propagation through optical fiber were largely based on the constant-coefficient NLSE, with the primary impacts of ASE noise understood to be Gordon-Haus timing jitter [39] and Gordon-Mollenauer phase jitter [40]. At the same time, however, the drive for higher throughput was leading researchers to couple multiple frequency (wavelength) channels into the same fiber for co-propagation using a technique known as wavelength-division multiplexing (WDM). The presence of multiple wavelengths in the same fiber leads to crosstalk through the fiber's nonlinearity, however, and this effect grows more severe as the number of wavelengths increases. Around this time, optical engineers developed a technique that is now known as *dispersion management* (DM), whereby sections of fiber with oppositely signed dispersion constants are concatenated in a periodic array [68] This was originally done to mitigate the timing and phase jitter in soliton propagation, however, it was quickly realized that DM also reduced crosstalk between the multiple frequencies that are present in a fiber using WDM. The essential idea behind DM is to have a high local value of dispersion in order to cause the pulses to broaden and reform periodically (i.e., to "breathe"), thus minimizing the conversion of frequency fluctuations to timing jitter through the Gordon-Haus effect, and a low path-averaged value of dispersion to balance the nonlinearity and form stable periodic solutions. While it is not altogether surprising that solitons of the constant-dispersion NLSE persist as periodic solitary waves in the perturbative limit of small-amplitude DM, it is rather more surprising that exact periodically breathing solutions of the NLSE+DM exist far away from the perturbative limit of the NLSE. These periodic

pulses are commonly referred to as *dispersion managed solitons*, even though they are *not* solutions of an integrable system.

In this chapter, the NLSE is extended to the case of varying dispersion and nonlinear coefficients to construct a model for DM soliton propagation in DM fiber, i.e., NLSE+DM. Although this equation is close to the form of the NLSE, the DM soliton solutions of NLSE+DM are very different from the soliton solutions of NLSE, due to the complex breathing dynamics resulting from the periodic variations in the dispersion coefficient. Fortunately, the periodicity that is characteristic of the NLSE+DM allows for a natural simplification through averaging, where the propagation a DM soliton is averaged over one dispersion map period, effectively separating the fast breathing from the slower evolution of the pulse "core". This core solution is shown to satisfy an averaged equation referred to as the DMNLSE [50], which no longer has periodically varying coefficients. However, the price paid for this simplification is the introduction of a non-local nonlinearity that takes the form of a convolution.

This chapter also presents a low-dimensional reduction for the pulse dynamics in a stochastically forced versions of DMNLSE. Like the NLSE, the core DM soliton solutions of the DMNLSE exhibit a radiation induce phase drift when noise is included in the propagation model. In addition, an extension of SPT to the DMNLSE for the derivation of a low-dimensional reduction for the DM soliton core dynamics fails to captures these effects in the same way it fail to capture radiation effects in the NLSE [51]. In preparation of constructing an ISMC method that accounts for this phase drift, the low-dimensional reduction constructed here uses the same variational approach presented in Chapter 4 for the NLSE, enabling the inclusion of radiation effects through the choice of ansatz for the evolution of the radiation. Unlike the NLSE, however, the localized pulse solutions to the DMNLSE can only be found

numerically, which introduces an additional source of error in the low-dimensional approximation.

5.1 The NLSE with Dispersion Management: NLSE+DM

Under dispersion management (DM), the transmission line consists of short segments of optical fiber characterized by large individual dispersion coefficients that alternate in sign, forming what is referred to as a dispersion map. For several reasons, the dispersion maps implemented in physical systems are neither exactly periodic nor symmetrically centered; in a typical scenario most of the dispersion period consists of single-mode fiber (SMF) with fixed anomalous (positive) dispersion, with a relatively short segment of dispersion-compensating fiber (DCF) with normal (negative) dispersion placed just prior to the amplifier. Nevertheless, it is often mathematically equivalent to treat an idealized map with equal amounts of fiber having dispersion coefficients that are nearly equal in absolute value, forming a periodic dispersion map that is symmetric around the midpoint between amplifiers and possessing a period congruent to the amplification cycle [69]. Figures A.2 and A.3 in Appendix A provides a comparison between realistic and idealized dispersion maps, along with the dynamics of a typical DM pulse. The dispersion coefficients in the dispersion map are chosen such that the accumulated dispersion over each map period nearly cancels, resulting in an average dispersion value that is much smaller than the dispersion coefficients of both fibers. In addition, the two fibers typically differ in the coefficients associated with nonlinearity and loss as well, however, these *do not* alternate in sign and thus, average out to a value very close to the local values of the individual fibers.

With the details discussed above, the evolution equation for optical pulses in DM fiber (denoted NLSE+DM) can be derived from a simple modification to the NLSE, where the constant coefficients are replaced with periodic functions that depend on the spacial coordinate z. Under the assumption of a piecewise constant dispersion

map, referred to as a two step map, this equation takes the dimensionless form

$$i\frac{\partial U}{\partial z} + ic_{\rm lg}\left(1 - \sum_{m=1}^{N_{\rm a}} \delta\left(\frac{z}{z_{\rm a}} - m\right)\right) U + \frac{1}{2}c_{\rm dm}(z)\frac{\partial^2 U}{\partial t^2} + c_{\rm nl}|U|^2 U = i\sum_{m=1}^{N_{\rm a}} \sigma n_m(t)\,\delta(z - mz_{\rm a})\,, \tag{5.1a}$$

with noise statistics that are equivalent to the constant coefficient case, i.e.,

$$\mathbb{E}[n_j(t)] = 0 \quad \text{and} \quad \mathbb{E}[n_j(t_1)\bar{n}_k(t_2)] = \delta_{jk}\,\delta(t_1 - t_2)\,, \tag{5.1b}$$

with σ^2 representing the dimensionless noise strength. The dimensionless dispersion map coefficient is given by

$$c_{\rm dm}(z) = \begin{cases} c_{\rm dm,1}, & 0 \le z < \frac{z_{\rm a}}{4} \\ -c_{\rm dm,2}, & \frac{z_{\rm a}}{4} \le z < \frac{3z_{\rm a}}{4} \\ c_{\rm dm,1}, & \frac{3z_{\rm a}}{4} \le z < z_{\rm a}, \end{cases} \tag{5.1c}$$

with numerical subscripts denoting the two types of fibers that comprise the dispersion map. The choice of dimensional scales that lead to this equation are given in Appendix A, and are chosen such that (i) the dimensionless dispersion map period appears as a small parameter, i.e., $z_{\rm a} \ll 1$, (ii) the dispersion and loss/gain coefficients are large, i.e., $c_{{\rm dm},j}, c_{\rm lg} \sim O(1/z_{\rm a})$ for $j = 1, 2$, and (iii) the nonlinear coefficients satisfy $c_{\rm nl} \sim O(1)$. Note that because the dispersion coefficient now depends on the evolution variable z, the NLSE+DM is does not possess the same invariances as the NLSE from which it was derived. In particular, it no longer has the Galilean invariance that produces the free frequency parameter seen in NLSE solitons, nor the amplitude invariance which involves rescaling both the time and space variables.

5.2 The DMNLSE

The small dimensionless period of the dispersion map, combined with the large changes in the local dimensionless dispersion coefficient, creates a natural setting for the application of asymptotic averaging. This procedure factors the varying coefficient NLSE+DM into an $O(1)$ linear equation capturing the "fast" breathing dynamics of the DM soliton, and a constant coefficient nonlinear integro-differential equation at $O(z_{\rm p})$ (DMNLSE) that governs the evolution of the DM soliton "core". One major benefit of using this approach is that many of the invariances of NLSE that where lost in the transition to NLSE+DM, are found to exist in the DMNLSE, which allows one to formulate a perturbation theory analogous to SPT for the constant coefficient NLSE [70, 51]. Finally, it should be noted that the DMNLSE notation is associated with the work of Ablowitz and Biondini [50] who reformulated the original derivation by Gabitov and Turitsyn [71]. Because of this, the DMNLSE is also referred to as the Gabitov-Turitsyn equation (GTE).

5.2.1 Averaging NLSE+DM: DMNLSE

Since the averaging procedure is fully detailed in Appendix A, only a brief outline the derivation is presented here. By introducing the variable $x = z/z_{\rm a}$ and rewriting the coefficient functions to explicitly reflect their order in $z_{\rm a}$, the NLSE+DM equation in (5.1a) can be written as

$$\begin{aligned} iz_{\rm a}\frac{\partial U}{\partial z} + i\frac{\partial U}{\partial x} + i\tilde{c}_{\rm lg}\left(1 - \sum_{m=1}^{N_{\rm a}} \delta(x-m)\right)U + z_{\rm a}\frac{1}{2}\frac{\partial^2 U}{\partial t^2} \\ + d(x)\frac{\partial^2 U}{\partial t^2} + z_{\rm a}c_{\rm nl}|U|^2U = iz_{\rm a}\sum_{m=1}^{N_{\rm a}} n_m(t)\,\delta(z - mz_{\rm a})\,, \end{aligned} \tag{5.2}$$

where $\tilde{c}_{\rm lg}$ and $d(x)$ are $O(1)$ quantities that are related to the original coefficients by

$$c_{\rm lg} = \frac{1}{z_{\rm a}}\tilde{c}_{\rm lg} \quad \text{and} \quad c_{\rm dm}(z/z_{\rm a}) = 1 + 2\frac{1}{z_{\rm a}}d(x). \tag{5.3}$$

Note that since the dispersion map function alternates in sign, it has an $O(1)$ mean term that is separated from its $O(1/z_{\rm a})$ variations according to

$$\begin{aligned} c_{\rm dm}(z) &= \frac{1}{z_{\rm a}}\int_0^{z_{\rm a}} c_{\rm dm}(z)\,dz + \frac{1}{z_{\rm a}}\left(c_{\rm dm}(z) - \int_0^{z_{\rm a}} c_{\rm dm}(z)\,dz\right) \\ &= 1 + \frac{2}{z_{\rm a}}d(x), \end{aligned} \tag{5.4}$$

where the non-dimensionalization given in Appendix A was chosen such that

$$\frac{1}{z_{\rm a}}\int_0^{z_{\rm a}} c_{\rm dm}(z)\,dz = 1. \tag{5.5}$$

By assuming a series solution of the form

$$U(z,x,t) = u_0(z,x,t) + z_{\rm a}u_1(z,x,t) + z_{\rm a}^2u_2(z,x,t) + \ldots, \tag{5.6}$$

the leading order solution of equation (5.2) is found to be

$$\hat{u}_0(z,x,t) = \hat{u}(z,\omega)\exp(-A(x))\,\exp\!\left(-i\omega^2D(x)\right), \tag{5.7}$$

where

$$A(x) = \tilde{c}_{\rm lg}\int_0^x \left(1 - \sum_{m=1}^{N_{\rm a}}\,\delta(y-m)\right)\,dy, \qquad D(x) = \int_0^x d(y)\,dy, \tag{5.8}$$

and the $\wedge$ superscript denotes the Fourier transform (FT) defined as

$$\begin{aligned} \mathcal{F}[f(t)] = \hat{f}(\omega) &= \int f(t)\exp(i\omega t)\;dt, \\ \mathcal{F}^{-1}\Big[\hat{f}(\omega)\Big] = f(t) &= \frac{1}{2\pi}\int \hat{f}(\omega)\exp(-i\omega t)\;d\omega. \end{aligned} \tag{5.9}$$

By applying Fredholm theory to the $O(z_{\rm a})$ equation resulting from the series expansion, the "core" pulse solution $\hat{u}(z,\omega)$ is seen to satisfy the FT version of the DMNLSE, which takes the form

$$\begin{aligned} i\frac{\partial\hat{u}}{\partial z} - \omega^2\frac{1}{2}\hat{u} + \int\int \hat{u}(z,\omega_1+\omega)\hat{u}(z,\omega_2+\omega)\bar{\hat{u}}(z,\omega_2+\omega_1+\omega)\times \\ \hat{K}(\omega_1,\omega_2;s_{\rm map})\,d\omega_1\,d\omega_2 = i\sum_{m=1}^{N_{\rm a}}\hat{n}_m(\omega)\,\delta(z-mz_{\rm a}), \end{aligned} \tag{5.10a}$$

where

$$\hat{K}(\omega_1, \omega_2; s_{\text{map}}) = \frac{1}{(2\pi)^2} \operatorname{sinc}\left(\omega_1 \omega_2 \frac{s_{\text{map}}}{2}\right). \tag{5.10b}$$

Note that this equation is parameterized by what is referred to as the map strength, denoted s_{map}, which is defined as the L^1-norm of the variations in the non-dimensional dispersion map, i.e.,

$$s_{\text{map}} = \int_0^1 |d(x)| \, dx. \tag{5.11}$$

By taking the inverse FT, the DMNLSE in the time domain is given as

$$\begin{aligned} i\frac{\partial u}{\partial z} + \frac{1}{2}\frac{\partial^2 u}{\partial t^2} + \int\int u(z, t_2 + t) u(z, t_1 + t) \bar{u}(z, t_1 + t_2 + t) \times \\ K(t_1, t_2; s_{\text{map}}) \, dt_1 dt_2 = i \sum_{m=1}^{N_{\text{a}}} n_m(t) \, \delta(z - m z_{\text{a}}), \end{aligned} \tag{5.12a}$$

where

$$K(t_1, t_2; s_{\text{map}}) = \int\int \exp(-i\omega_1 t_1) \exp(-i\omega_2 t_2) \, \hat{K}(\omega_1, \omega_2; s_{\text{map}}) \, d\omega_1 \, d\omega_2. \tag{5.12b}$$

5.2.2 Invariances of the DMNLSE

As previously stated, the averaging procedure which transforms the NLSE+DM to the DMNLSE also restores many of the invariances (more precisely the conserved quantities and rate equations that result from Noether's theorem [72, 73, 74, 75]) that are commonly associated with NLSE. In particular, the DMNLSE is phase, time and Galilean invariant, which implies the introduction of free parameters Φ, T and Ω, that appear in pulse solutions of the DMNLSE in a manner similar to the soliton solutions of the NLSE. Note that the DMNLSE is also invariant under translations in z; however, like the NLSE, this invariance can be shown to be equivalent to a combination of the above three for the pulse solutions discussed here.

One more invariance exists in the DMNLSE, that leads the inclusion of an amplitude scaling parameter A. As in the NLSE, this invariance is associated with conservation of energy (or photon number) given by

$$\frac{d}{dz}\int |u(z,t)|^2\,dt = 0. \tag{5.13}$$

In the case of the NLSE, this leads to the inverse relationship between the amplitude and width parameters in the soliton solution of Equation (2.4). However, since pulse solutions of the DMNLSE have an additional dependence on the map strength parameter, the relationship here is not as simple. When the map strength parameter of a DM soliton is varied, the functional form of the pulse solution changes, which subsequently changes the initial amplitude and width of the pulse. Although this variation in the pulse shape does not affect the other conserved quantities, it does affect the pulse energy. Therefore, any scaling of the amplitude parameter must also be accompanied by a change in the pulse form, which occurs through a scaling of the map strength parameter. This becomes evident in the exact form for the invariance, given by

$$\tilde{u}(z,t;s_{\text{map}},A) = Au(A^2z, At; A^2 s_{\text{map}}), \tag{5.14}$$

which states that, as long as $u(z,t;s_{\text{map}})$ is a solution of the DMNLSE, than $\tilde{u}(z,t;s_{\text{map}},A)$ is also a solution for any value of A.

Including all four invariances discussed above produces a family of DM solitons of the form

$$\begin{aligned} &u_{\text{dmsol}}(z,t;s_{\text{map}},A,\Omega,T,\Phi) = \\ &\qquad Au_0(A(t-T-\Omega z);A^2 s_{\text{map}})\exp\left(\Omega(t-T)+\frac{A^2+\Omega^2}{2}z+\Phi\right), \end{aligned} \tag{5.15}$$

where $u_0(z,t;s_{\text{map}})$ is the underlying pulse solution, i.e., functional form, which is analogous to the hyperbolic secant pulse form of the NLSE soliton. It is important

to note however, that although easily found numerically and well approximated analytically for large map strengths by a Gaussian pulse form, $u_0(z,t;s_{\text{map}})$ *does not* have a closed-form expression, complicating any attempt at forming a perturbation theory analogous to SPT of the NLSE.

5.2.3 Pulse Solutions of the DMNLSE

Because the DMNLSE does not admit closed-form solutions, the only way of producing a DM soliton is numerically, which is done here through a simple iterative method based on the DMNLSE in the Fourier domain given by Equation (5.10a) [70]. By neglecting the noise terms and assuming a DM soliton of the form

$$u_{\text{dmsol}}(z,t;s_{\text{map}},\lambda) = u_0(t;s_{\text{map}},\lambda)\exp\big(i\lambda^2 z/2\big), \tag{5.16}$$

the z derivative is replaced by the nonlinear eigenvalue λ, giving the equation

$$\begin{aligned}\frac{\lambda^2+\omega^2}{2}\hat{u}_0(\omega;s_{\text{map}},\lambda) = \int\int \hat{u}_0(\omega_1+\omega;s_{\text{map}},\lambda)\hat{u}_0(\omega_2+\omega;s_{\text{map}},\lambda)\times \\ \bar{\hat{u}}_0(\omega_2+\omega_1+\omega;s_{\text{map}},\lambda)\hat{K}(\omega_1,\omega_2;s_{\text{map}})\,d\omega_1\,d\omega_2,\end{aligned} \tag{5.17}$$

which can be solved using an iterative method [70]. Note, by comparing (5.16) to the general four-parameter family of DM solitons in (5.15), it is seen that the eigenvalue λ is a proxy for the amplitude parameter A, which implies that the resulting numerical solution $u_0(t;s_{\text{map}},\lambda)$ can be generated from the amplitude invariance [52], i.e.,

$$u_0(t;s_{\text{map}},\lambda) = \lambda u_0(\lambda t;\lambda^2 s_{\text{map}},1). \tag{5.18}$$

Thus, all DM soliton solutions of Equation (5.17) can be generated from the one parameter family of solutions $u_0(t;s_{\text{map}},1)$ which are plotted in Figure 5.1.

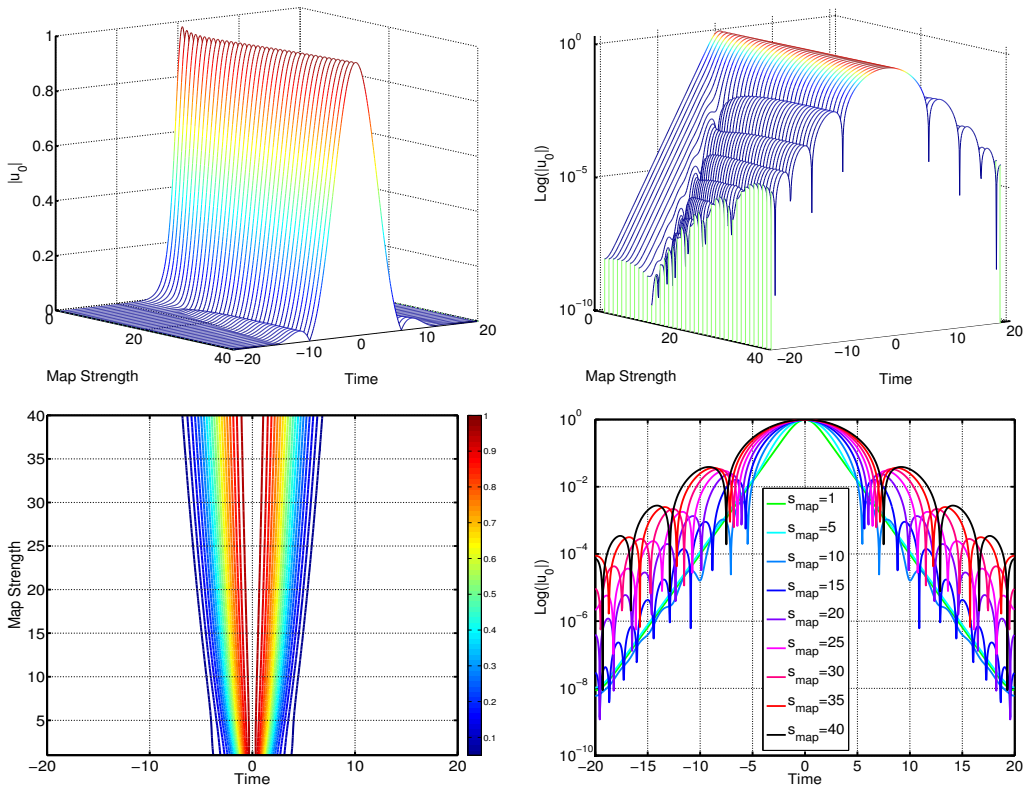


Figure 5.1 Top Left: DM soliton solutions of (5.17) as a function of the map strength parameter. Top Right: DM soliton solutions of (5.17) on Log scale as a function of the map strength parameter. Bottom Left: Contour plots of DM soliton solutions as a function of map strength. Bottom Right: Profile of DM solitons on a Log scale with varying map strength. In all plots $\lambda = 1$.

5.3 DMNLSE Phase Shift from Radiation

The radiation induced phase drift seen in the NLSE of Chapter 2 is also seen in simulations of the stochastically forced DMNLSE given in (5.12a) [51, 52]. In this section, the variational approach from Chapter 3 is extended to the DMNLSE, with the aim of deriving a low-dimensional reduction capable of accounting for this phase drift. However, this case is slightly more complicated by the lack of closed functional form for the DM soliton which, when combined with the approximation for the radiation, introduces additional error in the reduced system.

5.3.1 Reduced Equations through Variational Approach

Following the derivation in Chapter 3 for the NLSE, the solution we are approximating is first written as a sum of two distinct parts, $u = u_{\rm s} + u_{\rm r}$, where $u_{\rm s}$ and $u_{\rm r}$ represent the solitonic and radiative components of the solution, respectively. Putting this into equation (5.12a) and rewriting the resulting equation as a coupled system gives

$$\begin{aligned}
&i\frac{\partial u_{\rm s}}{\partial z} + \frac{1}{2}\frac{\partial^2 u_{\rm s}}{\partial t^2} + \int\int K(t_1,t_2)u_{\rm s}(z,t_1+t)u_{\rm s}(z,t_2+t)\bar{u}_{\rm s}(z,t_1+t_2+t)\,dt_1dt_2 \\
&\quad + 2\int\int K(t_1,t_2)u_{\rm s}(z,t_1+t)u_{\rm r}(z,t_2+t)\bar{u}_{\rm r}(z,t_1+t_2+t)\,dt_1dt_2 = i\sigma n_{\rm s}, \\
&i\frac{\partial u_{\rm r}}{\partial z} + \frac{1}{2}\frac{\partial^2 u_{\rm r}}{\partial t^2} \\
&\quad + 2\int\int K(t_1,t_2)u_{\rm s}(z,t_1+t)u_{\rm r}(z,t_2+t)\bar{u}_{\rm s}(z,t_1+t_2+t)\,dt_1dt_2 = i\sigma n_{\rm r},
\end{aligned} \tag{5.19}$$

where again, the radiation is explicitly split into a portion driving the pulse, $n_{\rm s}$, and portion driving the radiation, $n_{\rm r}$, and the negligible terms corresponding to SPM of the radiation and FWM are discarded. Note that the complete system also includes the conjugates of these equations, which are not shown. The functional representation for this system takes the same form as before, i.e.,

$$J(u_{\rm s},\bar{u}_{\rm s},u_{\rm r},\bar{u}_{\rm r}) = \int_0^{z_f}\int L_{den}(u_{\rm s},\bar{u}_{\rm s},u_{\rm r},\bar{u}_{\rm r})\,dtdz; \tag{5.20a}$$

however, the Lagrangian density is now given as

$$\begin{aligned}
L_{den}(u_{\rm s},\bar{u}_{\rm s},u_{\rm r},\bar{u}_{\rm r}) &= \mathrm{Re}\left[i\frac{\partial u_{\rm s}}{\partial z}\bar{u}_{\rm s}\right] + \mathrm{Re}\left[i\frac{\partial u_{\rm r}}{\partial z}\bar{u}_{\rm r}\right] - \frac{1}{2}\left|\frac{\partial u_{\rm s}}{\partial t}\right|^2 - \frac{1}{2}\left|\frac{\partial u_{\rm r}}{\partial t}\right|^2 \\
&+ \frac{1}{2}\mathrm{Re}\left[\int\int K(t_1,t_2)u_{\rm s}(z,t_1+t)u_{\rm s}(z,t_2+t)\bar{u}_{\rm s}(z,t_1+t_2+t)\bar{u}_{\rm s}(z,t)\,dt_1dt_2\right] \\
&+ 2\mathrm{Re}\left[\int\int K(t_1,t_2)u_{\rm s}(z,t_1+t)u_{\rm r}(z,t_2+t)\bar{u}_{\rm r}(z,t_1+t_2+t)\bar{u}_{\rm s}(z,t)\,dt_1dt_2\right] \\
&- 2\mathrm{Re}\left[\bar{u}_{\rm s}\,in_{\rm s}\right] - 2\mathrm{Re}\left[\bar{u}_{\rm r}\;in_{\rm r}\right].
\end{aligned} \tag{5.20b}$$

At this point, functional forms must be chosen for both the solitonic and radiative parts of the solution. Noting that we are interested in the large map strength regime, and recalling from Section 5.2.3 that the core of DM soliton solutions of the DMNLSE take the form of a Gaussian pulse as the map strength gets large, the solitonic portion of the solution is taken to be a chirped Gaussian pulse,

$$u_{\mathrm{s}}(E, W, T, \Omega, \Phi, C) = \sqrt{\frac{E}{W}} \exp\left(-\frac{\Sigma(t,z)^2}{2}\right) \exp(i\Psi(t,z)) \tag{5.21a}$$

where

$$\Sigma(t,z) = \frac{(t - T(z))}{W(z)} \tag{5.21b}$$

and

$$\Psi(t,z) = \Phi(z) + \Omega(z)(t - T(z)) + \frac{C(z)}{2}(t - T(z))^2. \tag{5.21c}$$

As before, the radiation is represented by a windowed Fourier decomposition,

$$u_{\mathrm{r}}(t, \alpha_1, \alpha_2, \ldots, \alpha_N) = \sum_{n=-N/2}^{N/2-1} \alpha_n(z) \exp(-i\omega_n t)\ \mathrm{H}(w - |t|)\,, \tag{5.22}$$

where the parameters $\alpha_n(z)$ are the Fourier coefficients of each mode and the frequencies are taken to match the numerical frequencies of $\omega_n = \frac{n\pi}{w}$.

Before stating the resulting averaged Lagrangian from these two functional forms, it is instructive to first explicitly state the calculations of the two integral terms in

(5.20b). The first is

$$\begin{aligned}
&\frac{1}{2}\mathrm{Re}\left[\int\int K(t_1,t_2)u_{\mathrm{s}}(z,t_1+t)u_{\mathrm{s}}(z,t_2+t)\bar{u}_{\mathrm{s}}(z,t_1+t_2+t)\bar{u}_{\mathrm{s}}(z,t)\,dt_1dt_2\right]\\
&\quad=\frac{1}{2}\left(\frac{E}{W}\right)^2\mathrm{Re}\left[\int\int\int K(t_1,t_2)\exp(-iCt_1t_2)\times\right.\\
&\quad\left.\exp\left(-\frac{1}{2W^2}\left[(t_1)^2+(t_2)^2+(t_1+t_2)^2\right]-\frac{2}{W^2}\left[(t_1{+}_2)(t-T)+(t-T)^2\right]\right)\,dt_1\,dt_2\,dt\right]\\
&\quad=\frac{\sqrt{\pi}}{2\sqrt{2}}E^2K_1(W,C)
\end{aligned}\tag{5.23a}$$

where

$$K_1(W,C)=\frac{1}{s_{\mathrm{map}}}\frac{1}{2\beta}\ln\left(\frac{\kappa_+-\left(\kappa_+^2+1\right)^{\frac{1}{2}}}{\kappa_+-\left(\kappa_+^2+1\right)^{\frac{1}{2}}}\frac{\kappa_--\left(\kappa_-^2+1\right)^{\frac{1}{2}}}{\kappa_-+\left(\kappa_-^2+1\right)^{\frac{1}{2}}}\right)\tag{5.23b}$$

with

$$\beta^2=\frac{(1+W^4C^2)}{W^2},\quad \kappa_\pm=W^2C\pm\frac{s_{\mathrm{map}}\beta^2}{2}.\tag{5.23c}$$

The second is

$$\begin{aligned}
&2\mathrm{Re}\left[\int\int K(t_1,t_2)u_{\mathrm{s}}(z,t_1+t)u_{\mathrm{r}}(z,t_2+t)\bar{u}_{\mathrm{r}}(z,t_1+t_2+t)\bar{u}_{\mathrm{s}}(z,t)\,dt_1dt_2\right]\\
&\quad\approx 2\frac{E}{W}\sum_{j=-N/2}^{N/2-1}|\alpha_j|^2\mathrm{Re}\left[\int\exp\left(-\frac{1}{W^2}(t-T)^2\right)\,dt\right]=2\sqrt{\pi}E\sum_{j=-N/2}^{N/2-1}|\alpha_j|^2,
\end{aligned}\tag{5.24}$$

where the approximation results from the two assumptions that (i) the soliton is much narrower than the window width and thus the Heaviside functions can be ignored by relying only on the support of the Gaussian soliton for convergence of the integral, i.e., $w \gg W(z)$, and (ii) the non-resonant terms coupling the radiation back to the soliton can be neglected, just as the FWM terms were in the construction of the coupled system in (5.19).

With these two calculations, the averaged Lagrangian takes the form

$$\begin{aligned} L_{avg} = & -\sqrt{\pi}E\dot{\phi} + \sqrt{\pi}E\Omega\dot{T} - \frac{\sqrt{\pi}}{4}EW^2\dot{C} - \frac{\sqrt{\pi}}{4}\frac{E}{W^2} - \frac{\sqrt{\pi}}{2}E\Omega^2 \\ & - \frac{\sqrt{\pi}}{4}EC^2W^2 + iw\sum_{j=-N/2}^{N/2-1}\left[\bar{\alpha}_j\dot{\alpha}_j - \alpha_j\dot{\bar{\alpha}}_j\right] - w\sum_{j=-N/2}^{N/2-1}|\alpha_j|^2\,\omega_j^2 \\ & + 2\sqrt{\pi}E\sum_{j=-N/2}^{N/2-1}|\alpha_j|^2 + \frac{\sqrt{\pi}}{2\sqrt{2}}E^2K_1(W,C) - 2\mathrm{Re}\left[\bar{u}_{\mathrm{s}}\,in_{\mathrm{s}}\right] - 2\mathrm{Re}\left[\bar{u}_{\mathrm{r}}\,in_{\mathrm{r}}\right], \end{aligned} \tag{5.25}$$

which, after applying the Euler-Lagrange equations for each parameter, i.e.,

$$\frac{\partial L_{avg}}{\partial X} - \frac{d}{dz}\left[\frac{\partial L_{avg}}{\partial \dot{X}}\right] = 0 \tag{5.26}$$

for $X = E$, W, T, Ω, Φ and C, gives SODEs of the form

$$\frac{dE}{dz} = \frac{2}{\sqrt{\pi}}\mathrm{Re}\left[\int \frac{\partial \bar{u}_s}{\partial \phi}(in(t,z))\,dt\right], \tag{5.27a}$$

$$\frac{d\Omega}{dz} = -\frac{2}{\sqrt{\pi}E}\left[\mathrm{Re}\left[\int \frac{\partial \bar{u}_s}{\partial T}(in(t,z))\,dt\right] + \Omega\mathrm{Re}\left[\int \frac{\partial \bar{u}_s}{\partial \Phi}(in(t,z))\,dt\right]\right], \tag{5.27b}$$

$$\frac{dT}{dz} = \Omega + \frac{2}{\sqrt{\pi}E}\mathrm{Re}\left[\int \frac{\partial \bar{u}_s}{\partial \Omega}(in(t,z))\,dt\right], \tag{5.27c}$$

$$\frac{dC}{dz} = \frac{1}{W^4} - C^2 + \frac{1}{\sqrt{2}}\frac{E}{W}\frac{\partial K_1}{\partial W} - \frac{4}{\sqrt{\pi}EW}\mathrm{Re}\left[\int \frac{\partial \bar{u}_s}{\partial W}(in(t,z))\,dt\right], \tag{5.27d}$$

$$\begin{aligned} \frac{dW}{dz} = & CW - \frac{1}{\sqrt{2}}\frac{E}{W}\frac{\partial K_1}{\partial C} \\ & + \frac{4}{\sqrt{\pi}EW}\left(\mathrm{Re}\left[\int \frac{\partial \bar{u}_s}{\partial C}(in(t,z))\,dt\right] - \frac{W^2}{4}\mathrm{Re}\left[\int \frac{\partial \bar{u}_s}{\partial \Phi}(in(t,z))\,dt\right]\right), \end{aligned} \tag{5.27e}$$

$$\begin{aligned} \frac{d\Phi}{dz} = & \frac{1}{2}\Omega^2 - \frac{1}{2}\frac{1}{W^2} + \frac{1}{\sqrt{2}}EK_1(W,C) - \frac{1}{4\sqrt{2}}EW\frac{\partial K_1}{\partial W} + 2\sum_{j=-N/2}^{N/2-1}|\alpha_j|^2 \\ & + \frac{W}{\sqrt{\pi}E}\mathrm{Re}\left[\int \frac{\partial \bar{u}_s}{\partial W}(in(t,z))\,dt\right] + \frac{2}{\sqrt{\pi}E}\Omega\mathrm{Re}\left[\int \frac{\partial \bar{u}_s}{\partial \Omega}(in(t,z))\,dt\right] \\ & - \frac{2}{\sqrt{\pi}}\mathrm{Re}\left[\int \frac{\partial \bar{u}_s}{\partial E}(in(t,z))\,dt\right], \end{aligned} \tag{5.27f}$$

where the radiation coefficients satisfy

$$\begin{aligned}\dot{\alpha}_j &= i\left(\frac{\sqrt{\pi}}{w}E - \frac{1}{2}\omega_j^2\right)\alpha_j + \frac{1}{2w}\int \exp(i\omega_j t)\ \mathrm{H}(w-|t|)\, n_{\mathrm{r}}(t,z)\, dt \\ &\approx i\left(\frac{\sqrt{\pi}}{w}E - \frac{1}{2}\omega_j^2\right)\alpha_j + \frac{1}{2w}\int \exp(i\omega_j t)\ \mathrm{H}(w-|t|)\, n(t,z)\, dt,\end{aligned} \tag{5.27g}$$

and

$$\begin{aligned}\dot{\bar{\alpha}}_j &= -i\left(\frac{\sqrt{\pi}}{w}E - \frac{1}{2}\omega_j^2\right)\bar{\alpha}_j + \frac{1}{2w}\int \exp(-i\omega_j t)\ \mathrm{H}(w-|t|)\ \ \bar{n}_r\, dt \\ &=\approx -i\left(\frac{\sqrt{\pi}}{w}E - \frac{1}{2}\omega_j^2\right)\bar{\alpha}_j + \frac{1}{2w}\int \exp(-i\omega_j t)\ \mathrm{H}(w-|t|)\, \bar{n}(t,z)\, dt.\end{aligned} \tag{5.27h}$$

5.3.2 The Approximate DMNLSE Soliton Modes

Equations (5.27) have terms representing the noise projected onto a basis formed from linear combinations of the derivatives of the Gaussian soliton ansatz with respect to each of the *six* parameters, i.e., $\partial u_{\mathrm{s}}/\partial X$ for $X = E$, W, T, Ω, Φ and C. This was also the case in Equations (3.26), however, there the functional form of the soliton ansatz matched the exact soliton solution (which is known in closed form), resulting in the equivalence between the true discrete eigenfunctions (and adjoints) of the linearized NLSE and those functions found through the variational method. Here, however, the DM soliton solutions are not known in closed form, but instead are approximated by a Gaussian ansatz, resulting in Gaussian approximations for the discrete eigenfunctions (and the adjoint eigenfunctions) of the linearized DMNLSE [70].

By explicitly calculating these derivatives and accounting for real, imaginary, even and odd parts, a convenient representation for the approximate eigenfunctions of the linearized DMNLSE is given by

$$\begin{aligned} v_E = \frac{1}{2E}u_{\mathrm{s}}, \quad v_W = \frac{1}{2W}(2\Sigma^2-1)u_{\mathrm{s}}, \quad v_T = \frac{1}{W}\Sigma u_{\mathrm{s}} \\ v_\Omega = iW\Sigma u_{\mathrm{s}}, \quad v_\Phi = iu_{\mathrm{s}}, \quad v_C = i\frac{W^2}{4}(2\Sigma^2-1)u_{\mathrm{s}},\end{aligned} \tag{5.28}$$

which, through the projection terms in Equations in (5.27), results in approximate adjoint eigenfunctions of

$$\begin{aligned} v_E^\dagger &= -i\frac{2}{\sqrt{\pi}}v_\Phi, \quad v_\Phi^\dagger = i\frac{2}{\sqrt{\pi}}v_E, \quad v_T^\dagger = -i\frac{2}{\sqrt{\pi}E}v_\Omega, \\ v_\Omega^\dagger &= i\frac{2}{\sqrt{\pi}E}v_T, \quad v_W^\dagger = -i\frac{4}{\sqrt{\pi}EW}v_C, \quad v_C^\dagger = i\frac{4}{\sqrt{\pi}EW}v_W. \end{aligned} \tag{5.29}$$

Note that under the inner product defined in Equation (2.16), i.e.,

$$\langle f, g\rangle = \mathrm{Re}\left[\int \bar{f}\, g\, dt\right], \tag{5.30}$$

the approximate eigenfunctions given in Equation (5.28) form an orthogonal set, i.e.,

$$\langle v_X, v_Y\rangle = \langle v_Y, v_X\rangle\, \delta_{XY}, \tag{5.31}$$

for $X, Y = E$, W, T, Ω, C and Φ. Likewise, the adjoint eigenfunctions given by Equations (5.29) are also orthogonal, and thus, the these two sets together form a bi-orthogonal basis, i.e.,

$$\left\langle v_X^\dagger, v_Y\right\rangle = \left\langle v_Y, v_X^\dagger\right\rangle \delta_{XY}. \tag{5.32}$$

Using these relations and the discrete form of noise given in Equation (2.2), the SODEs 5.27 can be written as

$$\frac{dE}{dz} = \sum_{k=1}^{N_\mathrm{a}} \Delta E_k\, \delta(z - kz_\mathrm{a}), \tag{5.33a}$$

$$\frac{d\Omega}{dz} = \sum_{k=1}^{N_\mathrm{a}} \Delta \Omega_k\, \delta(z - kz_\mathrm{a}), \tag{5.33b}$$

$$\frac{dT}{dz} = \Omega + \sum_{k=1}^{N_\mathrm{a}} \Delta T_k\, \delta(z - kz_\mathrm{a}), \tag{5.33c}$$

$$\frac{dW}{dz} = CW - \frac{1}{\sqrt{2}}\frac{E}{W}\frac{\partial K_1}{\partial C} + \sum_{k=1}^{N_\mathrm{a}} \Delta W_k\, \delta(z - kz_\mathrm{a}), \tag{5.33d}$$

$$\frac{dC}{dz} = \frac{1}{W^4} - C^2 + \frac{1}{\sqrt{2}}\frac{E}{W}\frac{\partial K_1}{\partial W} + \sum_{k=1}^{N_\mathrm{a}} \Delta C_k\, \delta(z - kz_\mathrm{a}), \tag{5.33e}$$

$$\begin{aligned}\frac{d\Phi}{dz} &= \frac{1}{2}\Omega^2 - \frac{1}{2}\frac{1}{W^2} + \frac{1}{\sqrt{2}}EK_1 \\ &\quad - \frac{1}{4\sqrt{2}}EW\frac{\partial K_1}{\partial W} + 2\sum_{j=-N/2}^{N/2-1} |\alpha_j|^2 + \sum_{k=1}^{N_\mathrm{a}} \Delta\Phi_k\, \delta(z - kz_\mathrm{a}),\end{aligned} \tag{5.33f}$$

$$\frac{d\alpha_j}{dz} \approx i\left(\frac{\sqrt{\pi}}{w}E - \frac{1}{2}\omega_j^2\right)\alpha_j + \sum_{k=1}^{N_\mathrm{a}} \Delta\alpha_{j,k}\, \delta(z - kz_\mathrm{a}), \tag{5.33g}$$

and

$$\frac{d\bar{\alpha}_j}{dz} \approx -i\left(\frac{\sqrt{\pi}}{w}E - \frac{1}{2}\omega_j^2\right)\bar{\alpha}_j + \sum_{k=1}^{N_\mathrm{a}} \Delta\bar{\alpha}_{j,k}\, \delta(z - kz_\mathrm{a}). \tag{5.33h}$$

where the stochastic jump terms are given by evaluating the projection functions

$$\begin{aligned}
\Delta E_k &= \sigma\mathrm{Re}\left[\int \bar{v}_E^\dagger(kz_\mathrm{a}, t)\exp(-i\Psi(kz_\mathrm{a}, t))\, n_k(t)\, dt\right], \\
\Delta\Omega_k &= \sigma\mathrm{Re}\left[\int \left(\bar{v}_\Omega^\dagger(kz_\mathrm{a}, t) + C(kz_\mathrm{a})\bar{v}_T^\dagger(kz_\mathrm{a}, t)\right)\exp(-i\Psi(kz_\mathrm{a}, t))\, n_k(t)\, dt\right], \\
\Delta T_k &= \sigma\mathrm{Re}\left[\int \bar{v}_T^\dagger(kz_\mathrm{a}, t)\exp(-i\Psi(kz_\mathrm{a}, t))\, n_k(t)\, dt\right], \\
\Delta W_k &= \sigma\mathrm{Re}\left[\int \bar{v}_W^\dagger(kz_\mathrm{a}, t)\exp(-i\Psi(kz_\mathrm{a}, t))\, n_k(t)\, dt\right], \\
\Delta C_k &= \sigma\mathrm{Re}\left[\int \bar{v}_C^\dagger(kz_\mathrm{a}, t)\exp(-i\Psi(kz_\mathrm{a}, t))\, n_k(t)\, dt\right], \\
\Delta\Phi_k &= \sigma\mathrm{Re}\left[\int \left(\bar{v}_\Phi^\dagger(kz_\mathrm{a}, t) + \Omega(kz_\mathrm{a})\bar{v}_T^\dagger(kz_\mathrm{a}, t)\right.\right. \\
&\qquad\qquad \left.\left. - \frac{W(kz_\mathrm{a})^2}{4}\bar{v}_C^\dagger(kz_\mathrm{a}, t)\right)\exp(-i\Psi(kz_\mathrm{a}, t))\, n_k(t)\, dt\right], \\
\Delta\alpha_{j,k} &= \frac{\sigma}{2w}\int_{-w}^{w} \exp(i\omega_j t)\; n_k(t)\, dt, \\
\Delta\bar{\alpha}_{j,k} &= \frac{\sigma}{2w}\int_{-w}^{w} \exp(-i\omega_j t)\; \bar{n}_k(t)\, dt,
\end{aligned} \tag{5.34}$$

at the amplification points kz_a. Like in the case of the NLSE, these jump terms are mean zero. Unlike the NLSE, however, there are many covariances between these

terms, which are best displayed as

$$\mathbb{E}\left[\begin{bmatrix}\Delta E_{k+1}\\ \Delta W_{k+1}\\ \Delta T_{k+1}\\ \Delta C_{k+1}\\ \Delta \Omega_{k+1}\\ \Delta \Phi_{k+1}\end{bmatrix}\begin{bmatrix}\Delta E_{k+1}\\ \Delta W_{k+1}\\ \Delta T_{k+1}\\ \Delta C_{k+1}\\ \Delta \Omega_{k+1}\\ \Delta \Phi_{k+1}\end{bmatrix}^T\right] = \tag{5.35}$$

$$\frac{\sigma^2}{\sqrt{\pi}}\begin{bmatrix} 2E_k & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{W_k^2}{E_k} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{W_k^2}{E_k} & 0 & \frac{C_k W_k^2}{E_k} & \frac{\Omega W_k^2}{E_k} \\ 0 & 0 & 0 & \frac{4}{E_k W_k^4} & 0 & -\frac{1}{2E_k W_k^2} \\ 0 & 0 & \frac{C_k W_k^2}{E_k} & 0 & \frac{1}{E_k}\left(\frac{1}{W_k^2}+C_k^2 W_k^2\right) & \frac{C_k \Omega_k W_k^2}{E_k} \\ 0 & 0 & \frac{\Omega_k W_k^2}{E_k} & -\frac{1}{2E_k W_k^2} & \frac{C_k \Omega_k W_k^2}{E_k} & \frac{1}{E_k}\left(\frac{3}{4}+\Omega_k^2 W_k^2\right) \end{bmatrix}, \tag{5.36}$$

and

$$\mathbb{E}[\Delta\alpha_{j,k+1}\Delta\bar{\alpha}_{j,k+1}] = \frac{\sigma^2}{2w}, \tag{5.37}$$

where the notation $X_k = X(kz_\mathrm{a})$ was used for brevity.

5.4 Results and Discussion

The phase evolution given by Equation (5.33f) can now be compared to the evolution of the phase of a DM soliton. The results of this are given in Figure 5.2, and are seen to approximate the phase evolution very well.

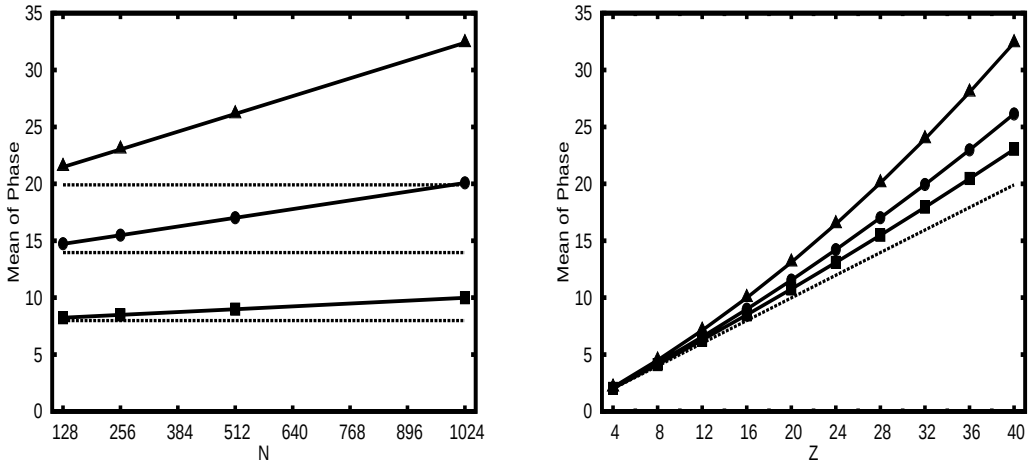


Figure 5.2 The plot on the left gives the mean phase vs. number of simulation modes using the SODEs in (5.27) (solid lines) and the DMNLSE given in (5.12a) (markers) at $z = 16$ (squares), 28 (circles) and 40 (triangles) with $s_{\text{map}} = 5$. The plot to the right given the mean of phase vs. transmission length using the SODEs in (5.27) (solid lines) and the DMNLSE given in (5.12a) (markers) for $N = 256$ (squares), 512 (circles), and 1024 (triangles) with $s_{\text{map}} = 5$. For comparison, the dashed line correspond to the first-order SPT approximation.

5.5 Summary

In this chapter, the NLSE was extended to the case of varying dispersion and nonlinear coefficients to construct a model for DM soliton propagation in DM fiber, i.e., NLSE+DM. This equation is close to the form of the NLSE, the resulting DM soliton solutions for this equation are very different from the soliton solutions of NLSE, due to the complex breathing dynamics resulting from the periodic variations in the dispersion coefficient. By applying an averaging technique, where the propagation of a DM soliton is averaged over one period of the variations, the fast breathing dynamics of the DM soliton was separated the from the slower evolution of the pulse "core", the latter of which was shown to satisfy an averaged equation referred to as the DMNLSE [50]. The DMNLSE no longer has periodically varying coefficients due to the averaging, which also resulted in the introduction of a non-local nonlinearity that takes the form of a convolution. This chapter continued by constructing a

low-dimensional reduction for the pulse dynamics in a stochastically forced versions of DMNLSE, which captures the radiation-induced phase drift when noise is included in the propagation model.

CHAPTER 6

REDUCED SYSTEMS FOR MODELS OF MODE-LOCKED LASERS

In previous chapters, the ISMC method was used to quantify rare event probabilities in optical communication systems modeled by variations of the stochastic NLSE. In this chapter, this ISMC method is applied to quantify failure rates in mode-locked lasers (MLL) operating in the soliton and dispersion-managed (DM) soliton (or stretched pulse) regimes. There are many similarities shared by the mathematical models commonly used to describe these two systems. Both include a balance between dispersion and nonlinearity that facilitates the propagation of optical pulses, and both include random perturbations from ASE noise. In addition, MLLs commonly implement the effects of dispersion and nonlinearity separately, through the combination of various separate optical components, i.e., lenses, mirrors and prisms, which results in a propagation model that includes space-dependent terms completely analogous to those of the NLSE+DM of optical communication presented in Chapter 5. However, there are also important differences between the two system, the majority of which stem from the amplification and modulation mechanisms that are essential in the formation of optical pulses in MLLs.

This chapter begins with a description of MLLs in both the time and frequency domains, which is followed by a review of several commonly used mathematical models for pulse propagation. This culminates with the presentation of a relatively new model, termed the dispersion-managed perturbed nonlinear Schrödinger equation (PNLSE+DM), that accounts for nonlinearity and dispersion variations inside the laser cavity, in addition to saturated modulation and bandwidth limited gain. In the soliton operating regime, this equation collapses to the constant-coefficient NLSE with non-conservative perturbations (PNLSE), which is amendable to an application

of the SPT presented in Chapter 2. This results in a low-dimensional reduction for the PNLSE, which is used to guide the ISMC method.

6.1 Mode-Locked Lasers in the Time Domain

In the simplest terms, a MLL is a laser that emits pulsed light in the time domain. The term mode-locking refers to a fixed phase relationship between axial modes in a laser cavity [76]. Under normal operation, the phase of each axial mode in a laser cavity is random, whereas in a mode-locked state, they are locked to a common value, thus aligning the modes and forming pulsed emissions. Though described as a resonant phenomenon occurring in the frequency domain, mode-locking is physically achieved through a power feed-back mechanism applied in the time domain. This was originally done with mechanically controlled power modulators, which is referred to as *active* mode-locking. However, the need for shorter pulses and higher repetition rates has lead to passive modulation designs, which rely on intensity-dependent attenuation, i.e., saturable absorption, to induce mode-locking. At the same time, the gain medium used for maintaining short pulses has also evolved to include an array of materials with extended bandwidths. Two of the more successful materials used for this purpose include Ti:Sapphire, commonly used in open cavity MLLs, and erbium-doped silica fiber, used in optical fiber-based MLLs [77].

The addition of a periodic modulation, in the form of saturable absorption, to a normally operating laser, forces the circulating radiation to coalesce into a pulse that is shortened and amplified on every pass through the cavity. The shortening process continues until the spectrum of the pulse approaches the finite bandwidth of the gain material and settles into a state where the effects of gain, loss, dispersion and nonlinearity are balanced in each round trip through the cavity. Thus, the combination of saturated loss from the modulation and broad-bandwidth gain from amplification can be thought of as creating a basin of attraction for the formation

and stabilization of an optical pulse. As a result, MLLs are extremely stable and accurate devices, making them useful in applications that require extreme precision over extended periods of time. However, these very attractive properties of MLLs also present a difficult challenge when it is necessary to quantify their failure rate in the presence of noise. Even the simplest models are sufficiently complex to prevent their analysis using probabilistic methods, and the precision and stability of MLLs demand that failures be extremely rare events, rendering ensemble approaches based on numerical simulations prohibitively expensive, thus the need for an extension of the ISMC method.

It is important to note that, not all MLLs produce optical pulses in the manner discussed above and, in general, there exists three regimes for optical pulse propagation within these devices. The first is the *soliton* regime, which primarily relies on a balance between dispersion and nonlinearity, with saturated loss and gain effects contributing as perturbations. However, self-phase modulation (SPM) causes the higher intensity portions of a soliton pulse to rotate faster than lower intensity tails, which results in wave breaking and/or multi-pulsing. Thus, there is a limit to the amount of power a single soliton pulse can carry in a MLL cavity. To create pulses with higher power, MLL laser designs started to incorporate the idea of dispersion management from optical communications, where different components of the laser contribute large amounts of either positive or negative dispersion. On average, the pulse receives just enough dispersion to counter the effects of nonlinearity, while the periodically alternating values of instantaneous dispersion forces the pulse to rapidly stretch and compress during one round trip through the laser cavity, which is referred to as stretched pulse or DM soliton propagation. During the time the pulse is stretched, the power is spread out over a longer time interval, thus reducing both the local intensity and SPM effect, and allowing for higher optical powers. Finally, a more recent approach to pulse propagation in MLLs has emerged that uses the self-similar

solution for pulse propagation in gain fiber, which results in the formation of so-called similariton pulses. These lasers are almost entirely constructed out of gain material, and incorporate rapid loss through chirped filtering [78, 79, 80]. The rapid transition to and from states of high gain and loss, results in complex pulse dynamics which are extremely difficult to capture with low-dimensional approximations. Thus, this regime is not considered here.

6.2 Mode-Locked Lasers in the Frequency Domain

As stated above, the term mode-locking refers to a fixed phase relationship between the axial modes in a laser cavity [76], which through interference, produces a train of pulses at integer multiples of the round trip time in the temporal domain. In the spectral domain, this output results in a series of sharp spikes at integer multiples of the repetition frequency, which is referred to as a *frequency comb* [81].

Importantly, the formation of a stable frequency comb requires mode-locking of two elements. First, the longitudinal cavity modes must be locked to produce pulsed output, and second, the phase differences between the carrier wave and the envelope must be locked to ensure that the spectral lines that form the comb have narrow enough line-width to be differentiated as individual spikes. To see this more clearly, consider

$$\begin{aligned} E(t) &= \sum_{n=0}^{N-1} u(t-n\tau)\exp(in\Delta\phi_e)\ \exp(i\omega_c t)\ \exp(i\phi_{r,n}) + c.c. \\ &= \sum_{n=0}^{N-1} u(t-n\tau)\exp(i(\omega_c(t-n\tau) + n\Delta\phi_{ce} + \phi_{r,n})) + c.c., \end{aligned} \tag{6.1}$$

which represents the pulse train leaving the cavity. The laser cavity is assumed to have a round trip time of τ, resulting in a laser output that consists of a sequence of pulses, each centered at integer multiples of τ. In this representation, $u(t)$ is the form of the pulse circulating in the laser cavity, $\exp(i\omega_c t)$ is the carrier wave which is assumed to have a zero initial phase, $\phi_{r,n}$ represents random phase perturbations

from noise and $n\Delta\phi_{ce} = n\Delta\phi_e + n\omega_c\tau$ is the phase at the peak of the nth pulse, which is the sum of phase rotations of the pulse ($n\Delta\phi_e$) and carrier wave ($n\omega_c\tau$). It should also be noted that, since we are primarily concerned with perturbations to the phase, this representation assumes no perturbations of the pulse timing, frequency or amplitude, which could be included by allowing the pulses to depend on timing, frequency and amplitude parameters that vary with the index n. A plot of Equation (6.1) in the absence of random phase perturbations is given in Figure 6.1.

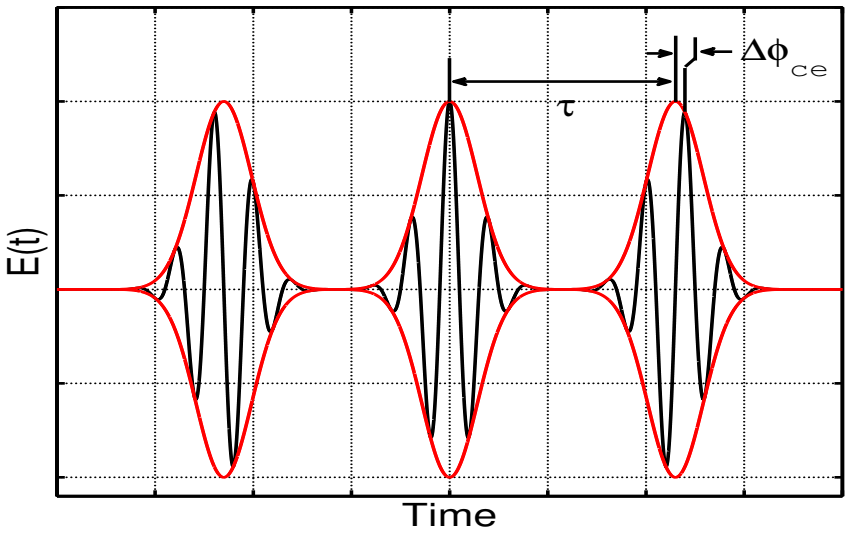


Figure 6.1 Illustration of the enveloped carrier wave output of a mode-locked laser. The red curve represents the envelope and the black curve represents the carrier wave.

The frequency comb that is generated by a train of N pulses can be found from calculating the Fourier transform (FT) of Equation (6.1),

$$\begin{aligned}\mathcal{F}[E(t)] = \hat{E}(\omega-\omega_c) &= \hat{U}(\omega-\omega_c)\sum_{n=0}^{N-1}\exp(-in\omega\tau + in\Delta\phi_{ce} + i\phi_{r,n}) + c.c. \\ &= \hat{U}(\omega-\omega_c)\hat{C}(\omega) + c.c.\end{aligned} \tag{6.2}$$

where the FT was defined as

$$\begin{aligned}\mathcal{F}[f(t)] = \hat{f}(\omega) &= \int f(t)\exp(-i\omega t)\ dt, \\ \mathcal{F}^{-1}\Big[\hat{f}(\omega)\Big] = f(t) &= \frac{1}{2\pi}\int \hat{f}(\omega)\exp(i\omega t)\ d\omega,\end{aligned} \tag{6.3}$$

and $\hat{C}(\omega)$ represents the *comb function*, defined as

$$\hat{C}(\omega) = \sum_{n=0}^{N-1} \exp(-in(\omega\tau - \Delta\phi_{ce}) + i\phi_{r,n})\,. \tag{6.4}$$

In the absence of random phase perturbations, i.e., $\phi_{ce,n} = 0$, the comb function takes the form

$$\begin{aligned}\hat{C}(\omega) &= \sum_{n=0}^{N-1} \exp(-in(\omega\tau - \Delta\phi_{ce})) \\ &= \exp\left(-i\frac{(\omega\tau - \Delta\phi_{ce})}{2}(N-1)\right) \frac{\sin\left(\frac{(\omega\tau-\Delta\phi_{ce})}{2}N\right)}{\sin\left(\frac{(\omega\tau-\Delta\phi_{ce})}{2}\right)},\end{aligned} \tag{6.5}$$

which gives a power spectral density (PSD) of

$$\begin{aligned}|\hat{E}(\omega)|^2 &= |\hat{U}(\omega-\omega_c)\hat{C}(\omega) + c.c.|^2 \approx 2|\hat{U}(\omega-\omega_c)|^2|\hat{C}(\omega)|^2 \\ &= 2|\hat{U}(\omega-\omega_c)|^2 \left(\frac{\sin\left(\frac{(\omega\tau-\Delta\phi_{ce})}{2}N\right)}{\sin\left(\frac{(\omega\tau-\Delta\phi_{ce})}{2}\right)}\right)^2.\end{aligned} \tag{6.6}$$

In the limit of $N \to \infty$, Equation (6.6) is seen to form sharp spectral peaks at $\omega_k = 2\pi k/\tau + \Delta\phi_{ce}/\tau = 2\pi f_{rep}k + \Delta\phi_{ce}f_{rep}$ with frequency spacing of $\Delta\omega = 2\pi/\tau = 2\pi f_{rep}$, where f_{rep} is repetition frequency of the pulses in the cavity. Thus, as illustrated in Figure 6.2, the PSD is in the form of many sharp, evenly spaced spectral peaks which are shifted to account for the carrier-envelope phase offset and enveloped by the broad spectral profile of the pulse centered at the carrier frequency.

6.3 Mode-Locked Laser Models

As previously stated, this chapter is focused on MLLs operating in the soliton and dispersion-managed (DM) soliton (stretched-pulse) regimes. It is assumed that the feedback (modulation) is supplied by a fast saturable absorber reacting instantaneously to changes in the pulse intensity, while the gain element is band-limited and slowly saturates with the total energy in the cavity [77]. As the resulting model

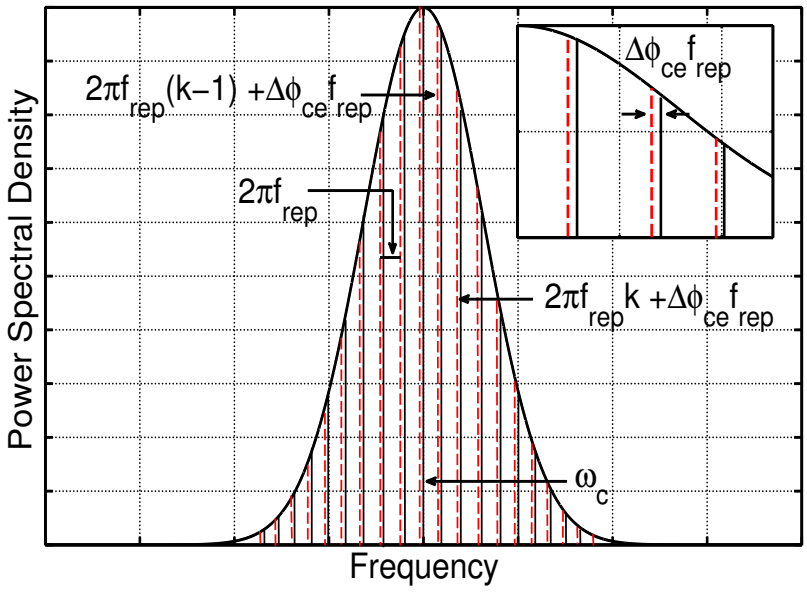


Figure 6.2 Illustration of the frequency comb output of a mode-locked laser. The red lines represent shifted combs spikes, enveloped by the spectral profile of the pulse.

will show, this combination of gain element and saturable absorber creates a larger positive net gain on the peak intensity portions of the pulse which acts to narrow the pulse until the spectral bandwidth of the gain medium is reached and the pulse settles into an equilibrium state. Figure 6.3 shows a diagram for the gain dynamics resulting from this combination.

It is important to note, however, that a complete description of a particular MLL design depends on much more than just the implementation of the modulation and gain. Moreover, there exist numerous MLL designs that use the same modulation and gain elements, but in different combinations and proportions, to produce a variety of different pulse dynamics. Because of this, the most common mathematical models for MLLs are constructed phenomenologically from existing models of optical pulse propagation in nonlinear materials and usually include a large number of parameters that allow the models to be "tuned". In addition, the majority of these models are categorized as *distributed models* in that they do not directly account for the evolution

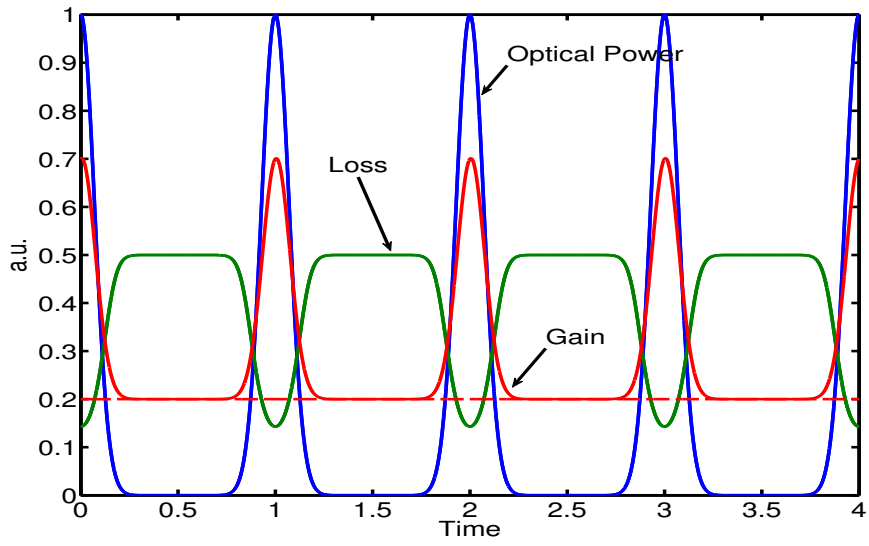


Figure 6.3 Gain dynamics for the combination of passive (fast) saturable absorber and (slow) saturable gain.

of the pulse through different components of the laser cavity, but instead assume that the effect of each component can be modeled as occurring simultaneously throughout the laser cavity. This treatment is formally equivalent to keeping the zeroth order equation from an averaging procedure, such as that presented in Chapter 4 for the derivation of DMNLSE.

6.3.1 Commonly Used MLL Models

The most widely known model for MLLs is the master mode-locking equation (MMLE),

$$i\frac{\partial u}{\partial z} + \frac{1}{2}\frac{\partial u}{\partial tt} + |u|^2u + ig_0u + ig_1|u|^2u + i\frac{2g_0(u - \tau\frac{\partial u}{\partial tt})}{1 + \|u\|^2/E_0} = 0, \tag{6.7}$$

originally derived by Haus [82, 83, 76, 77]. Note that $\|u\|^2 = \int |u|^2 dt$. The MMLE models the effects of nonlinearity and dispersion through same terms as in the NLSE for pulse propagation in optical fiber. It also models bandwidth-limited gain through an energy saturation term and intensity discrimination through a cubic nonlinearity. For a narrow range of these parameters, this equation has stable pulse solutions

with mode-locking evolution [77]. Otherwise the pulses are found to be unstable; either dispersing to radiation, evolving into non-localized quasi-periodic states, or grow rapidly under evolution [77]. Thus, the basic MMLE captures some qualitative aspects of pulse propagation in a laser cavity, but only possesses as a small range of the parameter space for which stable mode-locked pulses exist.

Various modifications have been proposed to extend the parameter space over which stable pulse are formed in the MMLE. These range from the inclusion of higher-order nonlinear terms to more complex models for the absorber [77]. A commonly used extension of the MMLE includes a quintic nonlinear term resulting in the quintic master mode-locking equation (QMMLE) given by

$$i\frac{\partial u}{\partial z} + \frac{1}{2}\frac{\partial u}{\partial tt} + |u|^2u + ig_0u + ig_1|u|^2u + ig_2|u|^4u + i\frac{2g_0(u - \tau\frac{\partial u}{\partial tt})}{1 + \|u\|^2/E_0} = 0. \tag{6.8}$$

The stability of the pulses in both the MMLE and the QMMLE are verified numerically, since the energy saturated gain term complicates the attempts at rigorous stability analysis of these equations. When rigorous stability statements are needed, it is common to approximate the saturating term as a constant, which leads to Ginzburg-Landau type equations, such as the Quintic Complex Ginzburg-Landau Equation (QCGLE) given by

$$i\frac{\partial u}{\partial z} + \left(\frac{1}{2} - ig_0\tau\right)\frac{\partial u}{\partial tt} - i\kappa u_{tttt} + (1 + ig_1)|u|^2u + (\nu + ig_2)\,|u|^4u + i2g_0u = 0, \tag{6.9}$$

which support many types of pulsating, chaotic, and periodically growing or decaying localized states, some of which can be seen in physical MLL systems [84, 76, 77, 55, 81].

6.3.2 The Perturbed NLSE as a Model for MLLs

The model that will be used here for the extension of the ISMC method to MLLs, is a stochastically forced, dispersion managed version of the perturbed nonlinearity

Schrödinger equation [85](PNLSE+DM), given in dimensionless form by

$$\begin{aligned} &i\frac{\partial u}{\partial z} + d(z)\frac{\partial^2 u}{\partial t^2} + \nu(z)|u|^2 u = \\ &i\left(\frac{g(z)}{1+\frac{E(u)}{E_{sat}}}\right)u + i\left(\frac{\tau(z)}{1+\frac{E(u)}{E_{sat}}}\right)\frac{\partial^2 u}{\partial t^2} - i\left(\frac{l(z)}{1+\frac{P(u)}{P_{sat}}}\right)u + i\sigma n(t,z). \end{aligned} \tag{6.10a}$$

where

$$P(u) = |u|^2 \quad \text{and} \quad E(u) = \int |u|^2 dt. \tag{6.10b}$$

The coefficient functions are all taken to be piecewise constant, with $d(z)$ and $\nu(z)$ accounting for variations in dispersion and nonlinearity, respectively, just as in the NLSE+DM discussed in Chapter 4. However, unlike NLSE+DM, the functions $g(z)$ and $\tau(z)$ quantify the strength of the energy-saturated gain and its bandwidth, respectively, while $l(z)$ accounts for power saturated loss. The noise term is written in a general form as $n(t,z)$ and is assumed to have statistics similar to what was used in Chapter 2 for the stochastic NLSE. This equation was originally considered, in the absence of noise, for propagation of large amplitude self-similar pulses, i.e., similaritons, in Ti:sapphire lasers with only normal dispersion [78, 86] and later extended to soliton and DM soliton-based lasers [85], which are the lasers of interest here.

6.4 The Soliton Propagation Regime

The assumption of operating in a soliton regime is equivalent to the assumption that the pulse shape remains relatively constant throughout propagation. To model this type of pulse propagation, the variation in the coefficients of the PNLSE+DM can approximated as perturbative constants. Thus, the equation used here for MLLs operating in a soliton regime (PNLSE) takes the form of a stochastic NLSE with

non-conservative perturbation terms,

$$\begin{aligned} &i\frac{\partial u}{\partial z}+\frac{1}{2}\frac{\partial^2 u}{\partial t^2}+|u|^2u= \\ &i\left(\frac{g}{1+\frac{E(u)}{E_{sat}}}\right)u+i\left(\frac{\tau}{1+\frac{E(u)}{E_{sat}}}\right)\frac{\partial^2 u}{\partial t^2}-i\left(\frac{l}{1+\frac{P(u)}{P_{sat}}}\right)u+i\sigma n(t,z), \end{aligned} \tag{6.11}$$

where $g \sim \tau \sim l \sim \sigma \ll 1$, and it is assumed that $d(z) = 1/2$ and $\nu(z) = 1$.

6.4.1 Derivation of a Reduced System

Because the system in Equation (6.11) is in the form a perturbed NLSE, the resulting solutions take the form of perturbed hyperbolic secant solitons, and thus, the SPT of Chapter 2 can be extended to derive a reduced system for the evolution of the pulse parameters. Now, however, inclusion of the non-conservative gain and loss terms introduce restoring forces to both the amplitude and frequency parameters. In the absence of stochastic effects, these parameter now converge to steady state values fixed by the coefficients g, τ and l.

Recalling that SPT is essentially an application of multiple scales, we introduce the length scale $z_1 = \sigma z$, which transforms Equation (6.11) to

$$\begin{aligned} &i\frac{\partial u}{\partial z}+i\sigma\frac{\partial u}{\partial z_1}+\frac{1}{2}\frac{\partial^2 u}{\partial t^2}+|u|^2u= \\ &i\sigma\left(\frac{\tilde{g}}{1+\frac{E(u)}{E_{sat}}}\right)u+i\sigma\left(\frac{\tilde{\tau}}{1+\frac{E(u)}{E_{sat}}}\right)\frac{\partial^2 u}{\partial t^2}-i\sigma\left(\frac{\tilde{l}}{1+\frac{P(u)}{P_{sat}}}\right)u+i\sigma n(t,z), \end{aligned} \tag{6.12}$$

where $\tilde{g}$, $\tilde{\tau}$ and $\tilde{l}$ are now $O(1)$ quantities. Inserting an expansion for the solution in the form

$$u = \left[v_0(t,z,z_1) + \sigma v_1(t,z,z_1)\right]\exp(i\Theta(t,z,z_1))\,, \tag{6.13}$$

gives the NLSE as the $O(1)$ equation, which as expected, gives the first order solution of a soliton with parameters that now depend on z_1,

$$v_0(t,z,z_1)\exp(i\Theta(t,z,z_1)) = u_{sol}(t,z,A(z_1),\Omega(z_1),T_0(z_1),\Phi_0(z_1)), \tag{6.14}$$

where u_{sol} is defined in Equation (2.4) and repeated here for convenience as

$$u_{sol}(z,t) = u_0(z,t)\exp(i\Theta(z,t))\,, \tag{6.15a}$$

where

$$u_0(z,t) = A\ \mathrm{sech}(A[t-T(z)])\,, \qquad \Theta(z,t) = \Omega\, t + \Phi(z), \tag{6.15b}$$

with

$$T(z) = T_0 + \Omega z \quad \text{ and } \quad \Phi(z) = \frac{A^2-\Omega^2}{2}z + \Phi_0. \tag{6.15c}$$

At $O(\sigma)$, the expansion gives

$$\begin{aligned}\left(\frac{\partial}{\partial z} + \Omega\frac{\partial}{\partial t}\right)v_1 - L(v_1;u_0) &= n(z,t)\exp(-i\Theta(t,z,z_1)) + \left(\frac{\tilde{g}}{1+\frac{E(u_0)}{E_{sat}}}\right)u_0 \\ &+ \left(\frac{\tilde{\tau}}{1+\frac{E(u_0)}{E_{sat}}}\right)\left[\frac{\partial^2 u_0}{\partial t^2} + 2i\Omega\frac{\partial u_0}{\partial t} - u_0\Omega\right] - \left(\frac{\tilde{l}}{1+\frac{P(u_0)}{P_{sat}}}\right)u_0 \\ &- \left[v_A\frac{dA}{dz_1} + v_T\frac{\partial T}{\partial z_1} + (v_\Omega + Tv_\Phi)\frac{d\Omega}{dz_1} + v_\Phi\frac{\partial\Phi}{\partial z_1}\right],\end{aligned} \tag{6.16a}$$

where as before,

$$L(v_1;u_0) = \frac{i}{2}\frac{\partial^2 v_1}{\partial t^2} - \frac{i}{2}A^2 v_1 + 2i|u_0|^2 v_1 + i(u_0)^2\bar{v}_1, \tag{6.16b}$$

is the reduced linearized operator resulting from the linearization of the NLSE about the soliton solution and v_X, for $X = A$, Ω, T and Φ, are the general eigenfunctions corresponding to the soliton parameters. Recall that these eigenfunctions form an orthonormal basis with respect to the corresponding adjoint eigenfunctions, denoted by $v_X^\dagger$ and defined in Equation (2.18), under the inner product defined by Equation (2.16). Thus, enforcement of the Fredholm orthogonality condition for Equation (6.16a) gives the reduces system

$$\frac{dA}{dz} = g_A(A,\Omega) + \sigma\mathrm{Re}\left[\int \bar{v}_A^\dagger \exp(-i\Theta)\, n(z,t)\, dt\right], \tag{6.17a}$$

$$\frac{dT}{dz} = \Omega + \sigma \mathrm{Re}\left[\int \bar{v}_T^\dagger \exp(-i\Theta)\, n(z,t)\, dt\right], \tag{6.17b}$$

$$\frac{d\Omega}{dz} = -g_\Omega(A,\Omega) + \sigma \mathrm{Re}\left[\int \bar{v}_\Omega^\dagger \exp(-i\Theta)\, n(z,t)\, dt\right], \tag{6.17c}$$

and

$$\begin{aligned}\frac{d\Phi}{dz} = \frac{A^2-\Omega^2}{2} &+ g_\Omega(A,\Omega)T \\ &+ \sigma \mathrm{Re}\left[\int \left(\bar{v}_\Phi^\dagger - T\bar{v}_\Omega^\dagger\right) \exp(-i\Theta)\, n(z,t)\, dt\right],\end{aligned} \tag{6.17d}$$

where $g_A(A,\Omega)$ and $g_\Omega(A,\Omega)$ are functions of the amplitude and frequency parameters given by

$$\begin{aligned}g_A(A,\Omega) = 2gE_{sat}\frac{A}{E_{sat}+2A} - 2l\frac{P_{sat}}{(P_{sat}+A^2)^{\frac{1}{2}}}\operatorname{arctanh}&\left(\frac{A}{(P_{sat}+A^2)^{\frac{1}{2}}}\right) \\ &- 2\tau E_{sat}\frac{A}{E_{sat}+2A}\left(\frac{A^3}{3}+\Omega^2\right),\end{aligned} \tag{6.17e}$$

and

$$g_\Omega(A,\Omega) = \frac{4}{3}\tau E_{sat}\frac{A^2}{E_{sat}+2A}\Omega, \tag{6.17f}$$

which originate from the new perturbation terms included in the reduction. Finally, using the discrete amplification model for the noise given in Equation (2.2), the stochasticity appears as discrete random jumps in the soliton parameters which take the form of projections between the noise realizations at each amplifier and the four adjoint modes of the linearization,

$$\frac{dA}{dz} = g_A(A,\Omega) + \sum_{k=1}^{N_a} \Delta A_k\, \delta(z - kz_a), \tag{6.18a}$$

$$\frac{dT}{dz} = \Omega + \sum_{k=1}^{N_a} \Delta T_k\, \delta(z - kz_a), \tag{6.18b}$$

$$\frac{d\Omega}{dz} = -g_\Omega(A,\Omega) + \sum_{k=1}^{N_a} \Delta\Omega_k\, \delta(z - kz_a), \tag{6.18c}$$

and

$$\frac{d\Phi}{dz} = \frac{A^2 - \Omega^2}{2} + g_\Omega(A,\Omega)T + \sum_{k=1}^{N_a} \Delta\Phi_k\, \delta(z - kz_a), \tag{6.18d}$$

where the stochastic jump terms are given by evaluating the projection functions

$$\begin{aligned}
\Delta A_k &= \sigma \mathrm{Re}\left[\int \bar{v}_A^\dagger(kz_a,t)\exp(-i\Theta(kz_a,t))\, n_k(t)\, dt\right],\\
\Delta \Omega_k &= \sigma \mathrm{Re}\left[\int \bar{v}_\Omega^\dagger(kz_a,t)\exp(-i\Theta(kz_a,t))\, n_k(t)\, dt\right],\\
\Delta T_k &= \sigma \mathrm{Re}\left[\int \bar{v}_T^\dagger(kz_a,t)\exp(-i\Theta(kz_a,t))\, n_k(t)\, dt\right],\\
\Delta \Phi_k &= \sigma \mathrm{Re}\left[\int \left(\bar{v}_\Phi^\dagger(kz_a,t) - T(kz_a)\bar{v}_\Omega^\dagger(kz_a,t)\right)\exp(-i\Theta(kz_a,t))\, n_k(t)\, dt\right],
\end{aligned} \tag{6.19}$$

at the amplification points kz_a. Note that because the addition of the new perturbative terms in no way effects the form of the linearized modes, these stochastic jump terms are equivalent to those in Chapter 2. In particular, they are mean zero and have variances at the $k+1$ amplifier given by

$$\begin{aligned}
\mathbb{E}\left[(\Delta A_{k+1})^2\right] &= \frac{\sigma^2}{2}\|v_A^\dagger(kz_a,t)\|^2 = \sigma^2 A_k,\\
\mathbb{E}\left[(\Delta T_{k+1})^2\right] &= \frac{\sigma^2}{2}\|v_T^\dagger(kz_a,t)\|^2 = \frac{\sigma^2\pi^2}{12A_k^3},\\
\mathbb{E}\left[(\Delta \Omega_{k+1})^2\right] &= \frac{\sigma^2}{2}\|v_\Omega^\dagger(kz_a,t)\|^2 = \frac{\sigma^2 A_k}{3},\\
\mathbb{E}\left[(\Delta \Phi_{k+1})^2\right] &= \frac{\sigma^2}{2}\|v_\Phi^\dagger(kz_a,t) - T_k v_\Omega^\dagger(kz_a,t)\|^2\\
&= \sigma^2 \frac{12\left(1 + (T_k + \Omega_k z_a)^2 A_k^2\right) + \pi^2}{36A_k},
\end{aligned} \tag{6.20}$$

where the notation $X_k = X(kz_a)$ was used for brevity. Figure 6.4 shows a comparison between the parameter evolution in the PNLSE of Equation (6.11) and the ODEs in Equations (6.18) in the absence of noise, demonstrating that this reduced system provides an excellent approximation for the deterministic pulse parameter evolution.

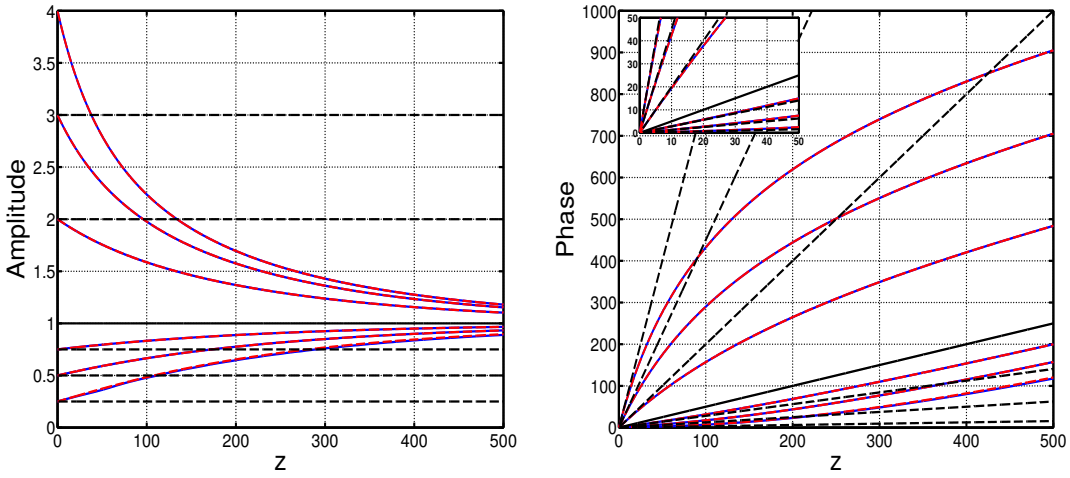


Figure 6.4 Comparison of the amplitude (Left) and phase (Right) parameters between noiseless numerical simulation of the PNLSE given in (6.11) and the system of ODEs given in (6.17). The blue and red curves represent the PDE and ODE results, respectively, whereas the dashed black curves represent the optimal paths under the original conservative SPT equations for NSLE and the solid black curves represent the fixed point, which is chosen to correspond to $A = 1$ by setting the parameter values to be $g = 0.022$, $\tau = 0.01$ and $l = 0.01$ in all simulations.

6.4.2 ISMC for the PNLSE in the Soliton Regime

The reduced system in Equations (6.18) can now be used to derive biasing distributions for the implementation of the ISMC method. This is only done for the phase parameter here, however, since this is the parameter of focus in previous chapters.

Recall that the task of constructing the biasing distributions for ISMC method involves solving two related problems. The first problem is to find the vectors (or functions in the continuous limit) that are most likely to impart unit changes through addition to the pulse. For the NLSE in Chapter 2, these were found to be normalized versions of the adjoint generalized eigenfunctions from SPT, i.e.,

$$f_X(z,t) = \frac{\left(v_X^\dagger(z,t) - Tv_\Omega^\dagger(z,t)\delta_{X\Phi}\right)}{\|v_X^\dagger(z,t) - Tv_\Omega^\dagger(z,t)\delta_{X\Phi}\|^2}\exp(i\Theta)\,, \tag{6.21}$$

for $X = A$, Ω, T and Φ. Since the new terms included in the reduction for PNLSE have *no effect* on the form of these linear modes, the same functions can be used in

this case as well. Thus, it only remains to determine the biasing weights $\Delta_X(z)$ that are used to construct the biasing vector

$$f(z,t) = \Delta_A(z)f_A(z,t) + \Delta_\Omega(z)f_\Omega(z,t) + \Delta_T(z)f_T(z,t) + \Delta_\Phi(z)f_\Phi(z,t), \tag{6.22}$$

that is added to the noise to drive the parameters along the *optimal* path to the targeted final values. Just as in the Chapter 2, the likelihood of realizing this biasing vector at any individual fixed point z is given by a Gaussian PDF and thus, the most likely biasing vector is the one with smallest *cumulative* L^2-norm given by,

$$\begin{aligned} S = \int_0^{z_f} \|f(z,t)\|^2 \, dz = \int_0^{z_f} \sum_Y \left(\Delta_Y(z)\right)^2 \|f_Y(z,t)\|^2 \\ + 2\Delta_\Phi(z)\Delta_\Omega(z) \left\langle f_\Phi(z,t), f_\Omega(z,t) \right\rangle \, dz, \end{aligned} \tag{6.23}$$

where $\sum_Y$ represents the sum over all four parameters. The functional in Equation (6.23) must be minimized subject to the constraint $X(z_f) = X_f$, where X represents the parameter of interest and X_f is the targeted final parameter value. As before, the four weighting functions $\Delta_Y(z)$ can be related to the optimal path through parameter space by finding the mean evolution of Equations (6.17) after the biasing has been applied. By replacing the mean zero noise $\sigma n(z,t)$ with the biased noise $\sigma n(z,t) + \sum_Y \Delta_Y(z) f_Y(z,t)$ and taking the expectation of these equations (noting that $\mathbb{E}[n(z,t)] = 0$), we arrive at equations that relate the biasing weights to the optimal path taken through parameter space, i.e.,

$$\dot{A}_{opt} = g_A(A_{opt}, \Omega_{opt}) + \Delta_A(z), \tag{6.24a}$$

$$\dot{\Omega}_{opt} = -g_\Omega(A_{opt}, \Omega_{opt}) + \Delta_\Omega(z) + \Delta_\Omega(z)\Delta_\Phi(z)\frac{T_{opt}A_{opt}^2}{1 + \frac{\pi^2}{12} + T_{opt}^2 A_{opt}^2}, \tag{6.24b}$$

$$\dot{T}_{opt} = \Omega_{opt} + \Delta_T(z), \tag{6.24c}$$

$$\dot{\Phi}_{opt} = \frac{A_{opt}^2 - \Omega_{opt}^2}{2} + g_\Omega(A_{opt}, \Omega_{opt})T_{opt} + \Delta_\Phi(z), \tag{6.24d}$$

where $\dot{X} = dX/dz$. Using these equations, the functional S given in Equation (2.56) is seen to depend on the optimal path each parameter takes through parameter space.

6.5 Results and Discussion

Just as in all previous calculations for the biasing path associated with the phase parameter, it is assumes here that the changes in both timing and frequency parameters have a negligible effect on the evolution of the phase parameter. This leaves only the amplitude and phase parameters in the construction of the cumulative L^2-norm, which takes the form

$$\begin{aligned} S_\Phi(A_{opt}, \dot{A}_{opt}, \dot{\Phi}_{opt}) &= \int_0^{z_f} \Delta_A(z)\|f_A(z,t)\|^2 + \Delta_\Phi(z)\|f_\Phi(z,t)\|^2\, dz \\ &= \int_0^{z_f} \frac{\left(\dot{A}_{opt} - g_A(A_{opt}, 0)\right)^2}{2A_{opt}} + \frac{18}{12+\pi^2} A_{opt}\left(\dot{\Phi}_{opt} - \frac{A_{opt}^2}{2}\right)^2 dz, \end{aligned} \tag{6.25}$$

under the constraints $A_{opt}(0) = A_i$, $\Phi_{opt}(0) = \Phi_i$ and $\Phi_{opt}(z_f) = \Phi_f$, where is it noted that $A_{opt}(z_f)$ is free to vary and will therefore be determined through the minimization process.

Taking variations of S_Φ gives the following two-dimensional BVP,

$$\ddot{A}_{opt} = \frac{\dot{A}_{opt}^2}{2A_{opt}} + h(A_{opt})\left(\frac{\partial h(A_{opt})}{\partial A_{opt}} - \frac{h(A_{opt})}{2A_{opt}}\right) + \frac{12+\pi^2}{72} c_\Phi^2 \frac{1}{A_{opt}} - c_\Phi A_{opt}^2, \tag{6.26a}$$

$$\dot{\Phi}_{opt} = \frac{A_{opt}^2}{2} + c_\Phi \frac{12+\pi^2}{36}\frac{1}{A_{opt}}, \tag{6.26b}$$

where c_Φ is a constant of integration and

$$\begin{aligned} h(A_{opt}) &= g_A(A_{opt}, 0) \\ &= 2gE_{sat}\frac{A(z_f)}{E_{sat}+2A_{opt}} - 2l\frac{P_{sat}}{(P_{sat}+A_{opt}^2)^{\frac{1}{2}}}\operatorname{arctanh}\left(\frac{A_{opt}}{(P_{sat}+A_{opt}^2)^{\frac{1}{2}}}\right) \\ &\quad - \frac{2}{3}\tau E_{sat}\frac{A_{opt}^4}{E_{sat}+2A_{opt}}. \end{aligned} \tag{6.26c}$$

The boundary conditions for this system are given by $A_{opt}(0) = A_i$, $\Phi_{opt}(0) = \Phi_i$, $\Phi_{opt}(z_f) = \Phi_f$ and $\dot{A}_{opt}(z_f) = h(A_{opt}(z_f))$. By integrating Equation (6.26b) and using the boundary conditions on Φ_{opt}, the integration constant c_Φ can be written in terms of integrals involving A_{opt}, i.e.,

$$c_\Phi = \frac{\Phi_f - \frac{1}{2}\int_0^z A_{opt}^2(z')\,dz'}{\frac{12+\pi^2}{36}\int_0^z \frac{1}{A_{opt}(z')}\,dz'}, \tag{6.27}$$

which gives a integro-differential equation for the optimal path for the amplitude parameter. Just as before, this equation can be solved by a modified shooting method, the form of which is more complex due to the boundary condition on $\dot{A}_{opt}(z_f)$ and the additional term in Equation (6.26a) that contains the function $h(A_{opt})$. These solutions are plotted in Figure 6.5 which, for comparison, also contains the optimal paths for the case of the NLSE in Chapter 2. In addition, the biasing weight functions $\Delta_A(z)$ and $\Delta_\Phi(z)$ are given in Figure 6.6 for both the PNLSE above and the NLSE.

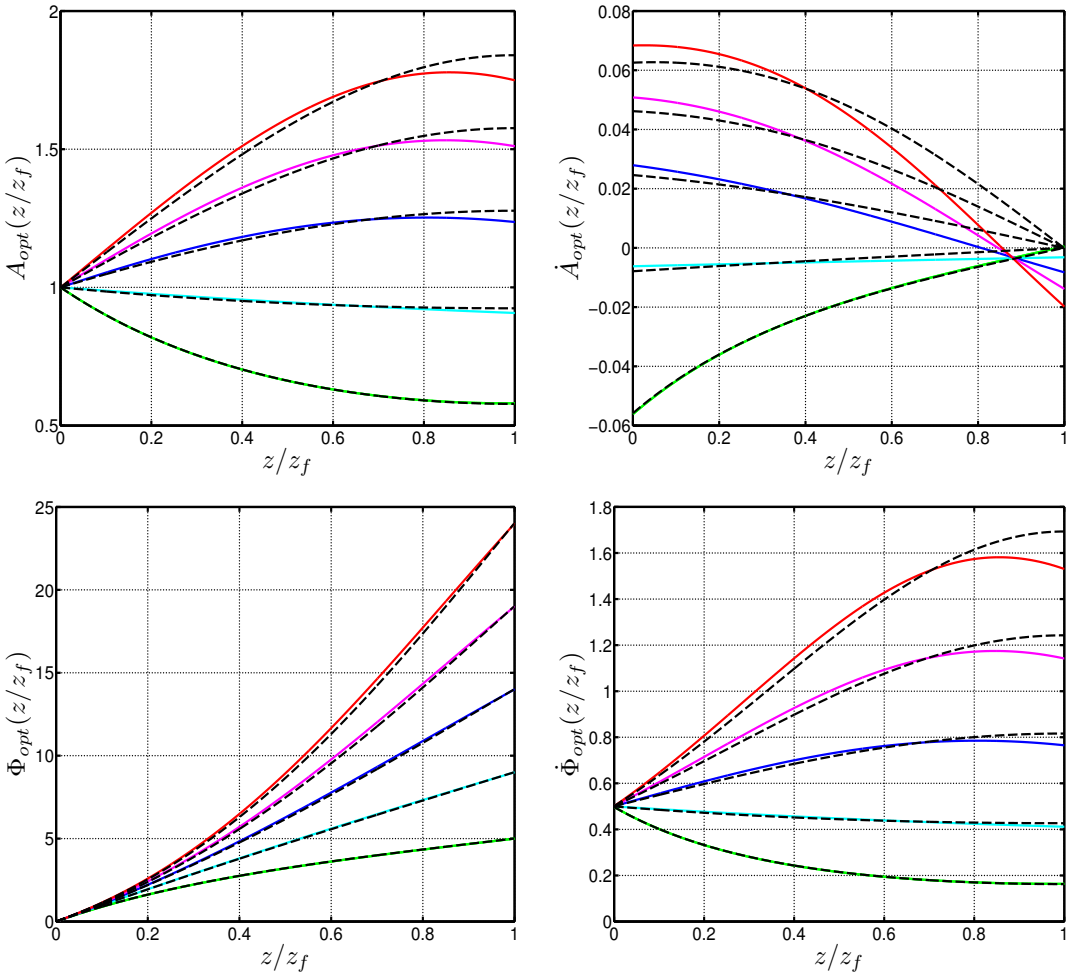


Figure 6.5 The optimal parameter paths for the amplitude (Top Left) and phase (Bottom Left) parameters are given for targeted final phase values of 5.0, 9.0, 14.0, 19.0, 24.0. The derivatives of these paths are given in the plots to the right. For comparison, the analogous parameter paths for the NLSE are given by the dashed black curves. In these calculations, the fixed point was chosen to correspond to $A = 1$ by setting the parameter values to be $g = 0.022$, $\tau = 0.01$ and $l = 0.01$ in all simulations. In the absence of biasing $\Phi_f = 10$.

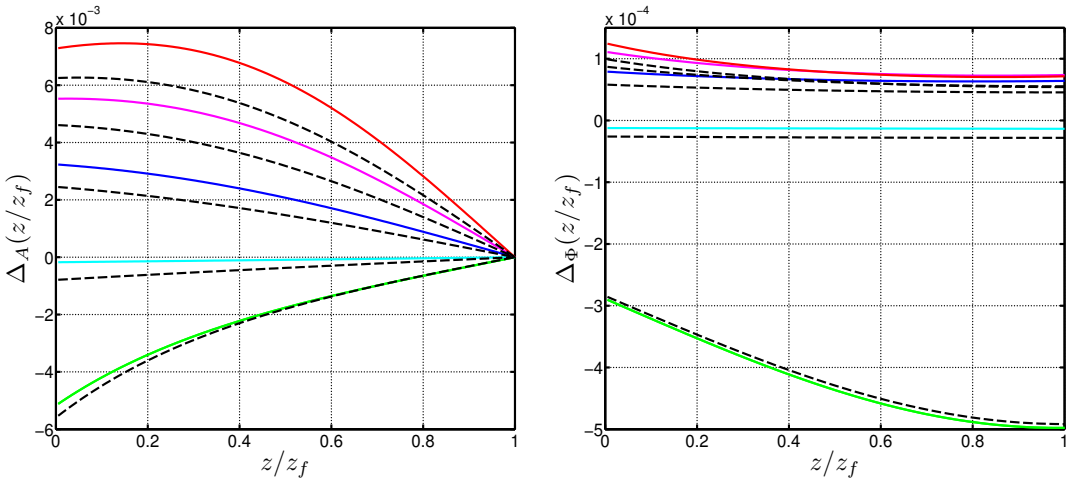


Figure 6.6 The biasing functions for the amplitude (Left) and phase (Right) parameters are given for targeted final phase values of 5.0, 9.0, 14.0, 19.0, 24.0. For comparison, the corresponding biasing functions for the NLSE are given by the dashed black curves. In these calculations, the fixed point was chosen to correspond to $A = 1$ by setting the parameter values to be $g = 0.022$, $\tau = 0.01$ and $l = 0.01$ in all simulations. In the absence of biasing $\Phi_f = 10$.

Now that the biasing paths for the amplitude and phase are known, an ISMC method is easily constructed in a manner analogous to what was presented in Chapter 2 for the NLSE. Figure 6.8 plots the PDFs for the phase using 128, 256, 512 and 1024 simulation modes and a PDF from the ISMC method applied to the NLSE in Chapter 2, which has been shifted to eliminate the radiation-induced phase shift. This plot confirms two expectations, (i) for the same noise strength, large phase deviations in the PNLSE are significantly less probable than in the NLSE due to the restoring forces which constantly work to restore the pulses amplitude to it fixed point value, and (ii) although the radiation induced phase shift is still present in the PNLSE, its effects are greatly reduced due to the smoothing of the noise from the filter in the gain model. To illustrate the effect a filter has on the propagation of a noisy soliton, Figure 6.7 plots the evolution of an initial noise soliton in both the NLSE and PNLSE. Finally, Figure 6.8 plots the PDFs obtained from the ISMC method

against the results of a standard MC method applied to both the SODEs in (6.19) and the PNLSE in Equation (6.11).

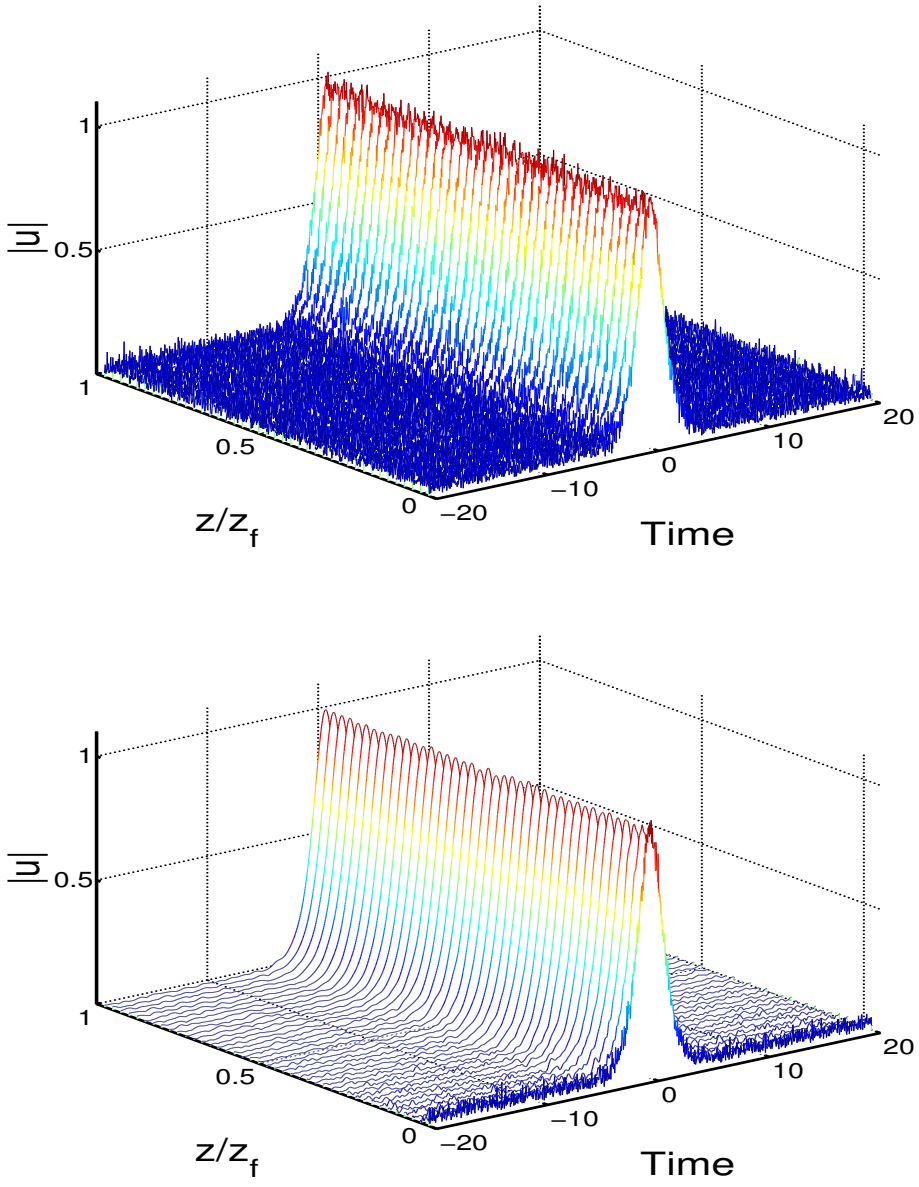


Figure 6.7 Comparison between noisy evolution of a soliton in the NLSE (Top) and the PNLSE (Bottom), which illustrates the smoothing effects of the filtering terms in the PNLSE.

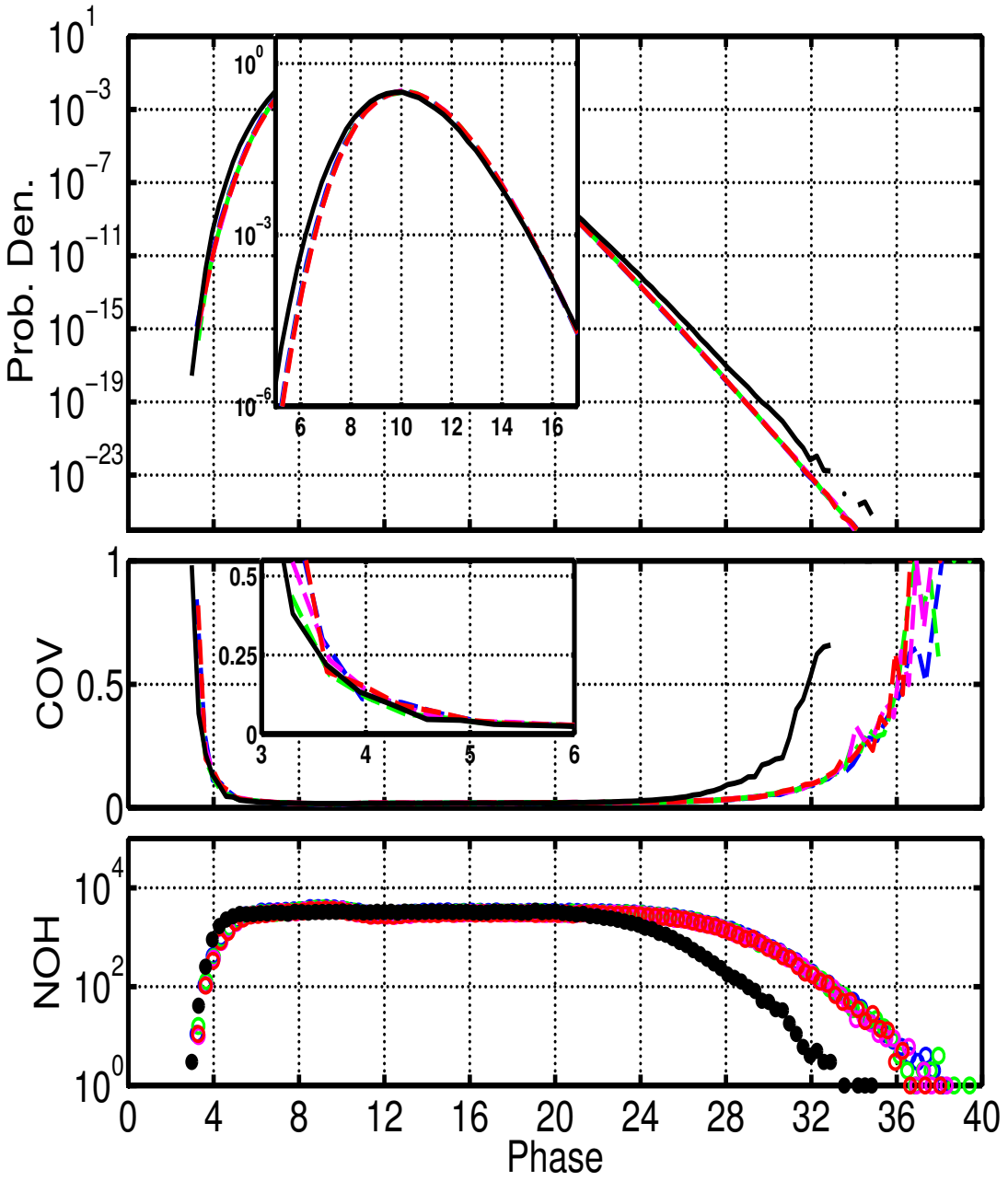


Figure 6.8 Top: Comparison of PDFs for the phase parameter while varying the number of simulation modes. The blue, green, magenta and red curves are the PDFs generated with 2×10^5 runs of the ISMC method applied to the PNLSE using 128, 256, 512 and 1024 simulation modes, respectively. For comparison, the solid black curve represents the results of 2×10^5 runs of the ISMC method applied to the NLSE, where the effects of the radiation drift have been removed. All curves in this plot are on a Log scale of base 10. Middle: Each colored curve represents the COV for the corresponding PDF in the top figure. Bottom: Each sequence of colored markers correspond to the NOH each bin received under the ISMC runs that produced to the PDF plots in the top figure. All curves in this plot are on a Log scale of base 10.

It is interesting to note the effects the saturable gain terms have on the resulting PDFs for the phase parameter. First, recall that in the NLSE, large phase deviations where primarily achieved through biasing the amplitude parameter, which integrated to changes in the phase. The same is true for the PNSLE, because the conservative part of the parameter dynamics are equivalent to those of the NSLE. However, with the addition of the loss terms, changes in the amplitude parameter are now countered by the restoring forces of the gain and loss. Thus, one would intuitively expect that larger biasing strengths are required to achieve the same final phase value. This expectation is partially validated by the plots in Figure 6.6, which show that the biasing curves for targeted phase values above the mean phase are larger than those required for the NLSE, in both the amplitude and phase parameters. However, this is not true for the targeted final phase values of 5.0 and 9.0, both of which are below the mean phase value of 10.0.

For phase value of 5.0 (green curves in Figures 6.6 and 6.5), both the amplitude and phase biasing curves in the case of the PNLSE are seen to remain very close to the biasing curves of the NLSE. Although this is result is surprising, it does not contradict intuition, since the reduced (in absolute value) biasing curve for the amplitude is compensated for by the increased (in absolute value) phase biasing curve. The results for the phase value of 9.0 (cyan curves in Figures 6.6 and 6.5), are even more surprising in that both the amplitude and phase biasing curves in the case of the PNLSE are seen to *decrease* (in absolute value), which implies that it is *more* probable to get a final phase value of 9.0 in the PNLSE than in the NLSE. To explain this, consider the result in Figure 6.8, which show that the PDFs of the final phase value under PNLSE fall inside of the PDF curve for NSLE, implying that large deviations in the PNLSE are less probable than in the NLSE or equivalently that the PDFs for the PNLSE have smaller variances than those of the NSLE. However, a subtle consequence of this is that the PDF around the mean must increase to ensure that the total probability

remains one, which confirms what was implied by the cyan curves in Figures 6.6, i.e., final phase values around the mean are more probable in PNLSE than in NLSE.

Finally, it is important to note that the effects of dispersive radiation are greatly diminished in the case of the PNLSE, as shown in the PDFs of Figure 6.8. As stated before, this can be attributed to the effect of filtering in the gain term, which damps many of the high frequency radiation modes and reduces their contribution to the phase shift effect seen in the NLSE. Therefore, it is expected that any improvements in the reduced system through the inclusion of radiation will not yield better performance of the ISMC method for the case of PNLSE or related equations that include filtering effects.

6.6 Summary

This chapter began by introducing the concept of mode-locking in both the temporal and spectral domains. In the temporal domain, the output of a MLL is short highly energetic pulses of light, whereas in the in the spectral domain, the output is a frequency comb of sharply spiked spectral lines. The chapter continued by introducing several models for the pulse propagation, including the PNSLE, which was the focus for the remainder of the chapter. Using the PNLSE, and assuming a laser operating in the soliton regime, SPT was used to derive a reduces system for the pulse dynamics, that was numerically shown to compare extremely well to simulations of the PNLSE. Finally, this reduced system was used to construct the biasing distributions necessary for the implementation of the ISMC method, the results of which were compared to the results of ISMC method applied to the NLSE in Chapter 2.

CHAPTER 7

SUMMARY AND FUTURE WORK

Summary. This dissertation has presented the results of applying the importance sampling Monte-Carlo (ISMC) method to various mathematical models associated with both optical communications and mode-locked lasers. Chapter 2 presented the ISMC method applied to a stochastic version of the nonlinear Schrödinger equation (NLSE) as a mathematical propagation model for noisy optical fiber. This included a discussion of the special solutions to the NLSE, known as solitons, and the perturbative technique known as soliton perturbation theory (SPT), which yields a low-dimensional approximation used to guide the ISMC method. The results of this chapter were used to illustrate that the reduction resulting from first order SPT lacks the capability of accounting for the effects of dispersive radiation, which can contribute significantly to the evolution of the phase parameter.

The inadequacy of SPT to correctly account for the evolution of the phase parameter was addressed in Chapter 3 through the construction of an improved low-dimensional reduction based on a variational reformulation of the stochastic NLSE. Under this approach, the perturbed soliton solutions of the stochastic NLSE are seen to correspond to stationary points of a functional. By assuming that the stochastic dynamics of the pulse stays close to that of a soliton and that the radiation can be represented as a windowed Fourier decomposition, stochastic ODEs are derived for the soliton parameters and the Fourier coefficients of the radiation, which are coupled through a term representing the radiation's power that appears in the evolutionary ODE for the soliton's phase. This reduced system was numerically verified through comparison with numerical simulations of the stochastic NLSE, which compares

both the fully nonlinear system of coupled stochastic ODEs (SODEs) and analytical solutions of the linear system resulting from a linearization around a soliton solution.

In Chapter 4, the improved low-dimensional approximation derived in Chapter 3 was used to construct an improved ISMC method that accounts for the interaction between the soliton and radiation. This method differs from the original by incorporating changes in both the soliton parameters and the power in the radiation, resulting in a more effective biasing scheme and a more efficient application of the ISMC method.

The effects of radiation seen in Chapter 2, and subsequently investigated in Chapter 3, also appear in more realistic propagation models for optical communication systems. Of particular interest was the case of the NLSE with a periodic modulation of the dispersion constant, referred to as dispersion management (DM), which yields a varying coefficient version of NLSE (NLSE+DM), and a related model where this modulation is averaged to give an autonomous, nonlocal equation (DMNLSE). The variational approach used in Chapter 3 was extended to the case of the DMNLSE in Chapter 5, with the aim of deriving an improved low-dimensional approximation for these systems that correctly accounts for the effects of radiation. Although this resulted in an improved low-dimensional reduction for the evolution of the DM soliton in the presence of radiation, the use of this reduction in the construction of a more efficient ISMC method remains limited by the lack of closed-form solutions of the DMNLSE.

In Chapter 6, the ISMC method was extended to the calculation of phase-slip probabilities in mode-locked lasers (MLL), which unlike the conservative systems considered in previous chapters, includes dissipative effects such as gain and loss, and non-local effects such as energy saturation. In their most general form, such models exhibit complicated pulse dynamics which cannot be accurately captured by a reduced system. Because of this, the investigation here was focused on a MLL operating in the

soliton propagation regime modeled by the constant coefficient perturbed nonlinear Schrödinger equation (PNLSE), which allows for a perturbative treatment of gain, loss and saturation effects. A reduced system for the PNLSE was derived using SPT, which is subsequently utilized to find the biasing paths needed by the ISMC method. From this, the PDF of the phase parameter was constructed.

Future Work. Other approaches have been considered for finding the biasing distributions need by the ISMC method. One of the more recent is based on a purely numerical implementation, which is capable of handing systems exhibiting complicated dynamics [87]. Although these methods are more general, they require significantly more computational resources and provide less insight into the transitions that lead to the detection of an error. The next step in this investigation is an alternative to both of these approaches, which is based on the extension of Wentzell-Freidlin theory to the stochastic PDE models. In this approach, the problem of finding an optimal path through sample space is transformed into a nonlinear optimization problem over a constrained spatiotemporal functional, which is the infinite-dimensional analog of the constrained minimization problem involving the cumulative L^2-norm of the biasing vector and the equations of the low-dimensional reduction. One of the primary objectives of this work will be to develop a numerical method capable of solving the constrained optimization problem when the endpoint constraints are not critical points of the functional, as is typically the case, but points in functional space that correspond to failure states of the light-wave system under consideration.

APPENDIX A

EVOLUTION EQUATIONS FOR OPTICAL FIBER

A.1 The Derivation of the NLSE

The derivation of NLSE can be found in many text on nonlinear optics. The derivation presented here is based on those found in references [33] and [1].

The propagation of electromagnetic fields in optical fiber, like all other optical material, is fundamentally governed by Maxwell's equations,

$$\nabla \cdot \boldsymbol{D} = 0, \tag{A.1a}$$

$$\nabla \cdot \boldsymbol{B} = 0, \tag{A.1b}$$

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}, \tag{A.1c}$$

$$\nabla \times \boldsymbol{H} = \frac{\partial \boldsymbol{D}}{\partial t}, \tag{A.1d}$$

along with the constitutive relations

$$\boldsymbol{D} = \epsilon_0 \left(\boldsymbol{E} + \boldsymbol{P}\right) \text{ and } \boldsymbol{B} = \mu_0 \left(\boldsymbol{H} + \boldsymbol{M}\right), \tag{A.1e}$$

where it is assumed that the material is void of free charge and surface currents. The external electric and magnetic fields are represented by the vectors $\boldsymbol{E}$ and $\boldsymbol{H}$, respectively, whereas the vectors $\boldsymbol{P}$ and $\boldsymbol{M}$ represent the electric and magnetic polarizations, respectively, which accounts for the material's response to the external fields. The constitutive relations in (A.1e) relate the field vectors with their corresponding polarization through the flux densities $\boldsymbol{D}$ and $\boldsymbol{B}$, where ϵ_0 and μ_0 represent the free

space permittivity and permeability, respectively, which are related to the speed of light in vacuum, denoted c, through the relation $1/c^2 = \epsilon_0\mu_0$.

A.1.1 Maxwell's Equations in Optical Fiber

Any specific application of Maxwell's equations requires knowledge of both the material composition (determining susceptibility of the material to both electric and magnetic polarization) and geometry of optical cavity (determining the boundary conditions that the fields must satisfy). Typical optical fiber consist two concentric cylinders, an inner core and an outer cladding. Both the core and cladding are composed of nearly pure silica glass with slightly different electrical susceptibilities, leading to slightly different indices of refraction. This is an important distinction since it is responsible for the phenomena of total internal reflection in which the electromagnetic (EM) field is primarily confined to the core, allowing EM waves to be transmitted without significant radiation from the surface of the fiber. The difference in electrical susceptibilities between the core and cladding are accounted for mathematically by taking the electrical susceptibility to consist of a piecewise constant function that takes on different values in these two regions. Using this approach, Maxwell's equations can be applied in both the core and cladding independently using a general electrical susceptibility, leading to two independent solutions that are coupled through the boundary condition at the interface. In addition, the magnetic susceptibility of silica glass in negligible, resulting in the absence of a magnetic polarization $\boldsymbol{M}$.

Since the electric polarization vector $\boldsymbol{P}$ is the material's response to an applied electric field, it is reasonable to assume a representation in the form of a power series expansion in powers of $\boldsymbol{E}$, where each term is related to a component of the

polarization vector through a susceptibility tensor, i.e.,

$$\begin{aligned} P_j = & \int_{-\infty}^{\infty} \chi_{jk}^{(1)}(t-t_1)E_k(t_1)dt_1 \\ & + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \chi_{jkl}^{(2)}(t-t_1)E_k(t_1)E_l(t_2)dt_1dt_2 \\ & + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \chi_{jklm}^{(3)}(t-t_1,t-t_2,t-t_3)E_k(t_1)E_l(t_2)E_m(t_3)dt_1dt_2dt_3 \\ & + \ldots, \end{aligned} \tag{A.2}$$

where summation is assumed over like indices. Note that since the equations in (A.1) are linear and the magnetic polarization $\boldsymbol{M}$ is assumed to be zero, the only source of nonlinearity comes from the polarization defined in equation (A.2). The representation of the electric polarization vector as convolution integrals in time is necessary to correctly model the non-instantaneous nature of the material's response to an external electric field. In addition, causality implies that the an induced electric polarization cannot depend on future values of the electric field, thus the susceptibility tensors are assumed to be zero for all negative arguments.

The polarization above is written in a very general form, which is much too complex to be used in the derivation of the NLSE. Fortunately, optical fiber posses two properties that significantly simplify this representation [33]. The first being the centrosymmetric nature of the fiber, which implies that the sign of the electric polarization must follow that of the electric field, i.e., if $\boldsymbol{E} \to -\boldsymbol{E}$, then it is required that $\boldsymbol{P} \to -\boldsymbol{P}$. An immediate consequence of this requirement is that all polarizations having an even power of $\boldsymbol{E}$ are identically zero, and thus, the first nonlinear term encountered is cubic in the electric field. The second is that the fiber is isotropic, which implies that the material's electrical susceptibility is essentially independent of the polarization of the external electric field, which reduces the number of independent susceptibility terms in the cubic nonlinearity. Applying these simplifications, the

polarization takes the form

$$\begin{aligned}\boldsymbol{P} &= \int_{-\infty}^{\infty} \chi^{(1)}(t-t_1)\boldsymbol{E}(t_1)dt_1 \\ &+ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \chi^{(3)}(t-t_1,t-t_2,t-t_3)\left(\boldsymbol{E}(t_1)\cdot\boldsymbol{E}(t_2)\right)\boldsymbol{E}(t_3)dt_1dt_2dt_3,\end{aligned} \tag{A.3}$$

where the expansion is truncated after the cubic term. Finally, note that $\chi^{(1)}(t-t_1)$ is a complex valued quantity, with an imaginary part that produces attenuation (loss) in the fiber. However, due to its small relative strength, this term is included as a perturbation, i.e., $\chi^{(1)}(t-t_1) = \chi_R^{(1)}(t-t_1) + i\varepsilon^2\chi_I^{(1)}(t-t_1)$, where ε is a small parameter which will be defined later. In addition, the third order susceptibility is also complex, however the imaginary component is typically too small to contribute at the orders considered in this derivation.

By taking the curl of Equations (A.1c) and (A.1d), and using the constitutive relation in Equation (A.1e), Maxwell's equation are decoupled into

$$\nabla^2\boldsymbol{E} - \nabla\left(\nabla\cdot\boldsymbol{E}\right) - \frac{1}{c^2}\left(\frac{\partial^2\boldsymbol{E}}{\partial t^2} + \frac{\partial^2\boldsymbol{P}_{\mathrm{L}}}{\partial t^2}\right) = \frac{1}{c^2}\frac{\partial^2\boldsymbol{P}_{\mathrm{NL}}}{\partial t^2}, \tag{A.4a}$$

and

$$\nabla\cdot\boldsymbol{D} = 0 \tag{A.4b}$$

for the electric field, and

$$\nabla^2\boldsymbol{H} - \frac{1}{c^2}\frac{\partial^2\boldsymbol{H}}{\partial t^2} = -\frac{1}{c^2}\left(\nabla\times\frac{\partial\boldsymbol{P}}{\partial t}\right), \tag{A.5a}$$

and

$$\nabla\cdot\boldsymbol{H} = 0 \tag{A.5b}$$

for the magnetic field, where $\boldsymbol{M} = \boldsymbol{0}$ was used to eliminate the divergence term in the magnetic field equation. Note that the electric polarization vector $\boldsymbol{P}$ is separated

into linear and nonlinear parts as $\boldsymbol{P} = \boldsymbol{P}_{\mathrm{L}} + \boldsymbol{P}_{\mathrm{NL}}$ with,

$$\boldsymbol{P}_{\mathrm{L}} = \int_{-\infty}^{\infty} \chi^{(1)}(t - t_1)\boldsymbol{E}(t_1)dt_1 \tag{A.6a}$$

and

$$\begin{aligned}\boldsymbol{P}_{\mathrm{NL}} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \chi^{(3)}(t - t_1, t - t_2, t - t_3)\times \\ (\boldsymbol{E}(t_1)\cdot\boldsymbol{E}(t_2))\,\boldsymbol{E}(t_3)dt_1dt_2dt_3.\end{aligned} \tag{A.6b}$$

Note also that the nonlinear polarization the electric flux density $\boldsymbol{D} = \boldsymbol{E} + \boldsymbol{P}_{\mathrm{L}} + \boldsymbol{P}_{\mathrm{NL}}$ is *not* linearly proportional to the electric field $\boldsymbol{E}$, and thus, the term $\nabla(\nabla\cdot\boldsymbol{E})$ must remain in equation (A.4a).

A.1.2 Approximation for the Spatial Mode

Before beginning the calculation of the NLSE, it is illustrative to derive an approximation for the first order solution of the expansion, which gives the spatial mode profile for the fiber. By neglecting the nonlinear terms, the Equations in (A.4) are written as

$$\nabla^2\boldsymbol{E} - \nabla(\nabla\cdot\boldsymbol{E}) - \frac{1}{c^2}\left(\frac{\partial^2\boldsymbol{E}}{\partial t^2} + \frac{\partial^2\boldsymbol{P}_{\mathrm{L}}}{\partial t^2}\right) = 0, \tag{A.7a}$$

and

$$\nabla\cdot(\boldsymbol{E} + \boldsymbol{P}_{\mathrm{L}}) = 0, \tag{A.7b}$$

which represent the linear evolution of the electric field through a medium with constant electric susceptibility. Because this equation is linear, the time derivatives can be removed by taking the FT, or equivalently inserting a trial solution of the form

$$\boldsymbol{E} = \hat{\boldsymbol{F}}(\boldsymbol{x}_\perp; k(\omega), \omega)\exp(i(kz - \omega t)) + c.c., \tag{A.8}$$

where *c.c.* represents the complex conjugate of the first term, making the solution a real quantity. Note, the *c.c.* will be assumed in each equation that follows, even when not explicitly written. Inserting (A.8) into the Equations in (A.7a) gives

$$\nabla^2 \hat{\boldsymbol{F}} + n^2(\omega)\frac{\omega^2}{c^2}\hat{\boldsymbol{F}} = \mathbf{0}, \tag{A.9a}$$

and

$$\nabla \cdot \hat{\boldsymbol{F}} = 0, \tag{A.9b}$$

where $n^2(\omega) = 1 + \hat{\chi}^{(1)}(\omega)$ is the index of refraction for the material, i.e., either in the core or the cladding.

At this point, two simplifications can be applied. First, as it commonly known, only two of the six independent terms (three for both the electric and magnetic fields) are independent, so with out loss of generality, only the equation for the z component of $\hat{\boldsymbol{F}}$ will be considered. Second, because of the cylindrical geometry of the fiber, it is advantageous to work in the cylindrical coordinate system, which results in gives the equation

$$\frac{\partial^2 \hat{F}_z}{\partial r^2} + \frac{1}{r}\frac{\partial \hat{F}_z}{\partial r} + \frac{1}{r^2}\frac{\partial^2 \hat{F}_z}{\partial \phi^2} - k^2 \hat{F}_z + \frac{n^2\omega^2}{c^2}\hat{F}_z = 0. \tag{A.10}$$

Using separation of variables, it is assumed that the solution of Equation (A.10) takes the form

$$\hat{F}_z(r, \phi; \omega) = \tilde{A}(\omega) R(r) \exp(im\phi) \tag{A.11}$$

which gives

$$\frac{\partial^2 R}{\partial r^2} + \frac{1}{r}\frac{\partial R}{\partial r} + \left(\frac{n^2\omega^2}{c^2} - k^2 - \frac{m^2}{r^2}\right) R = 0. \tag{A.12}$$

It is immediately apparent that Equation (A.12) is the well-known equation for a Bessel function. Defining r_{i} as the radius of the interface and letting n_{cr} and n_{cl} be the indicies of refraction for the core and cladding, respectfully, the equation in the core ($r < r_{\mathrm{i}}$) has a solution of

$$R(r) = J_m(\kappa_{\mathrm{cr}} r), \tag{A.13a}$$

where J_m is the Bessel function of the first kind and

$$\kappa_{\mathrm{cr}} = \frac{n_{\mathrm{cr}}^2 \omega^2}{c^2} - k^2. \tag{A.13b}$$

In addition, the solution in the cladding ($r > r_{\mathrm{i}}$) takes the form

$$R(r) = K_m(\kappa_{\mathrm{cl}} r), \tag{A.14a}$$

where K_m is the modified Bessel function that decays to zero with increasing argument and

$$\kappa_{\mathrm{cl}} = k^2 - \frac{n_{\mathrm{cl}}^2 \omega^2}{c^2}. \tag{A.14b}$$

The boundary condition at the interface are used to determine the constant k, and thus, the number of spacial modes the fiber can support [88]. Optical fiber based communication systems (almost) exclusively use single-mode fiber, i.e., fiber supporting only a single transverse mode, with typical core and cladding indices around 1.45 ± 0.005 and diameters of 5-10 μm and 60-140μm, respectively [1].

A.1.3 Simplifying Assumptions and the General Dispersion Relation

The NLSE will be derived through a perturbation approach, where the solution is represented by an asymptotic expansion of the electric field in a small parameter associated with the bandwidth of the pulse in the frequency domain. This is equivalent

to the slowly-varying-envelope (SVE) approximation, which assumes that the pulse is approximately constant, i.e., slowly-varying, over a large portion of the transverse (time) domain. For simplicity, the derivation presented here neglects the restrictions imposed by Equation (A.4b) and the boundary conditions at the interface between the core and cladding, leaving only Equation (A.4a). Because the boundary conditions are neglected, this derivation can be carried out in Cartesian coordinates, noting that each equation has an analog in the cylindrical coordinate system, which is more appropriate for derivations that include the interfacial effects between the core and cladding.

The General Dispersion Relation. As stated above, the expansion used in this derivation has a first order term that consists of a small-bandwidth wave packet solution. The modal structure for this term can found by neglecting the nonlinear electric polarization terms in Equation (A.4a) and inserting a general solution of the form

$$\boldsymbol{E} = \hat{\boldsymbol{F}}(\boldsymbol{x}_\perp; k, \omega) \exp(i(kz - \omega t)) \,, \tag{A.15}$$

which, in matrix-vector notation, gives

$$L\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, ik, -i\omega\right)\left[\hat{\boldsymbol{F}}(\boldsymbol{x}_\perp; k, \omega)\right] = \boldsymbol{0}, \tag{A.16a}$$

where

$$L\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, ik, -i\omega\right) = \begin{bmatrix} \frac{\partial^2}{\partial y^2} + (ik)^2 - \frac{n^2(\omega)(i\omega)^2}{c^2} & -\frac{\partial}{\partial x}\frac{\partial}{\partial y} & -ik\frac{\partial}{\partial x} \\ -\frac{\partial}{\partial x}\frac{\partial}{\partial y} & \frac{\partial^2}{\partial x^2} + (ik)^2 - \frac{n^2(\omega)(i\omega)^2}{c^2} & -ik\frac{\partial}{\partial y} \\ -ik\frac{\partial}{\partial x} & -ik\frac{\partial}{\partial y} & \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} - \frac{n^2(\omega)(i\omega)^2}{c^2} \end{bmatrix}, \tag{A.16b}$$

and $n^2(\boldsymbol{x}_\perp, \omega) = (1 + \hat{\chi}_{\mathrm{R}}^{(1)}(\boldsymbol{x}_\perp, \omega))$ with $\boldsymbol{x}_\perp = [x, y]^T$. For any particular value of ω, there exist nontrivial solutions, $\hat{\boldsymbol{F}}(\boldsymbol{x}_\perp; k(\omega), \omega)$, to the above operator, which along with the appropriate boundary conditions determine the dispersion relation $k(\omega) = f(\omega, n(\omega))$. In addition, by defining an inner product as

$$\langle \boldsymbol{X}, \boldsymbol{Y} \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \bar{\boldsymbol{X}} \cdot \boldsymbol{Y} \, dxdy, \tag{A.17}$$

integration-by-parts shows that

$$\langle \boldsymbol{X}, L[\boldsymbol{Y}] \rangle = \langle L[\boldsymbol{X}], \boldsymbol{Y} \rangle, \tag{A.18}$$

and thus, L is self-adjoint and $\langle \boldsymbol{X}, L[\boldsymbol{Y}] \rangle$ is real [1]. From Equation (A.16a), it is seen that

$$\left\langle \hat{\boldsymbol{F}}, L\left[\hat{\boldsymbol{F}}\right] \right\rangle = 0, \tag{A.19}$$

where $\hat{\boldsymbol{F}}$ is the solution to $L\left[\hat{\boldsymbol{F}}\right] = 0$. Since both L and $\hat{\boldsymbol{F}}$ depend parametrically on ω, Equation (A.19) can be differentiated with respect to ω, giving the useful relation

$$k' \left\langle \hat{\boldsymbol{F}}, \frac{\partial L}{\partial (ik)} \left[\hat{\boldsymbol{F}}\right] \right\rangle = \left\langle \hat{\boldsymbol{F}}, \frac{\partial L}{\partial (-i\omega)} \left[\hat{\boldsymbol{F}}\right] \right\rangle. \tag{A.20}$$

A.1.4 The Expansion

It is now assumed that the solution to the linearized problem, $\hat{\boldsymbol{F}}(\boldsymbol{x}_\perp; k(\omega), \omega)$ and therefore the dispersion relation $k(\omega)$, is known, which allows for the introduction of a series expansion for the nonlinear solution of

$$\boldsymbol{E}(z, t) = \boldsymbol{E}_0 + \varepsilon^2 \boldsymbol{E}_1 + \varepsilon^3 \boldsymbol{E}_2 + \ldots, \tag{A.21}$$

where

$$\hat{\boldsymbol{E}}_0 = \hat{\boldsymbol{F}}(\boldsymbol{x}_\perp; \omega) \hat{A}\left(Z_1, Z_2, Z_3, \ldots, \frac{\omega - \omega_0}{\varepsilon} \right) \exp(ik_0 z) + c.c., \tag{A.22}$$

[1] In both statements it is assumed that $\boldsymbol{X}$, $\boldsymbol{Y}$, $\frac{\partial \boldsymbol{X}}{\partial x}$, $\frac{\partial \boldsymbol{Y}}{\partial x}$, $\frac{\partial \boldsymbol{X}}{\partial y}$ and $\frac{\partial \boldsymbol{Y}}{\partial y} \to \boldsymbol{0}$ sufficiently fast as $\boldsymbol{x}_\perp \to \infty$

with $Z_j = \varepsilon^j z$, $T = \varepsilon t$ and $k_0 = k(\omega_0)$. Note that the FT version of the first term in this expansion consist of the linear solution of Equation (A.16a) enveloped by an amplitude coefficient, A, that is localized around a central wave number ω_0, where the localization is quantized by the small parameter ε. Also note that *c.c.* stands for the complex conjugate of the first term, which ensures that the field is real. Taking the inverse FT of Equation (A.22) gives

$$\boldsymbol{E}_0(z,t) = \sum_{n=0}^{\infty} \frac{i^n}{n!} \varepsilon^{n+1} \frac{\partial^n \hat{\boldsymbol{F}}}{\partial \omega^n}\bigg|_{\omega=\omega_0} \frac{\partial^n A}{\partial T^n} \exp(i(k_0 z - \omega_0 t)) + c.c. \tag{A.23}$$

Follow the usual procedure of inserting the expansion in Equation (A.21) into (A.4a) and solving order by order, gives the linear operator L applied to each perturbation $\boldsymbol{E}_j$ which is forced by the terms from $\boldsymbol{E}_0$ that are of the same order. By enforcing the expansion to stay well ordered, one can derive an equation for the amplitude envelope $A(Z_1, Z_2, \ldots, T)$.

A.1.5 The Polarization Terms

Before beginning the process of solving the linear equations at each order, it is advantageous to first indicate the polarization term that results from the assumed expansion in Equation (A.21).

Linear Polarization. Using the FT, the linear polarization term is written as

$$\hat{\boldsymbol{P}}_{\mathrm{L}} = \mathcal{F}\left[\int_{-\infty}^{\infty} \chi^{(1)}(t-t_1)\boldsymbol{E}(t_1)\, dt_1\right] = \left(\hat{\chi}_{\mathrm{R}}^{(1)}(\omega) + i\varepsilon^2 \hat{\chi}_{\mathrm{I}}^{(1)}(\omega)\right) \hat{\boldsymbol{E}}. \tag{A.24}$$

Inserting the FT of the expansion in Equation (A.21) for $\hat{\boldsymbol{E}}$ gives

$$\begin{aligned}
\hat{\boldsymbol{P}}_{\mathrm{L}} = &\sum_{n=0}^{\infty} \frac{1}{n!} \frac{\partial^n \left(\hat{\chi}_{\mathrm{R}}^{(1)}(\omega)\hat{\boldsymbol{F}}(\boldsymbol{x}_\perp;\omega)\right)}{\partial \omega^n}\bigg|_{\omega=\omega_0} (\omega-\omega_0)^n \hat{A}\left(\frac{\omega-\omega_0}{\varepsilon}\right) \exp(ik_0 z) \\
&+ i\varepsilon^2 \sum_{n=0}^{\infty} \frac{1}{n!} \frac{\partial^n \left(\hat{\chi}_{\mathrm{I}}^{(1)}(\omega)\hat{\boldsymbol{F}}(\boldsymbol{x}_\perp;\omega)\right)}{\partial \omega^n}\bigg|_{\omega=\omega_0} (\omega-\omega_0)^n \hat{A}\left(\frac{\omega-\omega_0}{\varepsilon}\right) \exp(ik_0 z) \\
&+ \varepsilon^2 \hat{\chi}_{\mathrm{R}}^{(1)}(\omega)\hat{\boldsymbol{E}}_1 + \varepsilon^3 \hat{\chi}_{\mathrm{R}}^{(1)}(\omega)\hat{\boldsymbol{E}}_2 + c.c. + O\left(\varepsilon^4\right),
\end{aligned} \tag{A.25}$$

which after taking the inverse FT yields

$$\begin{aligned}\boldsymbol{P}_{\mathrm{L}} = &\sum_{n=0}^{\infty} \frac{i^n}{n!} \frac{\partial^n \left(\hat{\chi}_{\mathrm{R}}^{(1)} \hat{\boldsymbol{F}}\right)}{\partial \omega^n}\Big|_{\omega=\omega_0} \varepsilon^{n+1} \frac{\partial^n A}{\partial T^n} \exp(i(k_0 z - \omega_0 t)) \\ &+ i\varepsilon^2 \sum_{n=0}^{\infty} \frac{i^n}{n!} \frac{\partial^n \left(\hat{\chi}_{\mathrm{I}}^{(1)} \hat{\boldsymbol{F}}\right)}{\partial \omega^n}\Big|_{\omega=\omega_0} \varepsilon^{n+1} \frac{\partial^n A}{\partial T^n} \exp(i(k_0 z - \omega_0 t)) \\ &+ \varepsilon^2 \int \chi_{\mathrm{R}}^{(1)}(t - t_1) \boldsymbol{E}_1(t_1)\, dt_1 + \varepsilon^3 \int \chi_{\mathrm{R}}^{(1)}(t - t_1) \boldsymbol{E}_2(t_1)\, dt_1 \\ &+ c.c. + O\big(\varepsilon^4\big)\,.\end{aligned} \tag{A.26}$$

Differentiating this twice with respect to t gives

$$\begin{aligned}\frac{\partial^2 \boldsymbol{P}_{\mathrm{L}}}{\partial t^2} = &\varepsilon \left[\hat{\chi}_{\mathrm{R}}^{(1)}(\omega_0) \hat{\boldsymbol{F}}(\omega_0) (i\omega_0)^2\right]_{\omega=\omega_0} A \exp(i(k_0 z - \omega_0 t)) \\ &+ \varepsilon^2 \left[-2i\omega_0 \left(\hat{\chi}_{\mathrm{R}}^{(1)} \hat{\boldsymbol{F}}\right) + i(i\omega_0)^2 \frac{\partial \left(\hat{\chi}_{\mathrm{R}}^{(1)} \hat{\boldsymbol{F}}\right)}{\partial \omega}\right]_{\omega=\omega_0} \frac{\partial A}{\partial T} \exp(i(k_0 z - \omega_0 t)) \\ &+ \varepsilon^3 \left[\hat{\chi}_{\mathrm{R}}^{(1)} \hat{\boldsymbol{F}} - i2(i\omega_0) \frac{\partial \left(\hat{\chi}_{\mathrm{R}}^{(1)} \hat{\boldsymbol{F}}\right)}{\partial \omega} - \frac{(i\omega_0)^2}{2} \frac{\partial^2 \left(\hat{\chi}_{\mathrm{R}}^{(1)} \hat{\boldsymbol{F}}\right)}{\partial \omega^2}\right]_{\omega=\omega_0} \frac{\partial^2 A}{\partial T^2} \exp(i(k_0 z - \omega_0 t)) \\ &+ \varepsilon^3 \left[i(i\omega_0)^2 \left(\hat{\chi}_{\mathrm{I}}^{(1)} \hat{\boldsymbol{F}}\right)\right]_{\omega=\omega_0} A \exp(i(k_0 z - \omega_0 t)) \\ &+ \varepsilon^2 \frac{\partial^2}{\partial t^2} \int \chi_{\mathrm{R}}^{(1)}(t - t_1) \boldsymbol{E}_1(t_1)\, dt_1 + \varepsilon^3 \frac{\partial^2}{\partial t^2} \int \chi_{\mathrm{R}}^{(1)}(t - t_1) \boldsymbol{E}_2(t_1)\, dt_1 \\ &+ c.c. + O\big(\varepsilon^4\big)\,.\end{aligned} \tag{A.27}$$

Nonlinear Polarization. The calculation of the nonlinear polarization vector is similar to the linear polarization above, however, since only terms up to $O(\varepsilon^3)$ are used for the derivation of the NSLE, only the first term of the expansion in Equation (A.21) is needed. In addition, from the assumption that $\chi^{(3)}$ is real, the FT of this quantity is seen to be even in each argument, which is a used to combine terms.

Recall that the nonlinear polarization is give by

$$\begin{aligned}\boldsymbol{P}_{\mathrm{NL}} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} &\chi^{(3)}(t - t_1, t - t_2, t - t_3) \times \\ &(\boldsymbol{E}_0(t_1) \cdot \boldsymbol{E}_0(t_2))\, \boldsymbol{E}_0(t_3)\, dt_1 dt_2 dt_3,\end{aligned} \tag{A.28}$$

which can be written as

$$\begin{aligned}\boldsymbol{P}_{\mathrm{NL}} = \frac{1}{(2\pi)^3}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\hat{\chi}^{(3)}(\omega_1,\omega_2,\omega_3)\times \\ \left(\hat{\boldsymbol{E}}_0(\omega_1)\cdot\hat{\boldsymbol{E}}_0(\omega_2)\right)\hat{\boldsymbol{E}}_0(\omega_3)\exp(-i(\omega_1+\omega_2+\omega_3)t)\,d\omega_1 d\omega_2 d\omega_3,\end{aligned} \tag{A.29}$$

by using the definition of the inverse FT, i.e.,

$$\boldsymbol{E}_0(t_j) = \frac{1}{2\pi}\int \hat{\boldsymbol{E}}_0(t_j)\exp(-i\omega_j t_j)\;d\omega_j. \tag{A.30}$$

Using the assumed form of $\hat{\boldsymbol{E}}_0$ from Equation (A.22), (A.29) can be written as

$$\begin{aligned}\boldsymbol{P}_{\mathrm{NL}} = \varepsilon^3\hat{\chi}^{(3)}(\omega_0)\left[\left(2(\hat{\bar{\boldsymbol{F}}}\cdot\hat{\boldsymbol{F}})\hat{\boldsymbol{F}}+(\hat{\boldsymbol{F}}\cdot\hat{\boldsymbol{F}})\hat{\bar{\boldsymbol{F}}}\right)|A|^2 A\exp(i(k_0 z-\omega_0 t))\right. \\ \left.+(\hat{\boldsymbol{F}}\cdot\hat{\boldsymbol{F}})\hat{\boldsymbol{F}}A^3\exp(3i(k_0 z-\omega_0 t))\right]+c.c.+O\left(\varepsilon^4\right),\end{aligned} \tag{A.31}$$

which immediately gives

$$\begin{aligned}\frac{\partial^2\boldsymbol{P}_{\mathrm{NL}}}{\partial t^2} = \varepsilon^3\hat{\chi}^{(3)}(\omega_0)(i\omega_0)^2\left[\left(2(\hat{\bar{\boldsymbol{F}}}\cdot\hat{\boldsymbol{F}})\hat{\boldsymbol{F}}+(\hat{\boldsymbol{F}}\cdot\hat{\boldsymbol{F}})\hat{\bar{\boldsymbol{F}}}\right)|A|^2 A\exp(i(k_0 z-\omega_0 t))\right. \\ \left.+9(i\omega_0)^2(\hat{\boldsymbol{F}}\cdot\hat{\boldsymbol{F}})\hat{\boldsymbol{F}}A^3\exp(3i(k_0 z-\omega_0 t))\right]+c.c.+O\left(\varepsilon^4\right).\end{aligned} \tag{A.32}$$

Order ε. The linear equation at $O(\varepsilon)$ is equivalent to Equation (A.16), which is satisfied by the choice of spacial mode $\hat{\boldsymbol{F}}(\boldsymbol{x}_\perp; k(\omega_0),\omega_0)$.

Order ε^2. At $O(\varepsilon^2)$, the expansion gives

$$\nabla^2\boldsymbol{E}_1-\nabla\left(\nabla\cdot\boldsymbol{E}_1\right)-\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\left(\boldsymbol{E}_1+\int_{-\infty}^{\infty}\chi_{\mathrm{R}}^{(1)}(t-t_1)\boldsymbol{E}_1(t_1)\,dt_1\right)=-\boldsymbol{G}_1 \tag{A.33a}$$

where

$$\begin{aligned}\boldsymbol{G}_1 = \left(iL\left[\frac{\partial\hat{\boldsymbol{F}}}{\partial\omega}\right]\frac{\partial A}{\partial T}+\frac{\partial L}{\partial(ik_0)}\left[\hat{\boldsymbol{F}}\right]\frac{\partial A}{\partial Z_1}\right. \\ \left.+\frac{\partial L}{\partial(-i\omega_0)}\left[\hat{\boldsymbol{F}}\right]\frac{\partial A}{\partial T}\right)\exp(i(k_0 z-\omega_0 t)).\end{aligned} \tag{A.33b}$$

Differentiating Equation (A.16a) with respect to ω gives the useful relation

$$L\left[\hat{\boldsymbol{F}}_{\omega}\right] = -ik_0' \frac{\partial L}{\partial(ik_0)}\left[\hat{\boldsymbol{F}}\right] + i\frac{\partial L}{\partial(-i\omega_0)}\left[\hat{\boldsymbol{F}}\right], \tag{A.34}$$

which when substituted into Equation (A.33b) gives

$$\boldsymbol{G}_1 = \frac{\partial L}{\partial(ik_0)}\left[\hat{\boldsymbol{F}}\right]\left(\frac{\partial A}{\partial Z_1} + k_0'\frac{\partial A}{\partial T}\right)\exp(i(k_0 z - \omega_0 t))\,. \tag{A.35}$$

Finally, by letting

$$\boldsymbol{E}_1 = \hat{\boldsymbol{F}}\exp(i(k_0 z - \omega_0 t))\,, \tag{A.36}$$

the orthogonality condition of Fredholm theory [42] gives the restriction of

$$\frac{\partial A}{\partial Z_1} + k_0'\frac{\partial A}{\partial T} = 0. \tag{A.37}$$

Order ε^3 The $O(\varepsilon^3)$ equation gives the first nonlinear term for A,

$$\nabla^2\boldsymbol{E}_2 - \nabla(\nabla\cdot\boldsymbol{E}_2) - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\left(\boldsymbol{E}_2 + \int_{-\infty}^{\infty}\chi_{\mathrm{R}}^{(1)}(t-t_1)\boldsymbol{E}_2(t_1)\,dt_1\right) = -\boldsymbol{G}_2 \tag{A.38a}$$

where

$$\begin{aligned}\boldsymbol{G_2} = \Bigg(&-\frac{1}{2}L\left[\frac{\partial^2\hat{\boldsymbol{F}}}{\partial\omega^2}\right]\frac{\partial^2 A}{\partial T^2} + \frac{\partial L}{\partial(ik_0)}\left[\hat{\boldsymbol{F}}\right]\frac{\partial A}{\partial Z_2} + \frac{(k_0')^2}{2}\frac{\partial L}{\partial(ik_0)}\left[\hat{\boldsymbol{F}}\right]\frac{\partial^2 A}{\partial T^2}\\
&+\frac{1}{2}\frac{\partial L}{\partial(-i\omega_0)}\left[\hat{\boldsymbol{F}}\right]\frac{\partial^2 A}{\partial T^2} - ik_0'\frac{\partial L}{\partial(ik_0)}\left[\frac{\partial\hat{\boldsymbol{F}}}{\partial\omega}\right]\frac{\partial^2 A}{\partial T^2} + i\frac{\partial L}{\partial(-i\omega_0)}\left[\frac{\partial\hat{\boldsymbol{F}}}{\partial\omega}\right]\frac{\partial^2 A}{\partial T^2}\\
&+\frac{i(i\omega_0)^2\hat{\chi}_I^{(1)}}{c^2}\hat{\boldsymbol{F}}A + \frac{(i\omega_0)^2\hat{\chi}^{(3)}}{c^2}\left[2(\hat{\bar{\boldsymbol{F}}}\cdot\hat{\boldsymbol{F}})\hat{\boldsymbol{F}} + (\hat{\boldsymbol{F}}\cdot\hat{\boldsymbol{F}})\hat{\bar{\boldsymbol{F}}}\right]|A|^2A\Bigg)\exp(i(k_0 z - \omega_0 t))\\
&+\frac{9(i\omega_0)^2\hat{\chi}^{(3)}}{c^2}(\hat{\boldsymbol{F}}\cdot\hat{\boldsymbol{F}})\hat{\boldsymbol{F}}A^3\exp(3i(k_0 z - \omega_0 t))\end{aligned} \tag{A.38b}$$

In general $3k(w_0) \neq k(3w_0)$, which implies that the last term does not satisfy the dispersion relation for the operator on the left. Hence, the term will not contribute

to any orthogonality restrictions on A and therefore can be neglected. Just as in the previous case, differentiating Equation (A.16a) twice with respect to ω gives the useful relation

$$\begin{aligned}\frac{1}{2}L\left[\frac{\partial^2\hat{\boldsymbol{F}}}{\partial\omega^2}\right] &= -i\frac{k_0''}{2}\frac{\partial L}{\partial(ik_0)}\left[\hat{\boldsymbol{F}}\right] + \frac{(k_0')^2}{2}\frac{\partial L}{\partial(ik_0)}\left[\hat{\boldsymbol{F}}\right] \\ &\quad + \frac{1}{2}\frac{\partial L}{\partial(-i\omega_0)}\left[\hat{\boldsymbol{F}}\right] - ik_0'\frac{\partial L}{\partial(ik_0)}\left[\frac{\partial\hat{\boldsymbol{F}}}{\partial\omega}\right] + i\frac{\partial L}{\partial(-i\omega_0)}\left[\frac{\partial\hat{\boldsymbol{F}}}{\partial\omega}\right],\end{aligned} \tag{A.39}$$

which gives

$$\begin{aligned}\boldsymbol{G_2} = \Bigg(&\frac{\partial L}{\partial(ik_0)}\left[\hat{\boldsymbol{F}}\right]\frac{\partial A}{\partial Z_2} + i\frac{k_0''}{2}\frac{\partial L}{\partial(ik_0)}\left[\hat{\boldsymbol{F}}\right]\frac{\partial^2 A}{\partial T^2} + \frac{1}{c^2}\left[i(i\omega_0)^2\hat{\chi}_I^{(1)}\right]\hat{\boldsymbol{F}}A \\ &+\frac{1}{c^2}\left[(i\omega_0)^2\hat{\chi}^{(3)}\right]\left[2(\hat{\bar{\boldsymbol{F}}}\cdot\hat{\boldsymbol{F}})\hat{\boldsymbol{F}} + (\hat{\boldsymbol{F}}\cdot\hat{\boldsymbol{F}})\hat{\boldsymbol{F}}^*\right]|A|^2A\Bigg)\exp(i(k_0z - i\omega_0t))\,.\end{aligned} \tag{A.40}$$

At this point, the solvability condition is obtained by projecting $\hat{\boldsymbol{F}}$ onto $\boldsymbol{G_2}$ and setting the result equal to zero, i.e.,

$$\begin{aligned}&\left\langle\hat{\boldsymbol{F}}, \frac{\partial L}{\partial(ik_0)}\left[\hat{\boldsymbol{F}}\right]\right\rangle\frac{\partial A}{\partial Z_2} + i\frac{k_0''}{2}\left\langle\hat{\boldsymbol{F}}, \frac{\partial L}{\partial(ik_0)}\left[\hat{\boldsymbol{F}}\right]\right\rangle A_{TT} - i\frac{\omega_0^2}{c^2}\left\langle\hat{\boldsymbol{F}}, \hat{\chi}_I^{(1)}\hat{\boldsymbol{F}}\right\rangle A \\ &- \frac{\omega_0^2}{c^2}\left\langle\hat{\boldsymbol{F}}, \hat{\chi}^{(3)}\left[2(\hat{\bar{\boldsymbol{F}}}\cdot\hat{\boldsymbol{F}})\hat{\boldsymbol{F}} + (\hat{\boldsymbol{F}}\cdot\hat{\boldsymbol{F}})\hat{\bar{\boldsymbol{F}}}\right]\right\rangle|A|^2A = 0,\end{aligned} \tag{A.41}$$

which can be written as

$$\frac{\partial A}{\partial Z_2} + \frac{ik_0''}{2}\frac{\partial^2 A}{\partial T^2} + \frac{\alpha}{2}A - i\gamma|A|^2A = 0, \tag{A.42a}$$

where the relation in (A.20) was used to write

$$\alpha = 2\omega_0^2k_0'\frac{\left\langle\hat{\boldsymbol{F}}, \hat{\chi}_I^{(1)}\hat{\boldsymbol{F}}\right\rangle}{\left\langle\hat{\boldsymbol{F}}, \frac{\partial(n(\omega_0)\omega_0)^2}{\partial\omega_0}\hat{\boldsymbol{F}}\right\rangle} \tag{A.42b}$$

and

$$\gamma = \omega_0^2k_0'\frac{\left\langle\hat{\boldsymbol{F}}, \hat{\chi}^{(3)}\left[2(\hat{\boldsymbol{F}}^*\cdot\hat{\boldsymbol{F}})\hat{\boldsymbol{F}} + (\hat{\boldsymbol{F}}\cdot\hat{\boldsymbol{F}})\hat{\boldsymbol{F}}^*\right]\right\rangle}{\left\langle\hat{\boldsymbol{F}}, \frac{\partial(n(\omega_0)\omega_0)^2}{\partial\omega_0}\hat{\boldsymbol{F}}\right\rangle}. \tag{A.42c}$$

A.2 The NLSE with Loss and Amplification

The above derivation gives two equation for the evolution of the slowly varying envelope $A(Z_1, Z_2, T)$,

$$\frac{\partial A}{\partial Z_1} + k_0' \frac{\partial A}{\partial T} = 0, \tag{A.43a}$$

and

$$\frac{\partial A}{\partial Z_2} + \frac{ik_0''}{2}\frac{\partial^2 A}{\partial T^2} + \frac{\alpha}{2}A - i\gamma|A|^2 A = 0. \tag{A.43b}$$

The first is solved by making the variable transformation $\tau = T + k_0' Z_1$, which is just a transformation into a reference frame moving at the group velocity in optics coordinates. The remaining equation becomes

$$i\frac{\partial A}{\partial \zeta} + i\frac{\alpha}{2}A - \operatorname{sgn}(k_0'')\frac{|k_0''|}{2}\frac{\partial^2 A}{\partial \tau^2} + \gamma|A|^2 A = 0, \tag{A.44}$$

where the evolution coordinate z_2 was renamed to ζ. This is a dimensional version of the nonlinear Schrödinger equation with a loss term represented by $i\frac{\alpha}{2}A$, where it is noted that α represents the loss in *power*, thus the factor of $1/2$ in the loss term. In addition, $\operatorname{sgn}(k_0'')$ is the sign function applied to k_0'', which represents dispersion in a reference frame moving at the group velocity, i.e., group velocity dispersion (GVD). For solitons to exist, the GVD must counter the effects of nonlinearity, requiring that $\operatorname{sgn}(k_0'') = -1$, which is referred to as the anomalous dispersion case. Although this restriction can be relaxed in the case of dispersion management [89], from this point forward anomalous dispersion is assumed.

A.2.1 Amplification and ASE Noise

Optical amplifiers can be classified into two categories based on the how the amplification is imparted, lumped and distributed. In distributed amplification, an external

power source, i.e., laser, introduces a pump wave which amplifies the signal via stimulated Raman scattering over a long length scale. Alternatively, lumped amplification uses stimulated emission from short segments of doped fiber, e.g., erbium doped fiber (EDF), which is excited (pumped) by an outside source to amplify the pulse at discrete amplification points. Figure A.1 shows a simplified diagram of an EDF amplifier, which is pumped with light from two laser diodes (bidirectional pumping), although unidirectional pumping in only the forward or backward direction (co-directional and counter-directional pumping) is also very common [2]. The pump light, which is

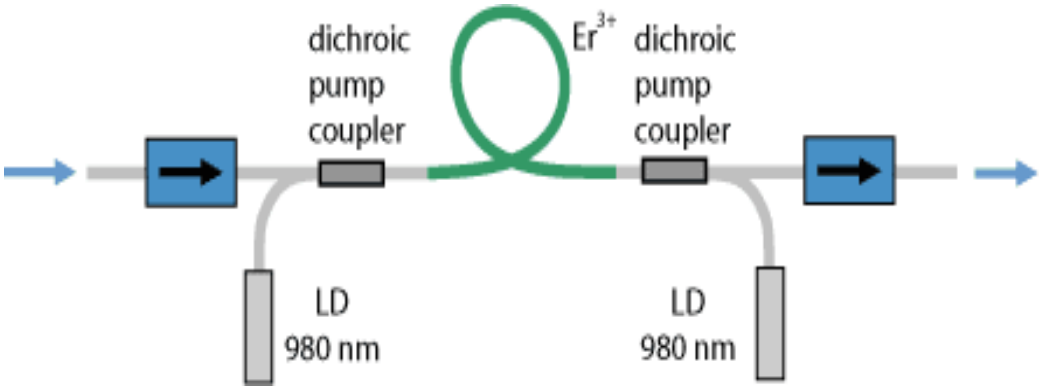


Figure A.1 Taken from [2]: Diagram of a typical erbium-doped-fiber amplifier. Two laser diodes provide the pump power, which are injected via (dichroic) fiber couplers. Optical isolators reduce the sensitivity of the device to back-reflections.

most often at the 980 nm wavelength, excites the erbium ions (Er^{3+}) into a higher energy state, from which they can amplify light via stimulated emission back to the ground-state. Of course, this also produces spontaneous emission in the form of noise, i.e., ASE noise, that propagates with the pulse.

Compared to all other length scales, the length of each lumped amplifier is extremely short (0.1m - 1m), indicating that the amplifiers can be modeled as point sources which are periodically positioned at a distance of L_{amp} apart, each contributing a (dimensionless) strength $g = \alpha L_{\text{amp}}$. In practice L_{amp} is a critical design parameter that is chosen based other system values, e.g., the loss coefficient α. Consequently, the ASE noise generated in the amplification process is modeled as point sources of

mean zero white noise, Gaussian distributed and delta correlated at individual values of τ. Additionally, the there is no correlation between distinct amplifiers, which implies a δ-correlation in both space and time. Thus, the propagation of pulses in the presence of amplification is governed by

$$i\frac{\partial A}{\partial \zeta} + i\frac{\alpha}{2}A - \operatorname{sgn}(k_0'')\,\frac{|k_0''|}{2}\frac{\partial^2 A}{\partial \tau^2} + \gamma|A|^2 A = i\frac{g}{2}\sum_{m=1}^{N_\mathrm{a}} \delta(\zeta - mL_\mathrm{amp})\,A + i\sum_{m=1}^{N_\mathrm{a}} N_m(\tau)\,\delta(\zeta - mL_\mathrm{amp})\,, \tag{A.45a}$$

where $\delta(x)$ and $\delta_{j,k}$ are the Kronecker and Dirac delta functions, respectively, and N_a is the number of amplifiers in the transmission line. The noise statistics are given by

$$\mathbb{E}[N_j(\tau)] = 0 \quad \text{and} \quad \mathbb{E}\big[N_j(\tau_1)\bar{N}_k(\tau_2)\big] = \tilde{\sigma}^2\delta_{jk}\,\delta(\tau_1 - \tau_2)\,, \tag{A.45b}$$

where the noise strength $\tilde{\sigma}^2$ is a dimensional quantity derived from analysis of the amplification process at the microscopic scale [29] and takes the form

$$\tilde{\sigma}^2 = \hbar\omega_0 n_\mathrm{sp}(G-1). \tag{A.45c}$$

Here $\hbar$ is the reduced Planck constant (Dirac constant), ω_0 is the frequency of the carrier wave, $G = \exp(g) = \exp(\alpha L_\mathrm{amp})$ is the dimensionless gain factor of the amplifier and $0 < n_\mathrm{sp} < 1$ is the dimensionless spontaneous emissions factor associated with the microscopic model of the amplifier.

A.2.2 Dimensionless NLSE

Since the investigations in this thesis are only concerned with the dynamics of a single pulse, the time variable and pulse amplitude are respectively scaled using the full-width-at-half-maximum (FWHM), denoted T_fwhm, and the peak launch power, denoted P_0, of a typical Schrödigner soliton[2]

[2]Note that the FWHM is related to the width through the relation $T_\mathrm{fwhm} = 2f^{-1}(1/2)T_0$, where the pulse is given by $f(t/T_0)$, with T_0 representing the normal pulse width.

The scaling for the distance variable is more complicated since there are multiple natural length scales from which to choose, the three most relevant being the amplifier spacing $L_{\rm amp}$ (which has specific relation to the loss length scale given by $L_{\rm loss} = 1/\alpha$), the dispersion length scale $L_{\rm dis} = T_{\rm fwhm}^2/|k_0''|$ (length over which the dispersion becomes $O(1)$) and the nonlinear length scale $L_{\rm nl} = 1/\gamma P_0$ (length over which nonlinear effects becomes $O(1)$). However, the effects of dispersion and nonlinearity must mutually compensate for solitons to propagate over significant distances, which is mathematically equivalent to the restriction $L_{\rm dis} \sim L_{\rm nl}$. Since the nonlinear length scale is in inverse proportion to the initial peak pulse power P_0, this restriction is often met by adjusting the peak power of the launched pulse. Assuming this is done here, there are only two effective lengths with which to scale the distance variable; the shorter amplification/loss length or the longer dispersion/nonlinear length. Since this thesis is concerned with the long distance propagation of optical pulses, it is appropriate to scale with respect to the dispersion/nonlinear length.

Introducing the non-dimensional variables

$$U(z,t) = \frac{A}{\sqrt{P_0}}, \quad t = \frac{\tau}{T_{\rm fwhm}} \quad \text{and} \quad z = \frac{\zeta}{L_{\rm dis}}, \tag{A.46}$$

results in the dimensionless equation,

$$\begin{aligned} i\frac{\partial U}{\partial z} + i\frac{1}{2}\left(\alpha L_{\rm dis} - g\sum_{m=1}^{N_{\rm a}} \delta(z - m z_{\rm a})\right) U + \frac{1}{2}\frac{\partial^2 U}{\partial t^2} + \frac{L_{\rm dis}}{L_{\rm nl}}|U|^2 U \\ = i\sum_{m=1}^{N_{\rm a}} n_m(t)\,\delta(z - m z_{\rm a})\,, \end{aligned} \tag{A.47a}$$

where

$$\mathbb{E}[n_j(t)] = 0 \quad \text{and} \quad \mathbb{E}[n_j(t_1)\bar{n}_k(t_2)] = \sigma^2 \delta_{jk}\,\delta(t_1 - t_2)\,, \tag{A.47b}$$

which has a non-dimensional noise strength of $\sigma^2 = \tilde{\sigma}^2/P_0 T_{\rm fwhm}$. Note that $z_{\rm a} = L_{\rm amp}/L_{\rm dis}$ is the non-dimensional period of the amplification cycle.

Table A.1 Typical Values for Coefficients of Single Mode Fiber in Long-haul Optical Communication Systems [1].

Quantity	Parameter	Value	Units
Absorption	α	0.2	dB/km
Pulse Width (FWHM)	T_{fwhm}	20	ps
Carrier Frequency	ω_0	150	THz
Group Velocity	$\frac{1}{k_0'}$	2.04×10^5	km/s
Group Velocity Dispersion	$\lvert k_0'' \rvert$	2.0	ps^2/km
Nonlinear Refractive Index	n_2	2.4×10^{-20}	m^2/W
Effective Core Area	A_{eff}	50	μm^2
Nonlinear Coefficient	γ	2.0	1/W km

Although Equation (A.47a) is in non-dimensional form, specific values for the dimensional coefficients have yet to be given. Typical values for these coefficients in a long-haul systems using single mode fiber are given in Table A.1[3] which indicate that $L_{\text{amp}} \approx 1$ km and $L_{\text{dis}} \approx 200$ km.

A.2.3 Path-Averaging NLSE

Equation (A.47a) is a dimensionless, stochastically forced, version of the NLSE in the presence of loss from absorption and gain from amplification. With the above scaling, the non-dimensional period of the amplification cycle, i.e., $z_{\text{a}} = L_{\text{amp}}/L_{\text{dis}} \approx 0.005$, forms a small parameter which can be used to average out the fast dynamics of absorption and amplification.

[3]Note that dB denotes the (dimensionless) scaling units of decibels which can be transformed to the original value through the relation $x = 10^{\frac{x_{\text{dB}}}{10}}$.

By introducing the variable, $x = z/z_{\rm a}$, equation (A.47a) can be rewritten as

$$\begin{aligned} iz_{\rm a}\frac{\partial U}{\partial z} + i\frac{\partial U}{\partial x} + i\frac{1}{2}\left(\alpha L_{\rm amp} - g\sum_{m=1}^{N_{\rm a}}\delta(x-m)\right)U \\ + z_{\rm a}\frac{1}{2}\frac{\partial^2 U}{\partial t^2} + z_{\rm a}\frac{L_{\rm dis}}{L_{\rm nl}}|U|^2 U = iz_{\rm a}\sum_{m=1}^{N_{\rm a}} n_m(t)\,\delta(z-mz_{\rm a})\,, \end{aligned} \tag{A.48}$$

where the noise terms are keep at $O(z_{\rm a})$ due to the small value of the noise strength. The solution to equation is assumed to take the form of a power series expansion in the small parameter, i.e.,

$$U(z,x,t) = u_0(z,x,t) + z_{\rm a}u_1(z,x,t) + z_{\rm a}^2 u_2(z,x,t) + \ldots, \tag{A.49}$$

which when inserted into equation (A.48) gives at $O(1)$,

$$i\frac{\partial u_0}{\partial x} + i\frac{1}{2}\left(\alpha L_{\rm amp} - g\sum_{m=1}^{N_{\rm a}}\delta(x-m)\right)u_0 = 0. \tag{A.50}$$

Assumed that the gain exactly counters the loss at the discrete amplification points, i.e., $g = \alpha L_{\rm amp}$, the solution of this equation takes the form

$$u_0(z,x,t) = \exp(-h(x))\,u(z,t), \tag{A.51a}$$

with

$$h(x) = \frac{\alpha L_{\rm amp}}{2}\left(x - \sum_{m=1}^{N_{\rm a}}{\rm H}(x-m)\right), \tag{A.51b}$$

which consists of a rapidly varying function ($\exp(-h(x))$) and a slowly varying function ($u(z,t)$) that remains undetermined.

At $O(z_{\rm a})$, the expansion gives

$$\begin{aligned} i\frac{\partial u_1}{\partial z} + i\frac{1}{2}\left(\alpha L_{\rm amp} - g\sum_{m=1}^{N_{\rm a}}\delta(x-m)\right)u_1 \\ = -\left[\exp(-h(x))\,i\frac{\partial u}{\partial z} + \exp(-h(x))\,\frac{1}{2}\frac{\partial^2 u}{\partial t^2} + \exp(-3h(x))\,\frac{L_{\rm dis}}{L_{\rm nl}}|u|^2 u\right] \\ + i\sum_{m=1}^{N_{\rm a}} n_m(t)\,\delta(z-mz_{\rm a})\,, \end{aligned} \tag{A.52}$$

where the solution found at $O(1)$ was used on the right hand-side. Following the normal procedure, an application of the Fredholm alternative [42] provides a solvability condition in the form of an orthogonality constraint, resulting in an evolution equation for $u(z,t)$. By defining the inner product as

$$\langle f, g\rangle = \mathrm{Re}\left[\int_0^1 \bar{f}(x)\, g(x)\, dx\right], \tag{A.53}$$

the homogeneous equation associated with the adjoint of the right-hand-side of equation (A.50) is found to be

$$\frac{\partial u_1^\dagger}{\partial z} - \frac{1}{2}\left(\alpha L_{\mathrm{amp}} - g\sum_{m=1}^{N_\mathrm{a}} \delta(x-m)\right) u_1^\dagger = 0, \tag{A.54}$$

which has the solution

$$u_1^\dagger(x) = \exp(h(x))\, u^\dagger(z,t). \tag{A.55}$$

Projecting this on both sides of equation (A.52) gives

$$i\frac{\partial u}{\partial z} + \frac{1}{2}\frac{\partial^2 u}{\partial t^2} + \frac{G-1}{G\ln(G)}\frac{L_{\mathrm{dis}}}{L_{\mathrm{nl}}}|u|^2 u = i\sum_{m=1}^{N_\mathrm{a}} n_m(t)\,\delta(z - mz_\mathrm{a}), \tag{A.56}$$

where as previously stated, $G = \exp(g) = \exp(\alpha L_{\mathrm{amp}})$. The last step in the process is determining the value of initial power P_0 to ensure that the nonlinear and dispersion coefficients are in the correct proportion to support solitons. This amounts to choosing

$$P_0 = \frac{G\ln(G)}{(G-1)}\frac{1}{\gamma L_{\mathrm{dis}}}. \tag{A.57}$$

However, since the dimensionless noise strength also depends on the initial power, the final equation becomes

$$i\frac{\partial u}{\partial z} + \frac{1}{2}\frac{\partial^2 u}{\partial t^2} + |u|^2 u = i\sum_{m=1}^{N_\mathrm{a}} n_m(t)\,\delta(z - mz_\mathrm{a}), \tag{A.58}$$

where

$$\mathbb{E}[n_j(t)] = 0 \quad \text{and} \quad \mathbb{E}[n_j(t_1)\bar{n}_k(t_2)] = \sigma^2 \delta_{jk}\, \delta(t - t')\,, \tag{A.59}$$

with

$$\sigma^2 = \frac{G-1}{G\ln(G)} \frac{\gamma L_{\text{dis}} \tilde{\sigma}^2}{T_{\text{fwhm}}} = \frac{(G-1)^2}{G\ln(G)} \left(\frac{\gamma L_{\text{dis}} \hbar \omega_0 n_{\text{sp}}}{T_{\text{fwhm}}} \right). \tag{A.60}$$

A.3 The NLSE+DM and DMNLSE for Optical Fiber

As discussed in Chapter 5, a transmission line with dispersion management consists of alternating segments of optical fiber with individual dispersion coefficients that are large in absolute value and opposite in algebraic sign. Assuming an ideal dispersion map (periodic, with a period equal to the amplification cycle), the equation for dispersion managed fiber in the presence of amplification takes the same form as that of (A.45), but with coefficients that piecewise constant functions representing the two different segments of fiber (the NLSE+DM),

$$\begin{aligned} i\frac{\partial A}{\partial \zeta} + i\frac{1}{2}\left(\alpha(\zeta) - g \sum_{m=1}^{N_{\text{a}}} \delta(\zeta - mL_{\text{amp}}) \right) A - \frac{1}{2} k(\zeta) \frac{\partial^2 A}{\partial \tau^2} + \gamma(\zeta)|A|^2 A \\ = i \sum_{m=1}^{N_{\text{a}}} N_m(\tau)\, \delta(\zeta - mL_{\text{amp}})\,, \end{aligned} \tag{A.61a}$$

$$\alpha(\zeta) = \begin{cases} \alpha_1, & 0 \le \zeta < \frac{(1-\theta)}{2} L_{\text{amp}} \\ \alpha_2, & \frac{(1-\theta)}{2} L_{\text{amp}} \le \zeta < \frac{(1+\theta)}{2} L_{\text{amp}} \\ \alpha_1, & \frac{(1+\theta)}{2} L_{\text{amp}} \le \zeta < L_{\text{amp}}, \end{cases} \tag{A.61b}$$

$$k(\zeta) = \begin{cases} |k_0''|_1, & 0 \le \zeta < \frac{(1-\theta)}{2} L_{\text{amp}} \\ -|k_0''|_2, & \frac{(1-\theta)}{2} L_{\text{amp}} \le \zeta < \frac{(1+\theta)}{2} L_{\text{amp}} \\ |k_0''|_1, & \frac{(1+\theta)}{2} L_{\text{amp}} \le \zeta < L_{\text{amp}}, \end{cases} \tag{A.61c}$$

and

$$\gamma(\zeta) = \begin{cases} \gamma_1, & 0 \leq \zeta < \frac{(1-\theta)}{2} L_{\text{amp}} \\ \gamma_2, & \frac{(1-\theta)}{2} L_{\text{amp}} \leq \zeta < \frac{(1+\theta)}{2} L_{\text{amp}} \\ \gamma_1, & \frac{(1+\theta)}{2} L_{\text{amp}} \leq \zeta < L_{\text{amp}}, \end{cases} \tag{A.61d}$$

where the noise statistics are equivalent to those defined in equation (A.45). Note that the subscripts on the dimensional values in the piece-wise continuous coefficient functions denote the two distinct fibers that comprise the dispersion map and θ represents the proportion of the dispersion map period that is occupied by the compensating fiber.

A.3.1 Dimensionless NLSE+DM

The non-dimensionalization of Equation (A.61) follows the non-dimensionalization of Equation (A.45) very closely. The time variable is scaled using the full-width-at-half-maximum (FWHM) T_{fwhm} of a typical (DM) pulse and the pulse power is scaled using the peak launch power P_0. However, since the parameters that define the length scales now vary with distance, the averaged values of dispersion and nonlinear coefficients are used to define the characteristic length scales for dispersion and nonlinear effects, i.e.,

$$L_{\text{dis}} = \frac{T_{\text{fwhm}}^2}{|k_0''|_1(1-\theta) - |k_0''|_2\theta} \quad \text{and} \quad L_{\text{nl}} = \frac{1}{(\gamma_1(1-\theta) + \gamma_2\theta)P_0}. \tag{A.62}$$

In addition, the averaged loss length scale is given as

$$L_{\text{loss}} = \frac{1}{\alpha_1(1-\theta) + \alpha_2\theta}, \tag{A.63}$$

which is used to define an appropriate amplifier spacing L_{amp}.

With these definitions, the same scalings can be used here as was used in the constant coefficient case, i.e.,

$$U(z,t) = \frac{A}{\sqrt{P_0}}, \quad t = \frac{\tau}{T_{\text{fwhm}}} \quad \text{and} \quad z = \frac{\zeta}{L_{\text{dis}}}, \tag{A.64}$$

which results in the dimensionless equation

$$\begin{aligned} i\frac{\partial U}{\partial z} + i\frac{1}{2}\left(\alpha(z)\frac{L_{\text{dis}}}{L_{\text{loss}}} - g\sum_{m=1}^{N_{\text{a}}}\delta(z-mz_{\text{a}})\right)U + \frac{1}{2}k(z)\frac{\partial^2 U}{\partial t^2} + \gamma(z)|U|^2U \\ = i\sum_{m=1}^{N_{\text{a}}} n_m(t)\,\delta(z-mz_{\text{a}})\,, \end{aligned} \tag{A.65a}$$

where as before,

$$\mathbb{E}[n_j(t)] = 0 \quad \text{and} \quad \mathbb{E}[n_j(t_1)\bar{n}_k(t_2)] = \sigma^2\delta_{jk}\,\delta(t_1-t_2)\,, \tag{A.65b}$$

with a non-dimensional noise strength of $\sigma^2 = \tilde{\sigma}^2/P_0T_{\text{fwhm}}$ and non-dimensional amplification period $z_{\text{a}} = L_{\text{amp}}/L_{\text{dis}}$. Now however, the non-dimensional coefficient functions are given by

$$\alpha(z) = \begin{cases} \frac{\alpha_1}{\alpha_1(1-\theta)+\alpha_2\theta}, & 0 \le z < \frac{(1-\theta)}{2}z_{\text{a}} \\ \frac{\alpha_2}{\alpha_1(1-\theta)+\alpha_2\theta}, & \frac{(1-\theta)}{2}z_{\text{a}} \le z < \frac{(1+\theta)}{2}z_{\text{a}} \\ \frac{\alpha_1}{\alpha_1(1-\theta)+\alpha_2\theta}, & \frac{(1+\theta)}{2}z_{\text{a}} \le z < z_{\text{a}}, \end{cases} \tag{A.65c}$$

$$k(z) = \begin{cases} \frac{|k_0''|_1}{|k_0''|_1(1-\theta)-|k_0''|_2\theta}, & 0 \le z < \frac{(1-\theta)}{2}z_{\text{a}} \\ -\frac{|k_0''|_2}{|k_0''|_1(1-\theta)-|k_0''|_2\theta}, & \frac{(1-\theta)}{2}z_{\text{a}} \le z < \frac{(1+\theta)}{2}z_{\text{a}} \\ \frac{|k_0''|_1}{|k_0''|_1(1-\theta)-|k_0''|_2\theta}, & \frac{(1+\theta)}{2}z_{\text{a}} \le z < z_{\text{a}}, \end{cases} \tag{A.65d}$$

Table A.2 Typical Values for SMF and DCF Coefficients for Optical Communication Systems with Dispersion Management [1].

Quantity	Parameter	SMF Values	DCF Values	Units
Absorption	α	0.2	0.5	dB/km
Group Velocity	$\frac{1}{k_0'}$	2.048×10^5	2.034×10^5	km/s
Group Velocity Dispersion	k_0''	-16	127	ps^2/km
Nonlinear Refractive Index	n_2	2.4×10^{-20}	2.8×10^{-20}	m^2/W
Effective Core Area	A_{eff}	50	20	μm^2
Nonlinear Coefficient	γ	2.0	5.0	1/W km

and

$$\gamma(z) = \frac{L_{\text{dis}}}{L_{\text{nl}}} \begin{cases} \frac{\gamma_1}{\gamma_1(1-\theta)+\gamma_2\theta}, & 0 \le z < \frac{(1-\theta)}{2} z_{\text{a}} \\ \frac{\gamma_2}{\gamma_1(1-\theta)+\gamma_2\theta}, & \frac{(1-\theta)}{2} z_{\text{a}} \le z < \frac{(1+\theta)}{2} z_{\text{a}} \\ \frac{\gamma_1}{\gamma_1(1-\theta)+\gamma_2\theta}, & \frac{(1+\theta)}{2} z_{\text{a}} \le z < z_{\text{a}}. \end{cases} \tag{A.65e}$$

As discussed before, a realistic DM transmission line consists mainly of single-mode fiber (SMF) over much of the dispersion map, which is compensated by a relatively short segment of dispersion compensating fiber (DCF) placed just prior to the amplifier. Typical parameter values for both SMF and DCF are given in Table A.2, which shows that the major differences in values between the two are in the dispersion. Figure A.2 shows a more realistic,i.e., unbalanced ($\theta \neq 1/2$) dispersion map used in implementation of dispersion management. For numerical reasons however, it is advantageous to used an idealization of this where local dispersion values that are nearly equivalent in absolute value and a dispersion map which is symmetrical around the midpoint between amplifiers ($\theta = 1/2$) as shown in Figure A.3. It should be noted that the map strength, i.e., the L^1-norm of the dispersion map, and the average dispersion are equivalent between both realistic and ideal dispersion maps.

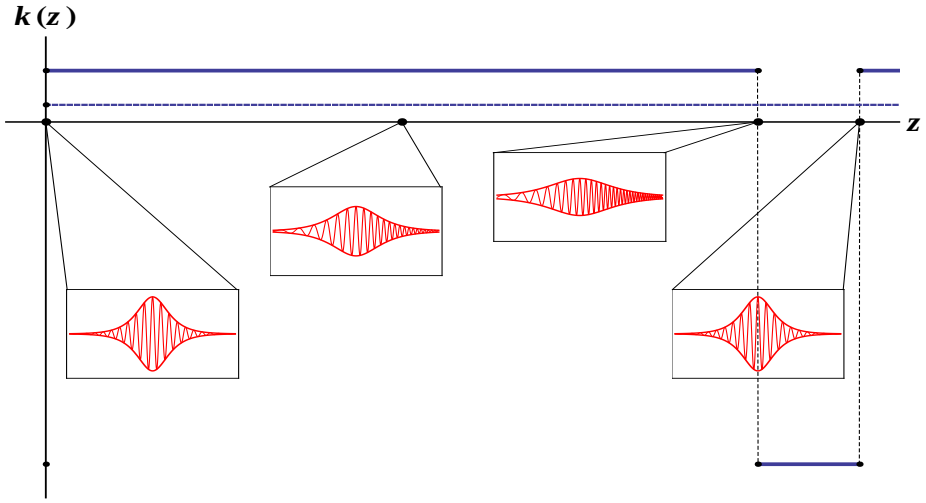


Figure A.2 Simplified diagram of a realistic dispersion map with pulse dynamics. The solid blue line indicates local dispersion values and the dashed line indicates the average dispersion over the entire map.

A.3.2 Averaged NLSE+DM: DMNLSE

Just as the small length scale variations from the amplitude and loss cycle could be factored out of the leading order solution of the constant coefficient case, an analogous averaging can be done here. However, in this case, the variations in the dispersion are large enough to to contribute to the fast dynamics at leading order.

Introducing the small length scale variable, $x = z/z_{\mathrm{a}}$ where $z_{\mathrm{a}} = L_{\mathrm{amp}}/L_{\mathrm{dis}}$, equation (A.65a) can be rewritten as

$$\begin{aligned} iz_{\mathrm{a}}\frac{\partial U}{\partial z} + i\frac{\partial U}{\partial x} + i\frac{1}{2}\left(\alpha(x) - g\sum_{m=1}^{N_{\mathrm{a}}}\delta(x-m)\right)U + \frac{z_{\mathrm{a}}}{2}\frac{\partial^2 U}{\partial t^2} + d(x)\frac{\partial^2 U}{\partial t^2} \\ + z_{\mathrm{a}}\gamma(x)|U|^2U = iz_{\mathrm{a}}\sum_{m=1}^{N_{\mathrm{a}}} n_m(t)\,\delta(z - mz_{\mathrm{a}})\,, \end{aligned} \tag{A.66a}$$

with new x dependent, $O(1)$ coefficients, of the form

$$\alpha(x) = \frac{L_{\mathrm{amp}}}{L_{\mathrm{loss}}}\begin{cases} \frac{\alpha_1}{\alpha_1(1-\theta)+\alpha_2\theta}, & 0 \le x < \frac{(1-\theta)}{2} \\ \frac{\alpha_2}{\alpha_1(1-\theta)+\alpha_2\theta}, & \frac{(1-\theta)}{2} \le x < \frac{(1+\theta)}{2} \\ \frac{\alpha_1}{\alpha_1(1-\theta)+\alpha_2\theta}, & \frac{(1+\theta)}{2} \le x < 1, \end{cases} \tag{A.66b}$$

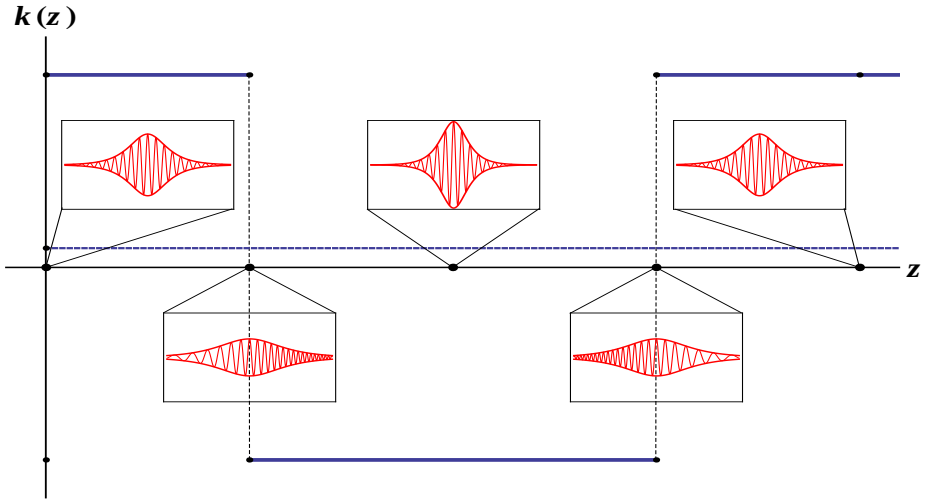


Figure A.3 Simplified diagram of an idealized dispersion map with pulse dynamics. The solid blue line indicates local dispersion values and the dashed line indicates the average dispersion over the entire map.

$$d(x) = \frac{z_{\mathrm{a}}}{2}\begin{cases} \theta\frac{|k_0''|_1+|k_0''|_2}{|k_0''|_1(1-\theta)-|k_0''|_2\theta}, & 0 \le x < \frac{(1-\theta)}{2} \\ -(1-\theta)\frac{|k_0''|_1+|k_0''|_2}{|k_0''|_1(1-\theta)-|k_0''|_2\theta}, & \frac{(1-\theta)}{2} \le x < \frac{(1+\theta)}{2} \\ \theta\frac{|k_0''|_1+|k_0''|_2}{|k_0''|_1(1-\theta)-|k_0''|_2\theta}, & \frac{(1+\theta)}{2} \le x < 1, \end{cases} \tag{A.66c}$$

and

$$\gamma(x) = \frac{L_{\mathrm{dis}}}{L_{\mathrm{nl}}}\begin{cases} \frac{\gamma_1}{\gamma_1(1-\theta)+\gamma_2\theta}, & 0 \le x < \frac{(1-\theta)}{2} \\ \frac{\gamma_2}{\gamma_1(1-\theta)+\gamma_2\theta}, & \frac{(1-\theta)}{2} \le x < \frac{(1+\theta)}{2} \\ \frac{\gamma_1}{\gamma_1(1-\theta)+\gamma_2\theta}, & \frac{(1+\theta)}{2} \le x < 1. \end{cases} \tag{A.66d}$$

It is important to note that due to the large variations in the dispersion, the function $k(x)/2$ in equation (A.65a) was separated into an $O(1)$ mean term and an $O(1/z_{\mathrm{a}})$ varying term by writing it in the form

$$\frac{1}{2}k(x) = \frac{1}{2}\int_0^1 k(x)dx + \frac{1}{z_{\mathrm{a}}}\left(\frac{z_{\mathrm{a}}}{2}k(x) - \frac{z_{\mathrm{a}}}{2}\int_0^1 k(x)dx\right) = \frac{1}{2} + \frac{1}{z_{\mathrm{a}}}d(x), \tag{A.67}$$

where $d(x)$ is the $O(1)$ function given in (A.66c).

The leading order solution to this equation can be found by assuming a solution in the form of an expansion in the small parameter, i.e.,

$$U(z,x,t) = u_0(z,x,t) + z_{\mathrm{a}} u_1(z,x,t) + z_{\mathrm{a}}^2 u_2(z,x,t) + \dots. \tag{A.68}$$

Inserting this back into equation (A.48) gives at $O(1)$,

$$i\frac{\partial u_0}{\partial x} + i\frac{1}{2}\left(\alpha(x) - g\sum_{m=1}^{N_{\mathrm{a}}}\delta(x-m)\right)u_0 + d(x)\frac{\partial^2 u_0}{\partial t^2} = 0, \tag{A.69}$$

Using the Fourier transform (FT), defined as

$$\begin{aligned}\mathcal{F}[f(t)] = \hat{f}(\omega) &= \int f(t)\exp(i\omega t)\ dt,\\ \mathcal{F}^{-1}\Big[\hat{f}(\omega)\Big] = f(t) &= \frac{1}{2\pi}\int \hat{f}(\omega)\exp(-i\omega t)\ d\omega,\end{aligned} \tag{A.70}$$

the FT of equation (A.69) and its corresponding solution is given by

$$i\frac{\partial \hat{u}_0}{\partial x} + i\frac{1}{2}\left(\alpha(x) - g\sum_{m=1}^{N_{\mathrm{a}}}\delta(x-m)\right)\hat{u}_0 - \omega^2 d(x)\hat{u}_0 = 0, \tag{A.71}$$

and

$$\hat{u}_0(z,x,\omega) = \hat{u}(z,\omega)\exp(-A(x))\ \exp\big(-i\omega^2 D(x)\big), \tag{A.72}$$

respectively, where

$$A(x) = \frac{1}{2}\int_0^x \alpha(y) - \frac{L_{\mathrm{amp}}}{L_{\mathrm{loss}}}\sum_{m=1}^{N_{\mathrm{a}}}\delta(y-m)\,dy, \quad D(x) = \int_0^x d(y)\,dy, \tag{A.73}$$

and $g = \int_0^1 \alpha(x)\,dx = \frac{L_{\mathrm{amp}}}{L_{\mathrm{loss}}}$ was used to indicate that the amplifier fully counters the absorption in the fiber. Note that $\hat{u}(z,\omega)$ is yet to be determined and represents the FT evolution of the "core" solution, whereas $\exp(-i\omega^2 D(x))$ captures the fast oscillations in the DM soliton's width and chirp, and $\exp(-A(x))$ captures the loss and gain.

The $O(z_p)$ equation and its FT is given by

$$\begin{aligned} i\frac{\partial u_1}{\partial x} + i\frac{1}{2}\left(\alpha(x) - \frac{L_{\text{amp}}}{L_{\text{loss}}}\sum_{m=1}^{N_{\text{a}}}\delta(x-m)\right)u_1 + d(x)\frac{\partial^2 u_1}{\partial t^2} = \\ -\left[i\frac{\partial u_0}{\partial z} + \frac{1}{2}\frac{\partial^2 u_0}{\partial t^2} + \gamma(x)|u_0|^2 u_0\right] + i\sum_{m=1}^{N_{\text{a}}} n_m(t)\,\delta(z - mz_{\text{a}})\,, \end{aligned} \tag{A.74a}$$

and

$$\begin{aligned} i\frac{\partial \hat{u}_1}{\partial x} + i\frac{1}{2}\left(\alpha(x) - \frac{L_{\text{amp}}}{L_{\text{loss}}}\sum_{m=1}^{N_{\text{a}}}\delta(x-m)\right)\hat{u}_1 - \omega^2 d(x)\hat{u}_1 = \\ -\left[i\frac{\partial \hat{u}_0}{\partial z} - \omega^2\frac{1}{2}\hat{u}_0 + \gamma(x)\mathcal{F}\big[|u_0|^2 u_0\big]\right] + i\sum_{m=1}^{N_{\text{a}}} \hat{n}_m(\omega)\,\delta(z - mz_{\text{a}})\,, \end{aligned} \tag{A.74b}$$

respectively. As expected, the right-hand side of equation (A.74b) contains the same operator as in the $O(1)$ equation, which implies that the evolution of $\hat{u}(z,\omega)$ can be found through the application of Fredholm theory, which as previously discussed, states that the right-hand side of equation (A.74b) must be orthogonal to the homogeneous solution of adjoint operator of the left-hand side [42], with orthogonality defined with respect to the inner product

$$\langle f, g\rangle = \text{Re}\left[\int_0^1 \bar{f}(x)\,g(x)\,dx\right]. \tag{A.75}$$

Using this inner product, the adjoint operator is found to be

$$i\frac{\partial \hat{u}_1^\dagger}{\partial x} - i\frac{1}{2}\left(\alpha(x) - \frac{L_{\text{amp}}}{L_{\text{loss}}}\sum_{m=1}^{N_{\text{a}}}\delta(x-m)\right)\hat{u}_1^\dagger - \omega^2 d(x)\hat{u}_1^\dagger, \tag{A.76}$$

which has the homogeneous solution of

$$\hat{u}_{\text{h}}^\dagger(z,x,t) = \hat{u}_{\text{h}}^\dagger(z,\omega)\exp(A(x))\,\exp\big(-i\omega^2 D(x)\big)\,. \tag{A.77}$$

Projecting this solution on both sides of equation (A.74b) gives

$$
\begin{aligned}
&i\frac{\partial \hat{u}}{\partial z} - \omega^2 \frac{1}{2}\hat{u} + \int_0^1 \gamma(x)\exp(-2A(x))\,\exp\big(i\omega^2 D(x)\big)\times \\
&\mathcal{F}\Big[\big|\mathcal{F}^{-1}\big[\hat{u}(z,\omega)\exp\big(-i\omega^2 D(x)\big)\big]\big|^2\,\mathcal{F}^{-1}\big[\hat{u}(z,\omega)\exp\big(-i\omega^2 D(x)\big)\big]\Big]\,dx = \\
&\quad i\sum_{m=1}^{N_\mathrm{a}} \hat{n}_m(\omega)\,\delta(z-mz_\mathrm{a})\,,
\end{aligned}
\tag{A.78}
$$

where the first order solution in (A.72) was used in the nonlinear term. The FT inside the integral can be rewritten as a convolution, i.e.,

$$
\begin{aligned}
&\mathcal{F}\Big[\big|\mathcal{F}^{-1}\big[\hat{u}(z,\omega)\exp\big(-i\omega^2 D(x)\big)\big]\big|^2\,\mathcal{F}^{-1}\big[\hat{u}(z,\omega)\exp\big(-i\omega^2 D(x)\big)\big]\Big] = \\
&\quad \frac{1}{(2\pi)^2}\int\int \hat{u}(z,\omega')\hat{u}(z,\omega'')\bar{\hat{u}}(z,\omega''+\omega'-\omega)\times \\
&\qquad\qquad \exp\big(i\left(2\omega''\omega' - 2\omega''\omega - 2\omega'\omega + \omega^2\right)D(x)\big)\;d\omega'\,d\omega'',
\end{aligned}
\tag{A.79}
$$

which when put back into the integral in equation (A.78), gives

$$
\begin{aligned}
&i\frac{\partial \hat{u}}{\partial z} - \omega^2\frac{1}{2}\hat{u} + \int\int \hat{u}(z,\omega_1+\omega)\hat{u}(z,\omega_2+\omega)\bar{\hat{u}}(z,\omega_2+\omega_1+\omega)\times \\
&\hat{K}(\omega_1,\omega_2)\,d\omega_1\,d\omega_2 = i\sum_{m=1}^{N_\mathrm{a}}\hat{n}_m(\omega)\,\delta(z-mz_\mathrm{a})\,,
\end{aligned}
\tag{A.80a}
$$

where

$$
\hat{K}(\omega_1,\omega_2) = \frac{1}{(2\pi)^2}\int_0^1 \gamma(x)\exp(-2A(x))\,\exp(i2\omega_1\omega_2 D(x))\,dx. \tag{A.80b}
$$

Equation (A.80a) is the DMNLSE in the Fourier domain which is converted to the time domain by taking the inverse FT which gives

$$
\begin{aligned}
&i\frac{\partial u}{\partial z} + \frac{1}{2}\frac{\partial^2 u}{\partial t^2} + \int\int u(z,t_2+t)u(z,t_1+t)\bar{u}(z,t_1+t_2+t)\times \\
&K(t_1,t_2)\,dt_1 dt_2 = i\sum_{m=1}^{N_\mathrm{a}} n_m(t)\,\delta(z-mz_\mathrm{a})\,,
\end{aligned}
\tag{A.81a}
$$

where

$$
K(t_1,t_2) = \int\int \exp(-i\omega_1 t_1)\,\exp(-i\omega_2 t_2)\,\hat{K}(\omega_1,\omega_2)\,d\omega_1\,d\omega_2. \tag{A.81b}
$$

A.3.3 Simplifying Assumptions

At this point it is advantageous to make a few assumptions to simplify the above equations. From the values in Table A.2, the loss coefficients for both types of fiber that comprise the dispersion map are approximately equal in value. Because of this, it is reasonable to let both coefficients equal their average values, i.e.,

$$\alpha_1 = \alpha_2 \approx \alpha_1(1-\theta) + \alpha_2\theta, \tag{A.82}$$

which implies that

$$A(x) = \frac{1}{2}\frac{L_{\text{amp}}}{L_{\text{loss}}}\left(x - \sum_{m=1}^{N_{\text{a}}} \text{H}(x-m)\right). \tag{A.83}$$

Now, if it assumed that the period of the amplification is much smaller than the loss length scale, i.e., $\frac{L_{\text{amp}}}{L_{\text{loss}}} \ll 1$, then the function $\exp(-2A(x))$ can be well approximated by its mean value,

$$\exp(-2A(x)) \approx \frac{G-1}{G\ln(G)}, \tag{A.84}$$

where $G = \exp\left(\frac{L_{\text{amp}}}{L_{\text{loss}}}\right)$. By the same argument, both nonlinear coefficients can be approximated by their mean value, giving

$$\gamma_1 = \gamma_2 \approx \gamma_1(1-\theta) + \gamma_2\theta. \tag{A.85}$$

Using both of these assumptions, the kernel of the convolution in equation (A.80b) can be written as

$$\begin{aligned}\hat{K}(\omega_1,\omega_2) &\approx \frac{1}{(2\pi)^2}\frac{G-1}{G\ln(G)}\int_0^1 \exp(i2\omega_1\omega_2 D(x))\,dx \\ &= \frac{1}{(2\pi)^2}\frac{G-1}{G\ln(G)}\,\text{sinc}\left(\omega_1\omega_2\frac{z_{\text{a}}\hat{d}}{2}\theta(1-\theta)\right).\end{aligned} \tag{A.86}$$

Since many of the non-dimensional functions, such as $d(x)$, $D(x)$ and $\hat{K}(\omega_1,\omega_2)$, depend on several different dispersion map parameters, it is advantageous to express

all dimensionless functions in terms of the L^1-norm of the dispersion variations, denoted s_{map}, which is referred to as the map strength, i.e.,

$$s_{map} = \int_0^1 |d(x)|\, dx = z_{\mathrm{a}} \hat{d}(1-\theta)\theta \quad \text{where} \quad \hat{d} = \frac{|k_0''|_1 + |k_0''|_2}{|k_0''|_1(1-\theta) - |k_0''|_2\theta}. \tag{A.87}$$

Using the definition of map strength in conjunction with the assumptions above, gives a DMNLSE in the Fourier domain of the form

$$\begin{aligned} i\frac{\partial \hat{u}}{\partial z} - \omega^2 \frac{1}{2}\hat{u} + \int\int \hat{u}(z,\omega_1+\omega)\hat{u}(z,\omega_2+\omega)\bar{\hat{u}}(z,\omega_2+\omega_1+\omega)\times \\ \hat{K}(\omega_1,\omega_2;s_{map})\, d\omega_1\, d\omega_2 = i \sum_{m=1}^{N_{\mathrm{a}}} \hat{n}_m(\omega)\, \delta(z - m z_{\mathrm{a}})\,, \end{aligned} \tag{A.88a}$$

where

$$\hat{K}(\omega_1,\omega_2;s_{map}) = \frac{1}{(2\pi)^2}\,\mathrm{sinc}\Big(\omega_1\omega_2\frac{s_{map}}{2}\Big)\,, \tag{A.88b}$$

which in the time domain becomes

$$\begin{aligned} i\frac{\partial u}{\partial z} + \frac{1}{2}\frac{\partial^2 u}{\partial t^2} + \int\int u(z,t_2+t)u(z,t_1+t)\bar{u}(z,t_1+t_2+t)\times \\ K(t_1,t_2;s_{map})\, dt_1 dt_2 = i \sum_{m=1}^{N_{\mathrm{a}}} n_m(t)\, \delta(z - m z_{\mathrm{a}})\,, \end{aligned} \tag{A.89a}$$

where

$$K(t_1,t_2,s_{map}) = \int\int \exp(-i\omega_1 t_1)\, \exp(-i\omega_2 t_2)\, \hat{K}(\omega_1,\omega_2;s_{map})\, d\omega_1\, d\omega_2. \tag{A.89b}$$

Notice that the initial power was chosen as

$$P_0 = \frac{G\ln(G)}{G-1}\frac{1}{L_{\mathrm{dis}}(\gamma_1(1-\theta)+\gamma_2\theta)} = \frac{G\ln(G)}{G-1}\frac{|k_0''|_1(1-\theta) - |k_0''|_2\theta}{T_{\mathrm{fwhm}}^2(\gamma_1(1-\theta)+\gamma_2\theta)}. \tag{A.90}$$

so the nonlinear coefficient is normalized. With this choice of P_0 the noise statistics are given as

$$\mathbb{E}[n_k(t)] = 0 \quad \text{and} \quad \mathbb{E}[n_k(t_1)\bar{n}_j(t_2)] = \sigma^2\, \delta(t_1 - t_2)\, \delta_{k,j} \tag{A.91}$$

where

$$\sigma^2 = \frac{(G-1)^2}{G\ln(G)}\left(\frac{(\gamma_1(1-\theta)+\gamma_2\theta)L_{\text{dis}}\hbar\omega_0 n_{\text{sp}}}{T_{\text{fwhm}}}\right). \tag{A.92}$$

A.4 Modulation Formats

We have shown that the soliton solution of the NLSE has four free parameters, corresponding to invariance under various transformations. When used as a bit carrier, information is encoded into the soliton parameters, transmitted, received, and decoded. Each bit occupies a given time window based on the data rate (e.g., 10Gb/s $\rightarrow$ 100ps window). At reception, the soliton parameters are read and the receiver determines if that window contains a 0 or 1. However, since there four free parameters, there is a number of coding schemes that one can use to encode the information. Three of the most used are discussed below.

A.4.1 On-Off Keying (OOK)

In the on-off keying scheme, the amplitude is the encoding parameter. The optical power integrated over a particular window is compared to an energy threshold to determine if window contains a 1 or a 0. The two main source for bit errors in this scheme are time shifts and loss.

A.4.2 Phase-Shift Keying (PSK)

In phase-shift keying, every window contains a pulse. At detection, the phase of each pulse is compared to that of a local oscillator. If the soliton is in phase, they are read as 1; if the pulse is out of phase it is read as 0.

A.4.3 Differential Phase-Shift Keying (DPSK)

Differential phase-shift keying is similar to PSK, in that the phase is used to encode information. But instead of depending directly of the phase, the bit is encoded in the

phase *difference* between two pulses, which does not require an external oscillator. If the phase difference is zero, the detector reads 1, if the phase change is π, the detector reads a 0.

APPENDIX B

NUMERICAL METHODS

This appendix contains an outline for many of the numerical methods used through out this thesis. In particular, it contains the derivation of a pseudo-spectral method based on the fourth order Runge-Kutta scheme, which is used to solve the various nonlinear evolution equation presented in this document. It should be noted that this technique relies heavily on the celebrated Fast Fourier Transform (FFT) which allows the numerical Fourier transform to be computed in $O(N\ln(N))$ operations as opposed to $O(N^2)$, which is number of operations a straightforward approach would require. In both approximations, N is the number of computational modes representing the numerical solution.

B.1 Integrating Factor Fourth-Order Runge-Kutta Method

This method, denoted by IFRK, is based on the method of lines [90] applied to the evolution equation written in the Fourier domain. For simplicity, the IFRK method will be derived for the case of a general nonlinearity, which can be substituted for as need to solve all other version of this equation.

Consider the equation,

$$\frac{\partial u}{\partial z} = i\frac{1}{2}\frac{\partial^2 u}{\partial t^2} + N(u) \tag{B.1}$$

which is to be solved on the infinite strip given by $t \in (\infty, -\infty)$ and $z \in [0, Z_f]$, where $N(u)$ represents a general nonlinear term. By taking the Fourier transform in the variable t, this equation takes form in the Fourier domain as

$$\frac{d\hat{u}}{dz} = -i\omega^2\frac{1}{2}\hat{u} + i\mathcal{F}[N(u)] \tag{B.2}$$

With the differentiation in time variable eliminated, this equation can be discretized over the frequency variable ω giving

$$\frac{d\hat{u}_j}{dz} = i\frac{\omega_j^2}{2}\hat{u}_j + i\mathcal{F}[N(u_j)] \tag{B.3}$$

which represents a system of ODEs, i.e., method of lines, that are parameterized by ω_j, where $u_j = u(z, \omega_j)$. Note that, for simple nonlinearities, these ODEs are decoupled, however, in general this is not the case. Also notice that the nonlinear term is left in its general form, indicating that it should always be evaluated in the time domain through the use of the FFT.

The reason for working with in Fourier space is to eliminate any temporal derivative, that would require the use of finite differences, which are replaced by products of the frequency variable. However, this results in system of ODEs that contain the term ω_j^2, which can range over several orders of magnitude. Thus, the resulting system of ODEs are stiff, which require such methods as a backward differentiation formulas (BDF) or implicit Runge-Kutta (RK) to do the stepping in z. However, this particular form of stiffness can be easily dealt with by wrapping the ω_j^2 term up into a phase rotation, which is the foundation of the integrating factor method. Letting

$$\hat{u}_j(z) = \exp\left(i\frac{1}{2}\omega_j^2 z\right)\hat{v}_j(z) \tag{B.4}$$

gives

$$\begin{aligned}\frac{d\hat{u}_j}{dz} &= i\frac{1}{2}\omega_j^2 \exp\left(i\frac{1}{2}\omega_j^2 z\right)\hat{v}_j(z) + \exp\left(i\frac{1}{2}\omega_j^2 z\right)\frac{\partial \hat{v}_j}{\partial z} \\ &= i\frac{\omega_j^2}{2}\exp\left(i\frac{1}{2}\omega_j^2 z\right)\hat{v}_j(z) + i\mathcal{F}\left[N\left(\mathcal{F}^{-1}\left[\exp\left(i\frac{1}{2}\omega_j^2 z\right)\hat{v}_j(z)\right]\right)\right]\end{aligned} \tag{B.5}$$

implying that

$$\frac{d\hat{v}_j}{dz} = i\exp\left(i\frac{1}{2}\omega_j^2 z\right)\mathcal{F}\left[N\left(\mathcal{F}^{-1}\left[\exp\left(i\frac{1}{2}\omega_j^2 z\right)\hat{v}_j(z)\right]\right)\right]. \tag{B.6}$$

Now that we have an equation that is no longer stiff, we can apply an explicit RK method, such as the commonly used fourth order RK method, which after discretization of the z variable as $z_n = n\Delta z$ and $\hat{v}_j(z_n) = \hat{v}_j^{(n)}$, gives

$$\hat{v}_j^{(n+1)} = \hat{v}_j^{(n)} + \frac{1}{6}\Delta z[K_1 + 2K_1 + 2K_3 + K_4] \tag{B.7a}$$

where

$$K_1 = f(z_n, \hat{v}_j^{(n)}), \tag{B.7b}$$

$$K_2 = f(z_n + \frac{1}{2}\Delta z, \hat{v}_j^{(n)} + \frac{1}{2}\Delta z K_1), \tag{B.7c}$$

$$K_3 = f(z_n + \tfrac{1}{2}\Delta z, \hat{v}_j^{(n)} + \frac{1}{2}\Delta z K_2), \tag{B.7d}$$

and

$$K_4 = f(z_n + \Delta z, \hat{v}_j^{(n)} + \Delta z K_3), \tag{B.7e}$$

where

$$f(z_n, \hat{v}_j^{(n)}) = i\exp\left(i\frac{1}{2}\omega_j^2 z_n\right)\mathcal{F}\left[N\left(\mathcal{F}^{-1}\left[\exp\left(i\frac{1}{2}\omega_j^2 z_n\right)\hat{v}_j^{(n)}\right]\right)\right]. \tag{B.7f}$$

Note that each stage requires two FFTs, so the computational cost is eight FFTs per evolutionary step. Figure B.1 plots the Log of the local error in the IFRK method as a function of the Log of the step size, confirming that the local error is $O(\Delta z^4)$.

B.2 Parameter Extraction

This section describes the methods used to extract the parameters of noisy soliton solution. There are three methods discussed, an integral representation, an iterative method based on projections and a method based on the inverse scattering transform

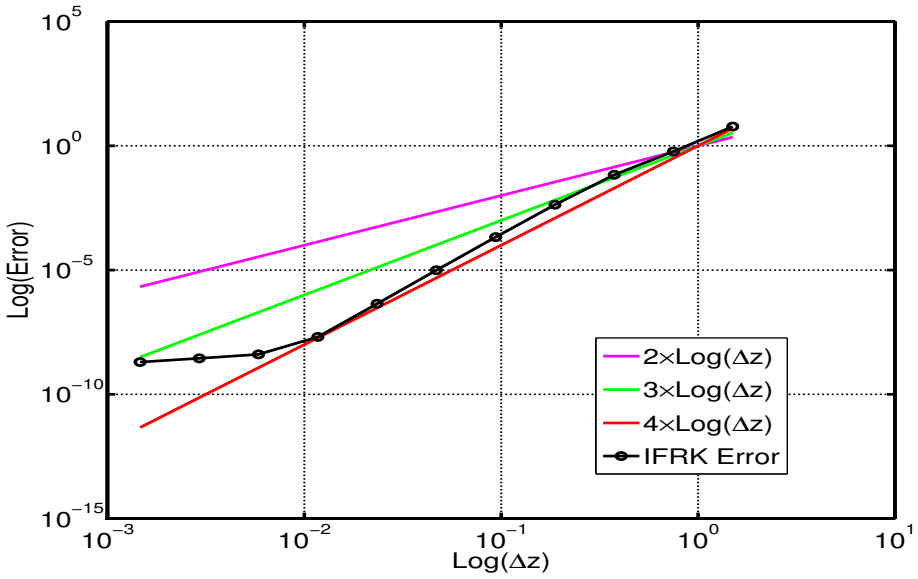


Figure B.1 Plot of the Log of the local error in the IFRK method as a function of the Log of the step size. The t domain is taken to be 40 dimensionless units wide and represented by 1024 computational points, yielding a time spacing of $\Delta t = 0.0195$.

B.2.1 Integral Representations

The fastest and most direct method for computing the parameters of a noisy soliton is to use integral approximations for each parameter. For the frequency, timing, and phase parameters, these are normalized moment equations, which weight the possible values of a given parameter, e.g., the phase, by the noisy soliton intensity [72, 74]. Alternatively, for the hyperbolic secant soliton of the NSLE, the amplitude parameter shares a relationship with the energy that can exploited to calculate its value [72, 74].

The pulse can be written as

$$u(t) = u_{\text{nsol}}(t; A, \Omega, T, \Phi) + R(t) \tag{B.8a}$$

represent the noisy soliton, where

$$u_{\text{nsol}}(t; A, \Omega, T, \Phi) = A\,\text{sech}(A(t-T))\,\exp(i\Omega t + i\Phi) \tag{B.8b}$$

is the numerical soliton and $R(t)$ is the radiative portion of the pulse containing all the noise after the changes to the soliton parameters have been projected out. The amplitude parameter can be approximated by calculating the energy,

$$\begin{aligned}\int \left|u(t)\right|^2 dt = A^2 \int (\,\mathrm{sech}(A(t-T)))^2\, dt + \int \left|R(t)\right|^2 dt \\ + 2\mathrm{Re}\left[\int A\,\mathrm{sech}(A(t-T))\,\exp(-i(\Omega t+\Phi))\,R(t)\,dt\right] \approx 2A,\end{aligned} \tag{B.9a}$$

where it is assumed that

$$\mathrm{Re}\left[\int A\,\mathrm{sech}(A(t-T))\,\exp(-i(\Omega t+\Phi))\,R(t)\,dt\right] \tag{B.9b}$$

can be neglected due to its mean zero contribution and

$$\int \left|R(t)\right|^2 dt \tag{B.9c}$$

can be neglected due to the small noise strength assumption. Thus,

$$A \approx \frac{1}{2}\int \left|u(t)\right|^2 dt, \tag{B.9d}$$

is the approximation used for the amplitude parameter, which on average, results in the amplitude being over estimated [45].

The remaining three parameters can be approximated by weighted moment equations of the form

$$\Omega \approx \frac{\int \omega \left|\hat{u}(\omega)\right|^2 d\omega}{\int \left|\hat{u}(\omega)\right|^2 d\omega}, \tag{B.10a}$$

$$T \approx \frac{\int t \left|u(t)\right|^2 dt}{\int \left|u(t)\right|^2 dt}, \tag{B.10b}$$

$$\Phi \approx \frac{\int \arctan\left(\frac{\mathrm{Im}[u]}{\mathrm{Re}[u]}\right) \left|u(t)\right|^2 dt}{\int \left|u(t)\right|^2 dt}, \tag{B.10c}$$

where $\hat{u}$ is the Fourier transformed version of u. Finally, it is noted that since the bandwidth of the radiation is larger than the pulse, the noisy pulse is often filtered prior to calculating the approximations above. To accomplish this, an estimate for the frequency value is first found through the moment equation

$$\Omega_{\text{est}} = \frac{\int \omega \left|\hat{u}(\omega)\right|^2 d\omega}{\int \left|\hat{u}(\omega)\right|^2 d\omega}, \tag{B.11a}$$

which is then used to construct the Gaussian filtered pulse

$$u_{\text{fil}}(t) = \mathcal{F}^{-1}\left[\exp\left(-\frac{(\omega-\Omega_{\text{est}})^2}{2W_{\text{fil}}}\right)\hat{u}(\omega)\right]. \tag{B.11b}$$

B.2.2 Projection Method

Recall that as a result of SPT, the stochastic perturbations to each parameter can be represented by (linear combinations of) projections of the form

$$\Delta X_k = \sigma \text{Re}\left[\int \bar{v}_X^\dagger(kz_{\text{a}},t)\exp(-i\Theta(kz_{\text{a}}))\,n_k(t)\,dt\right], \tag{B.12}$$

where k indexes the amplifier, $v_X^\dagger$ are the discrete adjoint eigenfunctions and n_k is the noise. Because of this, perturbations in the parameters (and by extension the pulse) are seen to originate from the portion of noise that projects onto the basis of these eigenfunctions, and more importantly, any noise that is orthogonal to this basis has no instantaneous effect on the pulse. Thus, the exact underlying soliton must be orthogonal to the radiation (defined as the difference between the soliton and noisy pulse) is must be orthogonal to each eigenfunction $v_X^\dagger$. From this, an iterative method can be constructed, the steps of which are given as:

1. Given the noisy pulse u, calculate approximations for the soliton parameters from the moment Equations (B.9d), (B.10a), (B.10b) and (B.10c).

2. Use these approximations to construct an approximation for the soliton, u_{sol}, and approximations for the adjoint eigenvectors $v_X^\dagger$.

3. Use u_{sol} to construct an approximation for the radiation through $R_{\text{app}} = u - u_{\text{sol}}$.

4. Calculate any changes in the parameters through the projections

$$\Delta X = \text{Re}\left[\int \bar{v}_X^{\dagger} \exp(-i\Theta)\, R(t)\, dt\right], \tag{B.13}$$

   and update the parameter values.

5. Calculate the error, e.g., $\max_X \Delta X$.

6. Return to step 2 and iterate until the error falls below a predefined value.

Finally, note that since this iteration method is only dependent on the construction of a basis of discrete eigenfunctions, it can be easily extended to the case of DMNLSE. Of course, it such settings, the pulse formation step is more complicated by the absence of a close form solution.

B.2.3 IST Reconstruction of the Numerical Soliton

The final method for recovering the pulse parameters is based on a discretized version of the trace formula from IST [57]. This allows one to reconstruct the pulse from the eigenvalues ζ and eigenfunctions (Jost functions) v_1 and v_2 of the forward scatter problem presented by Zakharov [46]

$$\begin{aligned} i\frac{\partial v_1}{\partial t} - iuv_2 &= \zeta v_1, \\ -i\frac{\partial v_2}{\partial t} - i\bar{u}v_1 &= \zeta v_2, \end{aligned} \tag{B.14}$$

where u (referred to as the potential) is the noisy pulse at some point along fiber. There exists various methods for solving Equations (B.14), the one used here was derived by Weideman and Herbst [91]. Assuming that the numerical object contains a single noisy soliton, the spectrum will consist of a continuum along the real axis corresponding to the radiation and one eigenvalue away from the axis corresponding

to the soliton, which is denoted ζ^*, with corresponding Jost functions of $v_1(t;\zeta^*)$ and $v_2(t;\zeta^*)$. After these eigenvalues and eigenfunctions are known, the soliton portion of the numerical pulse, u_{nsol}, can be exactly represented using the discrete part of the trace formula [57]

$$u_{\text{nsol}} = -\frac{v_1^2(t;\zeta^*)}{\int v_2(t;\zeta^*)v_1(t;\zeta^*)\,dt} - \frac{\bar{v}_2^2(t;\zeta^*)}{\int \bar{v}_2(t;\zeta^*)\bar{v}_1(t;\zeta^*)\,dt} \tag{B.15}$$

It is important to note, however, that this formula is *only* valid for soliton solutions of the NLSE, since it relies on integrability which is a property that is not shared by any of the varying coefficient equations, such as NLSE+DM, or the averaged equations, such as the DMNLSE.

REFERENCES

[1] G. Agrawal, *Nonlinear Fiber Optics*. Optics and Photonics, San Diego, California: Academic Press, 2006.

[2] R. Paschotta, *Encyclopedia of Laser Physics and Technology*, vol. 1. Hoboken, New Jersey: Wiley, 2008.

[3] R. O. Moore, G. Biondini, and W. L. Kath, "Importance sampling for noise-induced amplitude and timing jitter in soliton transmission systems," *Optics Letters*, vol. 28, pp. 105–107, 2003.

[4] R. Holzlöhner, V. S. Grigoryan, C. R. Menyuk, and W. L. Kath, "Accurate Calculation of Eye Diagrams and Bit Error Rates in Optical Transmission Systems Using Linearization," *Journal of Lightwave Technology*, vol. 20, p. 389, Mar. 2002.

[5] Y. Yadin, M. Shtaif, and M. Orenstein, "Bit-error rate of optical DPSK in fiber systems by multicanonical Monte Carlo simulations," *IEEE Photonics Technology Letters*, vol. 17, pp. 1355–1357, Jun. 2005.

[6] R. O. Moore, T. Schfer, and C. K. Jones, "Soliton broadening under random dispersion fluctuations: Importance sampling based on low-dimensional reductions," *Optics Communications*, vol. 256, no. 46, pp. 439 – 450, 2005.

[7] E. Spiller, W. Kath, R. Moore, and C. McKinstrie, "Computing large signal distortions and bit-error ratios in DPSK transmission systems," *IEEE Photonics Technology Letters*, vol. 17, pp. 1022 –1024, May 2005.

[8] A. Tonello, S. Wabnitz, I. Gabitov, and R. Indik, "Importance sampling of Gordon-Mollenauer soliton phase noise in optical fibers," *IEEE Photonics Technology Letters*, vol. 18, pp. 886 –888, Apr. 2006.

[9] K. Chepuri and T. H. de Mello, "Solving the vehicle routing problem with stochastic demands using the cross entropy method," *Annals of Operations Research*, vol. 134, pp. 153–181, 2005.

[10] M. Onorato, A. R. Osborne, M. Serio, and S. Bertone, "Freak waves in random oceanic sea states," *Physical Review Letters*, vol. 86, pp. 5831–5834, Jun. 2001.

[11] D. R. Solli, C. Ropers, P. Koonath, and B. Jalali1, "Optical rogue waves," *Nature*, vol. 450, pp. 1054–1057, 2007.

[12] W. J. Morokoff, "An importance sampling method for portfolios of credit risky assets," in *Proceedings of the 36th conference on Winter simulation*, WSC '04, pp. 1668–1676, Winter Simulation Conference, 2004.

[13] T. Driscoll and K. Maki, "Searching for rare growth factors using multicanonical Monte Carlo methods," *SIAM Review*, vol. 49, no. 4, pp. 673–692, 2007.

[14] R. Y. Rubinstein, *Simulation and the Monte Carlo method*. Wiley series in probability and mathematical statistics. Probability and mathematical statistics, Hoboken, New Jersey: Wiley, 1981.

[15] P. Smith, M. Shafi, and H. Gao, "Quick simulation: a review of importance sampling techniques in communications systems," *IEEE Journal on Selected Areas in Communications*, vol. 15, pp. 597 –613, May 1997.

[16] J. Chan and D. Kroese, "Rare-event probability estimation with conditional Monte Carlo," *Annals of Operations Research*, vol. 189, pp. 43–61, 2011. 10.1007/s10479-009-0539-y.

[17] J. Bucklew, *Introduction to Rare Event Simulation*. Springer Series in Statistics, New York, New York: Springer, 2004.

[18] P. Bratley, B. Fox, and L. Schrage, *A Guide to Simulation*. New York, New York: Springer-Verlag, 1983.

[19] G. S. Fishman, *Monte Carlo: Concepts, Algorithms, and Applications*. Springer Series in Operations Research, New York, New York: Springer-Verlag, 1996.

[20] M. C. Jeruchim, P. Balaban, and K. Shanmugan, *Simulation of Communication Systems*. Applications of communications theory, New York, New York: Plenum, 1992.

[21] D. Landau and K. Binder, *A Guide To Monte Carlo Simulations In Statistical Physics*. Cambridge, England: Cambridge University Press, 2005.

[22] M. Denny, "Introduction to importance sampling in rare-event simulations," *European Journal of Physics*, vol. 22, no. 4, p. 403, 2001.

[23] R. O. Moore, G. Biondini, and W. L. Kath, "A method to compute statistics of large, noise-induced perturbations of nonlinear Schrödinger solitons," *SIAM Review*, vol. 50, no. 3, pp. 523–549, 2008.

[24] G. Biondini and W. Kath, "Polarization-mode dispersion emulation with Maxwellian lengths and importance sampling," *IEEE Photonics Technology Letters*, vol. 16, pp. 789 –791, Mar. 2004.

[25] G. Biondini, W. Kath, and C. Menyuk, "Importance sampling for polarization-mode dispersion," *IEEE Photonics Technology Letters*, vol. 14, pp. 310 –312, Mar. 2002.

[26] G. Biondini, W. Kath, and C. Menyuk, "Importance sampling for polarization-mode dispersion: techniques and applications," *Journal of Lightwave Technology*, vol. 22, pp. 1201 – 1215, April 2004.

[27] S. Fogal, G. Biondini, and W. Kath, "Multiple importance sampling for first- and second-order polarization-mode dispersion," *IEEE Photonics Technology Letters*, vol. 14, pp. 1273 –1275, Sep. 2002.

[28] R. O. Moore, G. Biondini, and W. L. Kath, "A method to compute statistics of large, noise-induced perturbations of nonlinear Schrödinger solitons," *SIAM Journal on Applied Mathematics*, vol. 67, no. 5, pp. 1418–1439, 2007.

[29] E. Iannone, F. Matera, A. Mecozzi, and M. Settembre, *Nonlinear Optical Communication Networks*. Hoboken, New Jersey: John Wiley, 1998.

[30] P. C. Becker, N. A. Olsson, and J. R. Simpson, *Erbium-doped fiber amplifiers fundamentals and technology*. San Diego, California: Academic Press, 1999.

[31] A. Hasegawa and F. Tappert, "Transmission of stationary nonlinear optical pulses in dispersive dielectric fibers i: Anomalous dispersion," *Applied Physics Letters*, vol. 23, no. 3, pp. 142–144, 1973.

[32] A. Hasegawa and F. Tappert, "Transmission of stationary nonlinear optical pulses in dispersive dielectric fibers ii: Normal dispersion," *Applied Physics Letters*, vol. 23, no. 4, pp. 171–172, 1973.

[33] A. C. Newell and J. V. Moloney, *Nonlinear Optics*. Addison-Wesley, 1992.

[34] A. A. Hardy and R. Oron, "Amplified spontaneous emission and Rayleigh backscattering in strongly pumped fiber amplifiers," *Journal of Lightwave Technology*, vol. 16, p. 1865, Oct. 1998.

[35] E. Desurvire and J. Simpson, "Amplification of spontaneous emission in erbium-doped single-mode fibers," *Journal of Lightwave Technology*, vol. 7, pp. 835 –845, May 1989.

[36] M. Holmes, *Introduction to Perturbation Methods*. Texts in Applied Mathematics, New York, New York: Springer-Verlag, 1995.

[37] J. P. Gordon, W. H. Louisell, and L. R. Walker, "Quantum fluctuations and noise in parametric processes ii," *Physical Review*, vol. 129, pp. 481–485, Jan. 1963.

[38] C. Giles and E. Desurvire, "Modeling erbium-doped fiber amplifiers," *Journal of Lightwave Technology*, vol. 9, pp. 271 –283, Feb. 1991.

[39] J. P. Gordon and H. A. Haus, "Random walk of coherently amplified solitons in optical fiber transmission," *Optics Letters*, vol. 11, no. 10, pp. 665–667, 1986.

[40] J. P. Gordon and L. F. Mollenauer, "Phase noise in photonic communications systems using linear amplifiers," *Optics Letters*, vol. 15, no. 23, pp. 1351–1353, 1990.

[41] A. Hasegawa and M. Matsumoto, *Optical Solitons in Fiber*. Springer, 1990.

[42] E. Fredholm, "Sur une classe d'equations fonctionnelles," *Acta Math.*, vol. 27, p. 365390, 1903.

[43] A. Owen and Y. Zhou, "Safe and effective importance sampling," *Journal of the American Statistical Association*, vol. 95, no. 449, pp. 135–143, 2000.

[44] E. Veach, *Robust Monte Carlo Methods for Light Transport Simulation*. PhD thesis, Stanford University, 1997.

[45] J. Li and W. L. Kath, "Extracting solitons from noisy pulses," *SIAM Journal on Applied Mathematics*, vol. 72, no. 2, pp. 577–593, 2012.

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

[47] B. Fornberg and T. A. Driscoll, "A fast spectral algorithm for nonlinear wave equations with linear dispersion," *Journal of Computational Physics*, vol. 155, pp. 457–467, 1999.

[48] T. Y. Hou, J. S. Lowengrub, and M. J. Shelley, "Removing the stiffness from interfacial flows with surface tension," *Journal of Computational Physics*, vol. 114, no. 2, pp. 312 – 338, 1994.

[49] C. McKinstrie and H. Kogelnik, "Vector modulation interaction in a rapidly spun fiber," *Selected Topics in Quantum Electronics, IEEE Journal of*, vol. 18, pp. 726 –737, Mar. 2012.

[50] M. J. Ablowitz and G. Biondini, "Multiscale pulse dynamics in communication systems with strong dispersion management," *Optics Letters*, vol. 23, no. 21, pp. 1668–1670, 1998.

[51] E. T. Spiller and G. Biondini, "Phase noise of dispersion-managed solitons," *Physical Review A*, vol. 4, pp. 011805–1 – 011805–4, 2009.

[52] E. Spiller and G. Biondini, "Importance sampling for dispersion-managed solitons," *SIAM Journal on Applied Dynamical Systems*, vol. 9, no. 2, pp. 432–461, 2010.

[53] Y. Chen, F. X. Kärtner, U. Morgner, S. H. Cho, H. A. Haus, E. P. Ippen, and J. G. Fujimoto, "Dispersion-managed mode locking," *Journal of the Optical Society of America B*, vol. 16, pp. 1999–2004, Nov. 1999.

[54] A. I. Korytin, A. Y. Kryachko, and A. M. Sergeev, "Dissipative solitons in the complex Ginzburg-Landau equation for femtosecond lasers," *Radiophysics and Quantum Electronics*, vol. 44, pp. 428–442, 2001. 10.1023/A:1017949230737.

[55] G. Biondini, "The DMGLE and its application to femtosecond lasers," *Nonlinearity*, vol. 21, p. 28492870, 2008.

[56] D. S. Cargill, R. O. Moore, and C. J. McKinstrie, "Noise bandwidth dependence of soliton phase in simulations of stochastic nonlinear Schrödinger equations," *Optics Letters*, vol. 36, pp. 1659–1661, May 2011.

[57] M. J. Ablowitz and H. Segur, *Solitons and the Inverse Scattering Transform*. SIAM, 1981.

[58] D. Anderson, "Variational approach to nonlinear pulse propagation in optical fibers," *Physical Review A*, vol. 27, pp. 3135–3145, 1983.

[59] C. J. McKinstrie, C. Xie, and C. Xu, "Effects of cross-phase modulation on phase jitter in soliton systems with constant dispersion," *Optics Letters*, vol. 28, no. 8, pp. 604–606, 2003.

[60] B. G. Bale and J. N. Kutz, "Variational method for mode-locked lasers," *Journal of the Optical Society of America B*, vol. 25, pp. 1193–1202, Jul. 2008.

[61] J. Elgin, "Purterbations of optical solitons," *Physical Review A*, vol. 47, pp. 4331–4341, 1993.

[62] T. P. Horikis and J. N. Elgin, "Soliton radiation in an optical fiber," *Journal of the Optical Society of America B*, vol. 18, no. 7, pp. 913–918, 2001.

[63] J. Yang, "Complete eigenfunctions of linearized integrable equations expanded around a soliton solution," *Journal of Mathematical Physics*, vol. 41, no. 9, pp. 6614–6638, 2000.

[64] H. Goldstein, *Classical Mechanics*. Addison-Wesley series in physics, Boston, Massachusetts: Addison-Wesley, 1980.

[65] D. Kaup and T. Vogel, "Quantitative measurement of variational approximations," *Physics Letters A*, vol. 362, no. 4, pp. 289 – 297, 2007.

[66] D. J. Kaup, "A perturbation expansion for the Zakharov-Shabat inverse scattering transform," *SIAM Journal on Applied Mathematics*, vol. 31, no. 1, pp. pp. 121–133, 1976.

[67] W. Kath and N. Smyth, "Soliton evloution and radiation loss for the nonlinear shrodinger equation," *Physical Review E*, vol. 51, no. 4, pp. 1484–1492, 1995.

[68] M. Suzuki, I. Morita, N. Edagawa, S. Yamamoto, H. Taga, and S. Akiba, "Reduction of Gordon-Haus timing jitter by periodic dispersion compensation in soliton transmission," *Electronics Letters*, vol. 37, 1995.

[69] L. F. Mollenanuer and J. P. Gordon, *Solitons in Optical Fibers*. Amsterdam, Netherlands: ELSEVIER, 2006.

[70] J. Li, E. Spiller, and G. Biondini, "Noise-induced perturbations of dispersion-managed solitons," *Physical Review A*, vol. 75, pp. 053805– 053818, 2007.

[71] I. R. Gabitov and S. K. Turitsyn, "Averaged pulse dynamics in a cascaded transmission system with passive dispersion compensation," *Optics Letters*, vol. 21, pp. 327–329, Mar. 1996.

[72] D. J. Kaup and A. C. Newell, "Solitons as particles, oscillators, and in slowly changing media: A singular perturbation theory," *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, vol. 361, no. 1707, pp. 413–446, 1978.

[73] A. M. Kosevich, "Particle and wave properties of solitons: Resonant and non-resonant soliton scattering by impurities," *Physica D: Nonlinear Phenomena*, vol. 41, no. 2, pp. 253 – 261, 1990.

[74] W. L. Kath, "A modified conservation law for the phase of a nonlinear Schrödinger soliton," *Methods and Applications of Applied Analysis*, vol. 4, pp. 141–155, 1997.

[75] E. Noether, "Invariant variation problems," *Transport Theory and Statistical Physics*, vol. 1, no. 3, pp. 186–207, 1971.

[76] H. Haus, "Mode-locking of lasers," *IEEE Journal On Selected Topics In Quantum Electronics*, vol. 6, 2000.

[77] J. N. Kutz, "Mode-locked soliton lasers," *SIAM Review*, vol. 48, pp. 629–678, 2006.

[78] F. Ilday, F. Wise, and F. Kaertner, "Possibility of self-similar pulse evolution in a Ti:sapphire laser," *Optics Express*, vol. 12, pp. 2731–2738, Jun. 2004.

[79] F. O. Ilday, J. R. Buckley, W. G. Clark, and F. W. Wise, "Self-similar evolution of parabolic pulses in a laser," *Physical Review Letters*, vol. 92, p. 213902, May 2004.

[80] W. H. Renninger, A. Chong, and F. W. Wise, "Giant-chirp oscillators for short-pulse fiber amplifiers," *Optics Letters*, vol. 33, pp. 3025–3027, Dec. 2008.

[81] S. T. Cundiff, "Phase stabilization of ultrashort optical pulses," *Journal of Physics D: Applyed Physics*, vol. 35, p. R43R59, 2002.

[82] H. A. Haus and Y. Lai, "Quantum theory of soliton squeezing: A linearized approach," *Journal of the Optical Society of America B*, vol. 7, no. 3, pp. 386–392, 1990.

[83] H. Haus, J. Fujimoto, and E. Ippen, "Analytic theory of additive pulse and Kerr lens mode locking," *IEEE Journal Of Quantum Electronics*, vol. 28, pp. 2086–2096, 1992.

[84] N. N. Akhmediev, V. V. Afanasjev, and J. M. Soto-Crespo, "Singularities and special soliton solutions of the cubic-quintic complex Ginzburg-Landau equation," *Physical Review E*, vol. 53, pp. 1190–1201, Jan. 1996.

[85] M. J. Ablowitz, T. P. Horikis, and B. Ilan, "Solitons in dispersion-managed mode-locked lasers," *Physical Review A*, vol. 77, p. 033814, Mar. 2008.

[86] S. Wieczorek and W. W. Chow, "Chaos in practically isolated microcavity lasers," *Physical Review Letters*, vol. 92, p. 213901, May 2004.

[87] G. Donovan, *Rare Event Simulations for Lighwave Systems Using the Cross-Entropy Method*. PhD thesis, Northwestern University, 2008.

[88] J. Jackson, *Classical Electrodynamics*. Hoboken, New Jersey: Wiley, 1999.

[89] A. Hasegawa, Y. Kodama, and A. Maruta, "Recent progress in dispersion-managed soliton transmission technologies," *Optical Fiber Technology*, vol. 3, no. 3, pp. 197 – 213, 1997.

[90] S. C. Reddy and L. N. Trefethen, "Stability of the method of lines," *Numerische Mathematik*, vol. 62, pp. 235–267, 1992. 10.1007/BF01396228.

[91] J. Weideman and B. Herbst, "Finite difference methods for an AKNS eigenproblem," *Mathematics and Computers in Simulation*, vol. 43, pp. 77–88, 1997.