Soliton broadening under random dispersion fluctuations: Importance sampling based on low-dimensional reductions

Richard O. Moore, Tobias Schäfer, Christopher K. R. T. Jones

May 8, 2006

Abstract

We demonstrate that dispersion-managed solitons are less likely to experience critical broadening under the influence of random dispersion fluctuations than are solitons of the integrable nonlinear Schrödinger equation, and that this robustness increases with map strength from the constant-dispersion (integrable) limit to the large-map-strength limit. To achieve this, we exploit a separation of scales in dispersion-managed soliton dynamics to implement an importance-sampled Monte Carlo approach that determines the probability of rare broadening events. This approach reconstructs the tails (i.e., the regions of practical importance) of probability distribution functions with an efficiency that is several orders of magnitude greater than conventional Monte Carlo simulations. We further show that the variational approach with an appropriately scaled ansatz is surprisingly good at capturing the effect of random dispersion on pulse broadening; where it fails, it can still be used to guide very efficient simulation of the original equation.

1 Introduction

The interaction between chromatic dispersion and nonlinearity in optical fiber has been the subject of considerable investigation since the communications industry first began its migration from electronics to optics. The derivation of the nonlinear Schrödinger (NLS) equation [6] and the subsequent demonstration [10] of NLS solitons as effective bit carriers in optical fiber helped fuel this research, and

2

focused particular attention on ways in which the two influences could be manipulated to allow higher carrying capacities in optical fiber.

Problems such as Gordon-Haus jitter and four-wave mixing have rendered the conventional NLS soliton increasingly less useful, especially with the introduction of more dense transmission formats using wave-division-multiplexing (WDM) and time-division-multiplexing (TDM). This has led researchers to experiment with the physical values of dispersion [7] and nonlinearity [15] in the fiber, the most notable success being the development of dispersion management. In its simplest form, this is a piecewise-constant, periodic variation in the dispersion value of the fiber, with high local dispersion and low average dispersion. In TDM systems, dispersion management causes neighboring pulses to oscillate greatly in width, mitigating their interaction through the Kerr nonlinearity. In WDM systems, dispersion management also causes pulses in different frequency channels to travel back and forth through each other during this oscillation, again mitigating four-wave mixing and cross-phase modulation. Moreover, the interaction between dispersion management and nonlinearity has been shown numerically to lead to another form of stable pulse [12], referred to as a dispersion-managed (DM) soliton. Considerable study has been made of these DM solitons to determine their robustness to various types of disturbance.

One important example of these is found in the deviations of a fiber's dispersion value from its nominal or average value. In particular, random fluctuations of measurable magnitude have been shown to be present in optical fiber [4, 9] and these fluctuations can lead to pulse disintegration if strong enough to prevent the soliton from resting in its nonlinear potential [1]. By analyzing the low-order moments obtained from a Gaussian ansatz in the limit of large map strength, a recent study of DM solitons has suggested that stronger map strengths stabilize the pulse width against dispersion fluctuations [5]. It is of practical interest to confirm this robustness in the original partial differential equation (PDE) and to see how it extends into the moderate and weak map strength regimes, to the limiting case of an NLS soliton.

Furthermore, in the absence of forward-error correction, deployable system designs require that bit errors occur very infrequently, given that typical tolerances are as low as one bit error for every 10^9 or 10^{12} transmitted bits. This suggests that the most important feature of a parameter's probability distribution function (pdf) is not its center or, equivalently, its first few moments, but rather its tails. Given the considerable computational effort required to resolve these tails using conventional numerical methods, there is a pressing need for alternative, low-dimensional system representations and for faster and more intelligent numerical

3

approaches. Emphasis should be placed on the latter in particular, since low-dimensional ordinary differential equation (ODE) reductions are often rendered invalid by the large-deviation events that they would otherwise be used to capture.

In this article, we demonstrate numerically how the robustness of solitons to random fluctuations in dispersion increases with increasing dispersion map strength, over the full range from constant-dispersion nonlinear Schrödinger (NLS) solitons to DM solitons with strong dispersion maps. We thereby corroborate and extend the results of Ref. [5], essentially connecting them to the Kepler problem limit studied in Ref. [2]. The key to our approach is our exploitation of a separation of scales in the dynamics of DM solitons; by averaging over the fast dynamics, we obtain slow equations which can be used in a systematic way to obtain the optimal biasing of random dispersion for use in instantiations of the stochastic PDE.

It is important to note that, unlike in previous studies, we confirm robustness in the tails of the distribution, a region inaccessible to conventional numerical methods. Furthermore, we compare the results obtained through simulations of the (DM) NLS equation with those obtained through simulations of a low-dimensional reduction to determine their correspondence. Finally, we demonstrate that a straightforward application of importance sampling can be used to reduce the computational expense of generating pdfs for the pulse width.

1.1 Basic equations and measure

Picosecond pulses propagating in a dispersion-managed fiber evolve according to the dispersion-managed NLS equation, given in dimensionless form by

$$iu_z + \frac{1}{2}d(z)u_{tt} + |u|^2 u = 0. \tag{1}$$

Here, we take z to be propagation distance in dispersion map spans, t to be time relative to a typical pulse width t_0, d to be dispersion relative to a characteristic dispersion strength set by t_0 and the map span, and u to be the electromagnetic field amplitude rescaled to set the nonlinear coefficient to unity. We have chosen this rescaling to allow us to concentrate on the varying component of the dispersion, $d(z)$, while preserving the effect of nonlinearity for the comparison cases.

The dispersion map $d(z)$ consists of a mean component $\bar{d}$ and a two-step map

$d_p(z)$ comprised of the following:

$$d_p(z) = \begin{cases} +\hat{d} & 0 \leq z < 1/4 \\ -\hat{d} & 1/4 \leq z < 3/4 \, , \\ +\hat{d} & 3/4 \leq z < 1 \end{cases} \qquad (2)$$

where $d_p(z+1) = d_p(z)$ for all $z \geq 0$. To implement an idealized model of random dispersion, we add a Gaussian-distributed white noise element to give $d(z) = \bar{d} + d_p(z) + d_r(z)$, with

$$\langle d_r(z)d_r(z') \rangle = D\delta(z - z'). \qquad (3)$$

The strength D of this random process is a measurable physical parameter. Note also that these fluctuations are "frozen" into the fiber and do not change with time. This reflects the fact that while correlation lengths in a fiber's dispersion value are determined by the length scales of deformations in the cylindrical geometry, for instance, correlation times are determined by environmental conditions such as daily changes in temperature. Such changes occur over much slower time scales than the propagation time for information through an optical link, so that the statistical properties of the fluctuations are essentially static.

In broad terms, a pulse can fail to be detected and lead to a bit error as a result of three possible mechanisms: a change in energy, a change in position, or a change in width. We study the third of these by considering an obvious measure of mean-square pulse width,

$$\tau \equiv \frac{1}{P} \int t^2 |u|^2 dt, \qquad (4)$$

that is insensitive to changes in the pulse energy, $P = \int |u|^2 dt$. To measure the effect of random dispersion on τ, we take the difference between $\Delta\tau$ and its reference value in the absence of random dispersion:

$$\Delta\tau = \frac{1}{P} \int t^2 |u|^2 dt - \frac{1}{P_0} \int t^2 |u_0|^2 dt. \qquad (5)$$

Here, $u(z,t)$ and $u_0(z,t)$ are the perturbed and unperturbed pulse envelopes, respectively.

5

1.2 Representation of NLS and DM solitons

If $d(z) = \bar{d}$ is constant, then Eqn. 1 is integrable and supports a four-parameter family of solitons. We are interested in only one of these parameters, so we write

$$u_0(z,t) = \eta \operatorname{sech}(t/T_{\text{sech}}) \exp\left(i\eta^2 z/2 + i\phi\right), \tag{6}$$

where $\eta = \sqrt{\bar{d}}/T_{\text{sech}}$. Equation 4 then gives

$$\tau = \frac{\pi^2 T_{\text{sech}}^2}{12} = \frac{\pi^2 T_{\text{FWHM}}^2}{48[\ln(1+\sqrt{2})]^2}, \tag{7}$$

where T_{FWHM} is the full width at half-maximum. It should be noted that the moment $\Delta\tau$ is insensitive to parameters contained in the phase of Eqn. 6.

When $\hat{d} > 0$, Eqn. 1 is no longer integrable; nevertheless, solutions that are periodic to within a rotating phase have been shown numerically to exist. When $\hat{d} \gg \bar{d}$, these solutions assume a Gaussian core,

$$u_0(z,t) \approx a(z) \exp\left[-t^2/2T_{\text{G}}^2(z) + i\phi(z,t)\right], \tag{8}$$

where the z-dependence of T_{G} reflects the breathing. In this case,

$$\tau \approx \frac{1}{2}T_{\text{G}}^2 = \frac{T_{\text{FWHM}}^2}{8\ln 2}. \tag{9}$$

Between these two limits, DM solitons have no closed functional form; their shapes, and therefore the transformation from T_{FWHM} to τ, can only be obtained numerically [13]. As the dispersion map strength increases from zero, the periodic solution continuously deforms between the two limits discussed above. Figure 1 illustrates this transition from hyperbolic secant to Gaussian with increasing map strength at fixed T_{FWHM}.

1.3 Low-dimensional reduction

In order to obtain a low-dimensional reduction of the NLS and DM soliton dynamics, we insert the following ansatz into the Lagrangian density of Eqn. 1:

$$u(z,t) = a(z)f\left(\frac{t}{T(z)}\right)\exp\left(i\lambda(z) + i\frac{M(z)}{T(z)}t^2\right), \tag{10}$$

6

(a)   (b)
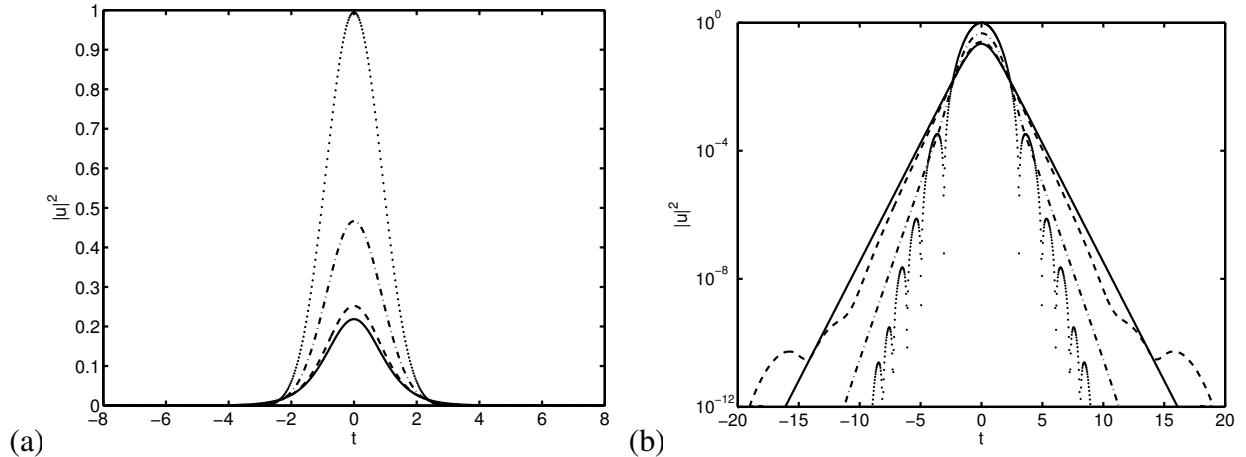
Figure 1: NLS/DM solitons with average dispersion fixed at $\bar{d} = 0.3$, at four differ-
ent map strengths. Solid line is $\hat{d} = 0$ (NLS), dashed line is $\hat{d} = 2$, dashed-dotted
line is $\hat{d} = 6$, dotted line is $\hat{d} = 10$. The widths are fixed at $T_{\text{FWHM}} = 2.06$. Optical
power is plotted against time in normal scale (a) and logarithmic scale (b).

where $T(z) \equiv T_{\text{FWHM}}(z)$ and where $f(x)$ is defined to have unit full-width at half-
maximum and unit L_2 norm. This gives

$$T'(z) = 2d(z)M(z) \quad \text{and} \quad M'(z) = \frac{\Gamma d(z)}{2T^3(z)} - \frac{E\Omega}{4T^2(z)}, \tag{11}$$

with $a^2(z)T(z) = \int |u|^2 dt \equiv E$. The evolution of the phase $\lambda(z)$ is not relevant for
our purposes. The constants are based on integrals of $f(x)$:

$$\Gamma = \frac{\int [f'(x)]^2 dx}{\int x^2 f^2(x) dx} \quad \text{and} \quad \Omega = \frac{\int f^4(x) dx}{\int x^2 f^2(x) dx}. \tag{12}$$

In the limit of zero dispersion management (i.e., the classical NLS equation), these
constants evaluate to

$$\Gamma = \frac{64}{\pi^2} \left(\ln(\sqrt{2}+1) \right)^4 \quad \text{and} \quad \Omega = \frac{32}{\pi^2} \left(\ln(\sqrt{2}+1) \right)^3, \tag{13}$$

while in the opposite limit of large dispersion management, they evaluate to

$$\Gamma = 16(\ln 2)^2 \quad \text{and} \quad \Omega = 8\sqrt{\frac{2}{\pi}}(\ln 2)^{3/2}. \tag{14}$$

For DM solitons with low or moderate dispersion management, these constants
must be evaluated numerically using the periodic pulse form.

7

2   Importance sampling

Importance sampling is a technique wherein the probability distribution functions used to generate the random Monte Carlo samples are biased to make errors occur more frequently than would be the case otherwise [17, 18]. Before we delve into the implementation of importance sampling for random dispersion fluctuations, let us briefly present the basic ideas in a general setting.

Let X denote a collection of random variables (RVs) identifying a particular system realization. A measurable quantity $y(X)$ can be associated with each of these realizations. (In our case, X is a matrix containing the values of a noise process discretized in t, z at all gridpoints along a fiber, and y is a final pulse parameter such as amplitude, timing, or, in this case, width.) Suppose that we are interested in calculating the probability P that $y(X)$ falls in some prescribed range at the end of the fiber line. This probability can be represented as the expectation value of an *indicator function* $I(y(X))$, such that $I(y) = 1$ if the random variable y falls in the prescribed range and $I(y) = 0$ otherwise. That is, the probability P is represented by the multi-dimensional integral,

$$P = \int I(y(x)) p_X(x) dx = E[I(y(X))], \tag{15}$$

where $p_X(x)$ is the joint probability density function (pdf) of the RV matrix X, $E[\cdot]$ denotes the expectation value with respect to $p_X(X)$, and the integral is over all allowed noise configurations. In the vast majority of interesting cases, a direct calculation of the integral in Eq. (15) is impossible. One then often resorts to Monte Carlo simulations to produce an estimator $\hat{P}$ for P, replacing the integral in Eq. (15) with

$$\hat{P}_{\text{mc}} = \frac{1}{M} \sum_{m=1}^{M} I(y(X_m)), \tag{16}$$

where M is the total number of Monte Carlo samples, and each X_m is drawn from $p_X(x)$. Equation (16) simply expresses the relative number of samples falling in the range of interest. If one is interested in low-probability events (where $P \ll 1$), however, an impractically large number of samples is often necessary in order to register even a single event, and an even larger number is required in order to obtain an accurate estimate. In these cases, importance sampling can be used to improve the efficiency of the Monte Carlo approach. We first rewrite the probability P in Eq. (15) as

$$P = \int I(y(X)) r(x) p^*(x) dx = E^*[I(y(X)) r(X)], \tag{17}$$

where $E^*[\cdot]$ denotes the expectation value with respect to the *biased distribution* $p^*(x)$, and where $r(x) = p_X(x)/p^*(x)$ is called the *likelihood ratio* [17]. As before, we then estimate the corresponding integral via Monte Carlo simulations; that is, we write an importance-sampled Monte Carlo estimate for P as

$$\hat{P}_{\text{is}} = \frac{1}{M} \sum_{m=1}^{M} I(y(X_m^*))r(X_m^*), \tag{18}$$

where now the samples X_m^* are drawn according to $p^*(x)$. By design, the estimator $\hat{P}_{\text{is}}$ is unbiased; i.e., $E^*[\hat{P}_{\text{is}}] = P$. Naturally, if $p^*(x) \equiv p_X(x)$ (that is, for unbiased Monte Carlo simulations), we have $r(x) = 1$, and Eq. (18) agrees with Eq. (16). The use of a biasing pdf allows the desired regions of sample space to be visited much more frequently. At the same time, the likelihood ratio automatically adjusts each contribution so that all of the different realizations contribute as appropriate to the final estimate.

The crucial step when applying importance sampling is to determine a biasing distribution $p^*(x)$ that reduces the variance of the estimator $\hat{P}_{\text{is}}$. Use of the *ideal* biasing distribution requires knowledge of the very distribution we are seeking to calculate; other biasing distributions vary in effectiveness depending on how well they approximate this distribution. A particularly simple and intuitive option is to increase the overall noise variance, in an attempt to increase the probability of generating errors. It is well-known, however, that this biasing method (often referred to as *variance scaling*) is only effective in low-dimensional systems [17]. In general, in order for importance sampling to be effective, $p^*(x)$ should concentrate the Monte Carlo samples near the regions that are most likely to generate rare events of interest, which in our case means determining the noise instantiations most likely to produce large width variations at the fiber output. We will use this knowledge to bias the *mean* of each noise distribution, giving $p^*(x) = p(x - b)$, where $b = b(z)$. The problem is thereby reduced to finding the optimal mean biasing b.

If one seeks to reconstruct a broad region of the pdf for the quantity of interest, no single choice of biasing distribution can be expected to capture with high efficiency all the regions of sample space that give rise to the events of interest. In this case, several different biasing distributions $p_q^*(x)$ can be used, and their results combined using a method known as *multiple importance sampling* [14, 19]. With this technique, a weight $w_q(x)$ is associated with each biasing distribution. An importance-sampled estimator for P is then written as

$$\hat{P}_{\text{mis}} = \frac{1}{Q} \sum_{q=1}^{Q} \frac{1}{M_q} \sum_{m=1}^{M_q} w_q(X_{mq}^*)I(y(X_{mq}^*))r_q(X_{mq}^*), \tag{19a}$$

9

where Q is the total number of biasing distributions, M_q is the number of samples drawn from $p_q^*(x)$, X_{mq}^* is the mth such sample and $r_q(x) = p_X(x)/p_q^*(x)$. Several strategies are possible for selecting the weights; the estimator $\hat{P}$ will be unbiased as long as $\sum_{q=1}^{Q} w_q(x) = 1$ for all x. A particularly simple and effective choice is the *balance heuristic* [19]:

$$w_q(x) = \frac{M_q p_q^*(x)}{\sum_{q'=1}^{Q} M_{q'} p_{q'}^*(x)} . \tag{19b}$$

Note that $M_q p_q^*(x)$ corresponds to the expected number of hits from the qth distribution. Thus, the sample weight dictated by the balance heuristic is given by the likelihood of having realized the sample using the qth distribution relative to the likelihood of having realized the sample at all.

2.1 Biasing based on fast dynamics

Determination of the optimal mean biasing begins with the understanding that DM solitons manifest themselves in finite-dimensional form as periodic solutions of Eqns. 11. In this context, an optimal biasing of the noise process $d_r(z)$ means a deterministic function $d_b(z)$ that represents the *most likely* noise instantiation leading to a given width deviation $T(z_f) - T(0) = T_f - T_p$ at the end of a fiber line, denoted by $z = z_f$, where $(T_p, 0)^T$ is the initial condition (in a symmetric map) that gives a periodic solution of Eqns. 11.

Our assumption of simple Gaussian-distributed noise means that the probability of observing a noise instantiation given by $d_b(z)$ is inversely proportional to the squared L_2 norm of $d_b(z)$. In other words, the most likely such noise to produce a prescribed deviation $T_f - T_p$ is that which minimizes

$$||d_b(z)||^2 \equiv \int_0^{z_f} d_b(z)^2 \, dz \tag{20}$$

under the constraint that $d(z) = \bar{d} + d_p(z) + d_b(z)$ satisfies what is now a boundary value problem (BVP) given by Eqns. 11 with initial condition $(T_p, 0)^T$ and terminal condition $T(z_f) = T_f$. This is a nonlinear optimal control problem which, given a sufficiently small target deviation $T_f - T_p$ from the periodic solution, could be approximated well by linearizing about the periodic solution. Our interest, however, is in determining the optimal biasing for *large* (i.e., $O(1)$) deviations. We therefore resort to an iterative approach using successive linearizations. The jth iteration begins with a linearization about the solution $(T_{j-1}, M_{j-1})^T$ obtained by

numerically solving Eqns. 11 with $d(z) = \bar{d} + d_p(z) + d_{j-1}(z)$, where d_{j-1} is the linear optimization result from the previous iteration:

$$\begin{pmatrix} T \\ M \end{pmatrix} = \begin{pmatrix} T_{j-1}(z) \\ M_{j-1}(z) \end{pmatrix} + \begin{pmatrix} t_j(z) \\ m_j(z) \end{pmatrix}, \tag{21}$$

The dispersion is then written as $d(z) = \bar{d} + d_p(z) + d_{j-1}(z) + \Delta d_j$, where $\Delta d_j = d_j(z) - d_{j-1}(z)$. The nonlinear control problem is thus approximated by the optimization (minimization) of $||d_j(z)||^2$ under the constraint that the following linear BVP has a solution:

$$L\begin{pmatrix} t_j \\ m_j \end{pmatrix} = \vec{f}_{j-1}\Delta d_j \equiv \begin{pmatrix} 2M_{j-1} \\ \Gamma/2T_{j-1}^3 \end{pmatrix}\Delta d_j, \tag{22}$$

where

$$\begin{pmatrix} t_j \\ m_j \end{pmatrix}\bigg|_{z=0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} t_j \\ m_j \end{pmatrix}\bigg|_{z=z_f} = \begin{pmatrix} T_f - T_{j-1}(z_f) \\ \tilde{m} \end{pmatrix}, \tag{23}$$

with $\tilde{m}$ unspecified and with linear operator

$$L = \begin{pmatrix} d/dz & -2\left(\bar{d} + d_p(z) + d_{j-1}(z)\right) \\ \frac{3\Gamma}{2T_{j-1}^4(z)}\left(\bar{d} + d_p(z) + d_{j-1}(z)\right) - \frac{E\Omega}{2T_0^3(z)} & d/dz \end{pmatrix}. \tag{24}$$

The solution to this Lagrange multiplier problem is

$$d_j(z) = \frac{\vec{v}^T(z;z_f)\vec{f}_{j-1}(z)}{||\vec{v}^T(s;z_f)\vec{f}_{j-1}(s)||^2}\left(T_f - T_{j-1}(z_f) + \int_0^{z_f} \vec{v}^T(s;z_f)\vec{f}_{j-1}(s)d_{j-1}(s)\,ds\right), \tag{25}$$

where $\vec{v}(z;\zeta)$ is the Green's function for the nonhomogeneous linear problem above, satisfying

$$L^\dagger\vec{v}(z;\zeta) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}\delta(z-\zeta) \tag{26}$$

with boundary condition $\vec{v}(z_f;\zeta) = (0,0)^T$.

This procedure, which can be regarded as a Newton-Raphson iteration, is initialized with the periodic solution for $(T_0(z), M_0(z))^T$ and $d_0(z) \equiv 0$, and terminates at $j = k$ with $d_b(z) \approx d_k(z)$ when $|T_k(z_f) - T_f|$ is suitably small. We must naturally be concerned about the convergence properties of this procedure, and about the risk of finding a local rather than global optimizer. The scope of this paper is to discuss the application of importance sampling to the case at hand, however, so we will not address these issues here. We simply comment that the method converged to a reasonable solution in all cases contained in this article.

11

2.2 Biasing based on slow dynamics

Since one of our goals is to deduce simple biasing rules for the application of importance sampling to critical broadening events in solitons, we note that we can avoid the above numerical optimization routine by exploiting a disparity in the time scales associated with the linear breathing of DM solitons and with their nonlinearity-induced phase rotations. Eqns. 11 can be thus be treated perturbatively using a multiple scales approach to reflect the fact that the dominant pulse dynamics lives on a torus characterized by the two time scales. The fast time scale produces a coupled nonlinear system that captures the linear evolution of Eqn. 1:

$$T' \;=\; 2d_{\mathrm{p}}(z)M \qquad \text{and} \tag{27}$$

$$M' \;=\; \frac{\Gamma d_{\mathrm{p}}(z)}{2T^3}, \tag{28}$$

with integration constants $\mathcal{T}$ and $\mathcal{M}$ representing the values of T and M, respectively, at the beginning of each symmetric dispersion map segment. These equations integrate to produce closed, periodic trajectories. The equations for the slow nonlinear envelope are

$$\mathcal{T}' \;=\; 2D(\mathcal{T},\mathcal{M})\mathcal{M} + K(\mathcal{T},\mathcal{M})\mathcal{T} \quad \text{and} \tag{29}$$

$$\mathcal{M}' \;=\; \frac{D(\mathcal{T},\mathcal{M})\Gamma}{2\mathcal{T}^3} - K(\mathcal{T},\mathcal{M})\mathcal{T}, \tag{30}$$

with

$$D(\mathcal{T},\mathcal{M}) \;=\; \bar{d} + d_r(z) + \frac{E\Omega F(\mathcal{T},\mathcal{M})}{\hat{d}\Gamma^{3/2}}\left(\frac{1}{\mathcal{T}^2} + \frac{4\mathcal{M}^2}{\Gamma}\right)^{-3/2}, \tag{31}$$

$$K(\mathcal{T},\mathcal{M}) \;=\; \frac{2E\Omega}{\hat{d}\Gamma}\left(\frac{1}{\mathcal{T}^2} + \frac{4\mathcal{M}^2}{\Gamma}\right)^{-1/2}\left(\frac{1}{\sqrt{4+s_-^2}} - \frac{1}{\sqrt{4+s_+^2}}\right), \tag{32}$$

and

$$F(\mathcal{T},\mathcal{M}) \;=\; \sinh^{-1}(s_+/2) - \sinh^{-1}(s_-/2) \tag{33}$$

$$+(s_- - s_+)\left(\frac{1}{\sqrt{4+s_-^2}} + \frac{1}{\sqrt{4+s_+^2}}\right), \tag{34}$$

where

$$s_{\pm} = \frac{4}{\sqrt{\Gamma}} \mathcal{T} \mathcal{M} \pm \frac{\sqrt{\Gamma} \hat{d}}{2} \left(\frac{1}{\mathcal{T}^2} + \frac{4\mathcal{M}^2}{\Gamma} \right).$$

We refer the reader to Ref. [16] for a more complete derivation of these equations.

The term $d_r(z)$ in Eqn. 31 merits a brief explanation. Formally, this term is the original random term $d_r(z)$ averaged over each unit period of the dispersion map. This produces a piecewise-constant, Gaussian-distributed random variable with zero mean and variance D, i.e., a different random number drawn for each dispersion map to represent the average random dispersion contribution for that map. In practice, however, the $O(1)$ variation of this random term is faster than the slow dynamics described by Eqns. 29 and 30, and it is therefore reasonable to approximate this process by one that is identical to the original process, $d_r(z)$. The fact that the random term enters the slow dynamics through a simple integral $\int^z d_r(s) \, ds$, rather than occurring with a nontrivial kernel arising from the fast dynamics, will be significant in determining the correct biasing to effect importance sampling.

Having determined the slow evolution of the soliton width $\mathcal{T}$ at each return point of the dispersion map, we can now proceed with the optimal biasing calculation. Again, this amounts to a nonlinear control problem for Eqns. 29 and 30 (although the ODEs are autonomous in this case, unlike Eqns. 11). Rather than repeating the numerical procedure used in Sec. 2.1, we seek a simpler biasing rule by linearizing about the fixed point of the slow equations. We will see that even the first order approximation often does a very good job in obtaining the proper biasing toward large width events.

We first note that, in the absence of random dispersion, a fixed point of Eqns. 29 and 30 occurs at $\mathcal{M} = \bar{\mathcal{M}} = 0$ and $\mathcal{T} = \bar{\mathcal{T}}$, where $\bar{\mathcal{T}}$ satisfies

$$\bar{\mathcal{T}}^3 F(\bar{\mathcal{T}}, 0) = -\frac{\Gamma^{3/2} \bar{d}\hat{d}}{E\Omega}. \tag{35}$$

An analysis of this equation shows that at least one such point must exist. This fixed point corresponds to a periodic orbit of Eqns. 11, and therefore to a (DM) soliton. In order to obtain the most probable routes to large deviations of T, and therefore of $\mathcal{T}$, we must find the most likely instantiations of d_r that provide these large deviations. By linearizing, we again reduce the nonlinear system to a simpler optimization problem, with the goal of minimizing $||d_b(z)||^2$ under the constraint

13

that the linear nonhomogeneous BVP given by

$$\text{Ł}\begin{pmatrix} \mathcal{T} - \bar{\mathcal{T}} \\ \mathcal{M} \end{pmatrix} = \begin{pmatrix} 0 \\ \Gamma d_{\text{b}}(z)/2\bar{\mathcal{T}}^3 \end{pmatrix} \tag{36}$$

has a solution with boundary conditions

$$\begin{pmatrix} \mathcal{T} - \bar{\mathcal{T}} \\ \mathcal{M} \end{pmatrix}\Bigg|_{z=0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \mathcal{T} - \bar{\mathcal{T}} \\ \mathcal{M} \end{pmatrix}\Bigg|_{z=z_{\text{f}}} = \begin{pmatrix} (\Delta\mathcal{T})_{\text{targ}} \\ \mathcal{M}_f \end{pmatrix}, \tag{37}$$

where $\mathcal{M}_f$ is left unspecified and

$$\text{Ł}\psi \equiv \begin{pmatrix} d/dz & -g(\bar{\mathcal{T}}) \\ h(\bar{\mathcal{T}}) & d/dz \end{pmatrix} \tag{38}$$

with

$$g(\bar{\mathcal{T}}) = \frac{8E\Omega\bar{\mathcal{T}}}{\Gamma\left(4 + \frac{\Gamma\hat{d}^2}{4\bar{\mathcal{T}}^4}\right)^{3/2}} \quad \text{and} \tag{39}$$

$$h(\bar{\mathcal{T}}) = \frac{3\Gamma\bar{d}}{2\bar{\mathcal{T}}^4} - \frac{E\Omega}{\bar{\mathcal{T}}^3}\frac{\left(4 - \frac{\Gamma\hat{d}^2}{4\bar{\mathcal{T}}^4}\right)}{\left(4 + \frac{\Gamma\hat{d}^2}{4\bar{\mathcal{T}}^4}\right)^{3/2}}. \tag{40}$$

The result of this Lagrange multiplier problem is

$$d_{\text{b}}(z) = \frac{4(\Delta\mathcal{T})_{\text{targ}}\bar{\mathcal{T}}^3}{\Gamma}\sqrt{\frac{h}{gz_{\text{f}}}}\left[1 - \frac{\sin(2\sqrt{gh}\,z_{\text{f}})}{2\sqrt{gh}\,z_{\text{f}}}\right]^{-1}\sin\left[\sqrt{gh}\,(z_{\text{f}} - z)\right].$$

As a final word about this calculation, it should be noted that the functional form of $d_{\text{b}}(z)$ obtained in this section is based solely on the slow dynamics of Eqns. 29 and 30; formally, this d_{b} has no dependence on the fast dynamics of linear Eqns. 27 and 28. That this is the case is simply a reflection of the fact that the random process obtained on the slow scale in Eqn. 31 is obtained as a straightforward integral of the random process on the fast scale. Generically, one must do a similar optimization as the one performed above to obtain the correct variation of d_{b} on the *fast* scale as well as on the slow scale (and, indeed, on every scale as far as the perturbation theory is carried). In this case, however, the optimization is trivial and results in a constant optimal biasing over the fast scale. To see an example where a stochastic parametric driving term must be biased on *both* scales, one could consider a parametrically driven Duffing oscillator, for example.

14

3 Results and discussion

In order to gauge the effectiveness of the different biasing techniques used in the application of importance sampling, to determine the accuracy with which our ODE reduction captures the broadening of solitons due to random dispersion, and to establish the impact of increasing map strength on robustness to random dispersion fluctuations, we have numerically integrated Eqn. 1 using an NLS soliton and three different DM solitons. To ensure a fair comparison among the different solitons as bit carriers in a fixed-rate communication line, the pulses have all been chosen to have a full width at half-maximum of $T_{\mathrm{FWHM}} = 2.06$. Their relative peak values in Fig. 1 clearly show the well-studied "power enhancement" that arises from increased map strength. The pulses are propagated to a distance of $z = 10$, corresponding to ten maps in the dispersion-managed cases. The strength of the random dispersion component is set at $D = 0.01$, and the continuous random process is approximated by a delta-correlated discrete process with standard deviation $\sigma^2 = D/\Delta z$, where Δz is the (fixed) discrete numerical evolution step in a pseudo-spectral 4th-order Runge-Kutta method.

3.1 Comparison of biasing methods

The mean-biasing profiles obtained using the Newton-Raphson iteration applied to the ODEs given by Eqns. 11 and those obtained using the linearized average evolution given by Eqns. 30 agree quite well, as demonstrated by Fig. 2. These profiles represent the prescribed mean-biasing for four different target values of $\Delta\tau$, one of which is $\Delta\tau = 0$, i.e., unbiased Monte Carlo simulations. Predictably, the biasing technique based on average evolution does not capture the fast-scale oscillations seen in the biasing yielded by the full optimization; however, it clearly reproduces the slow behavior which is presumably most important in obtaining the large excursions of $\Delta\tau$ necessary for successful implementation of importance sampling.

These biasing profiles were used in importance-sampled Monte Carlo simulations on the ODE model given by Eqns. 11, and recombined using the balance heuristic described above to form the pdfs depicted in Fig. 3. It is evident from this comparison that the biasing profiles produce pdfs that agree exceedingly well with each other, down to probabilities well below 10^{-12}. Furthermore, the accuracy of these pdfs is confirmed by comparison with unbiased Monte Carlo simulations of Eqns. 11, indicated by the squares in Fig. 3. It is important to note that, despite a reduction of two orders of magnitude in the number of runs, the importance-
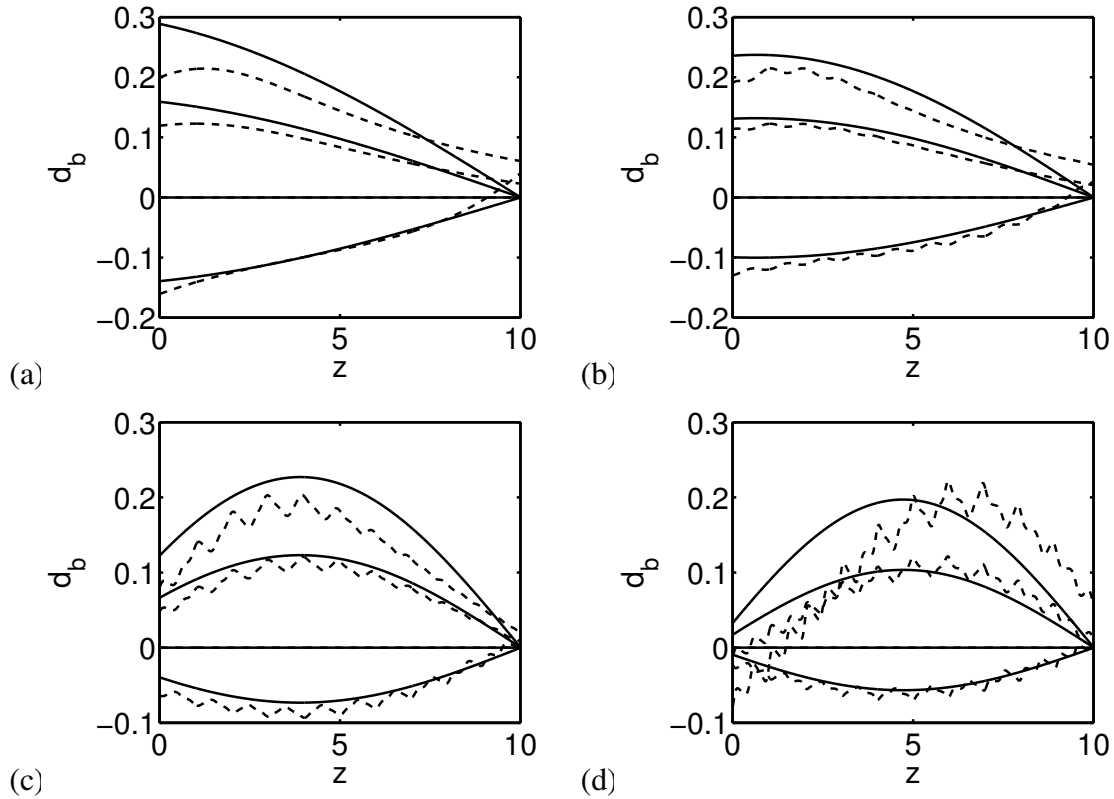
Figure 2: Biasing profiles generated using a Newton-Raphson iteration applied to the fast dynamics (dashed lines) and using the linearized average equations (solid lines), with four different targets of $\Delta\tau$, in the cases (a) $\hat{d} = 0$ (NLS), (b) $\hat{d} = 2$, (c) $\hat{d} = 6$, and (d) $\hat{d} = 10$.

sampled simulations reach several orders of magnitude farther into the tails of the pdfs.

3.2   Validity of ODE reductions

These ODE (Eqns. 11) results are compared with full PDE (Eqn. 1) simulations in Fig. 4. As in the case of the ODE simulations, the PDE simulation results using either of the ODE-based biasing techniques described above compare very well to each other, and to the unbiased PDE simulations. It is interesting to note, however, that while the PDE and ODE simulations agree well in either the zero-map limit (i.e., the case of integrable NLS) or in the large-map-strength limit
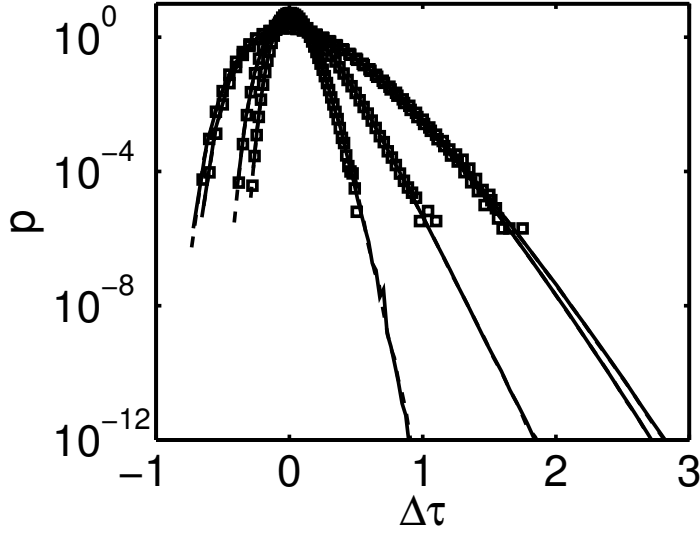
16

Figure 3: Mean-square width pdfs for NLS soliton, soliton with $\hat{d} = 2$, soliton with $\hat{d} = 6$, and soliton with $\hat{d} = 10$ (from broadest to narrowest). Squares represent 10^7 unbiased Monte Carlo simulations, dashed lines represent 4×10^5 biased simulations based on fast dynamics, and solid lines represent 4×10^5 biased simulations based on slow dynamics.

($\hat{d} = 10$), this agreement is particularly bad in the case of small but finite map strength ($\hat{d} = 2$). In this case, the pdfs obtained using the PDE simulations diverge noticeably from the pdf obtained from ODE simulations. Moreover, *even though the ODE reduction is not adequate in capturing the effect of random dispersion directly, it can successfully be used in guiding simulations of the PDE.* Finally, almost all of the curves show relatively poor agreement between the PDE and ODEs when comparing the region over which $\Delta \tau < 0$. While it is not immediately clear why this should be the case, these "narrowing" events are not the events in which we are interested.

3.3 Effect of increasing dispersion map strength

Finally, Fig. 5 demonstrates that increasing dispersion map strength strongly mitigates the effect of random dispersion on soliton broadening. This increased robustness continues from the NLS limit of no dispersion management to the strong-

17

dispersion-map limit. In addition, the importance-sampled PDE simulations afford the observation that this robustness is not only in the first few moments of the distributions, but persists well into the tails, where the critical events that determine bit error ratios are to be found.

The increased robustness to dispersion fluctuations is a fundamental characteristic of the nonlinear oscillator, given in Eqns. 11, that dominates the behavior of NLS and dispersion-managed solitons, as noted in Ref. 5. It is *not* simply a reflection of the well-known power enhancement experienced by solitons when the dispersion map strength is increased. Whereas power enhancement buffers a soliton against external noise, in this case the forcing is parametric, and is therefore unaffected by an increase in soliton "mass". To better illustrate this point, Fig. 5 includes the χ^2 distribution that results from a Gaussian pulse with $T_{\text{FWHM}} = 2.06$ launched in Eqn. 1 with the nonlinear term removed. The pdf has the same form regardless of the amplitude (i.e., mass) of the Gaussian pulse.

4 Conclusion

We have analyzed the effect of random dispersion fluctuations on the mean-square width of solitons of the dispersion-managed nonlinear Schrödinger equation, from the integrable limit to the limit of large dispersion map strength. We have demonstrated that the increased robustness shown in Ref. 5 to hold for low-order moments of the width distribution upon increasing map strength persists well down into the tails of the pdf, which is the most relevant region for the determination of bit error ratios.

In order to study the tails of the pdf, we have applied importance sampling in the generation of biased Monte Carlo simulations, where we have used a reliable ODE reduction (Eqns. 11) of Eqn. 1 to obtain the correct mean-biasing. We have shown that, while the ODEs themselves can often be used to capture the effects of random dispersion, they are sometimes inadequate. In those cases, they can still be used to guide simulations of the PDE so as to increase the efficiency of pdf regeneration by *several orders of magnitude*.

Furthermore, we have used a separation of scales in the low-dimensional dynamics of DM solitons to obtain a simple, closed-form mean-biasing formula which appears to be equally as effective as the true mean-biasing obtained numerically from the ODEs.

18

References

[1] F. Kh. Abdullaev and B. B. Baizakov. Disintegration of a soliton in a dispersion-managed optical communication line with random parameters. *Opt. Lett.*, 25(2):93–95, 2000.

[2] F. Kh. Abdullaev, J. C. Bronski, and G. Papanicolaou. Soliton perturbations and the random Kepler problem. *Physica D*, 135:369–386, 2000.

[3] G. Biondini, W. L. Kath, and C. R. Menyuk. Importance sampling for polarization-mode dispersion. *Photon. Technol. Lett.*, 14:310–312, 2002.

[4] I. Brener, P. P. Mitra, D. D. Lee, D. J. Thomson, and D. L. Philen. High-resolution zero-dispersion wavelength mapping in single-mode fiber. *Opt. Lett.*, 23(19):1520–1522, 1998.

[5] J. Garnier. Stabilization of dispersion-managed solitons in random optical fibers by strong dispersion management. *Opt. Comm.*, 206:411–438, June 2002.

[6] A. Hasegawa and F. Tappert. Transmission of stationary nonlinear optical physics in dispersive dielectric fibers I: Anomalous dispersion. *Appl. Phys. Lett.*, 23(3):142–144, 1973.

[7] C. Lin, H. Kogelnik, and L. G. Cohen. Optical pulse equalization and low dispersion transmission in single-mode fibers in the 1.3–1.7 mm spectral region. *Opt. Lett.*, 5(11):476–478, 1980.

[8] N. Madras. *Lectures on Monte Carlo Methods*. Fields Institute Monographs, American Mathematical Society, Providence, RI, 2002.

[9] L. F. Mollenauer, P. V. Mamyshev, and M. J. Neubelt. Method for facile and accurate measurement of optical fiber dispersion maps. *Opt. Lett.*, 21(21):1724–1726, 1996.

[10] L. F. Mollenauer, R. H. Stolen, and J. P. Gordon. Experimental observation of picosecond pulse narrowing and solitons in optical fibers. *Phys. Rev. Lett.*, 45(13):1095–1098, September 1980.

[11] R. O. Moore, G. Biondini, and W. L. Kath. Importance sampling for noise-induced amplitude and timing jitter in soliton transmission systems. *Opt. Lett.*, 28:105–107, January 2003.

[12] J. H. B. Nijhof, N. J. Doran, W. Forysiak, and A. Berntson. Energy enhancement of dispersion-managed solitons and WDM. *Electr. Lett.*, 34:481–482, 1998.

[13] J. H. B. Nijhof, W. Forysiak, and N. J. Doran. The averaging method for finding exactly periodic dispersion-managed solitons. *IEEE J. Sel. Top. Quantum Electron.*, 6(2):330–336, 2000.

[14] A. Owen and Y. Zhou, Safe and effective importance sampling. *J. Amer. Stat. Assoc.*, 95:135, 2000.

[15] C. Pare, A. Villeneuve, P. A. Belanger, and N. J. Doran. Compensating for dispersion and the nonlinear Kerr effect without phase conjugation. *Opt. Lett.*, 21(7):459–461, April 1996.

[16] T. Schäfer, R. O. Moore, and C. K. R. T. Jones. Pulse propagation in media with deterministic and random dispersion variations. *Opt. Commun.*, 214(1-6):353–362, December 2002.

[17] P. J. Smith, M. Shafi and H. Gao, Quick simulation: A review of importance sampling techniques in communications systems. *IEEE J. Select. Areas Commun.*, 15:597–613, 1997.

[18] R. Srinivasan, *Importance Sampling: Applications in Communications and Detection*. Springer-Verlag, New York, 2002.

[19] E. Veach, *Robust Monte Carlo Methods for Light Transport Simulation*. Ph.D. thesis, Stanford University, 1997.
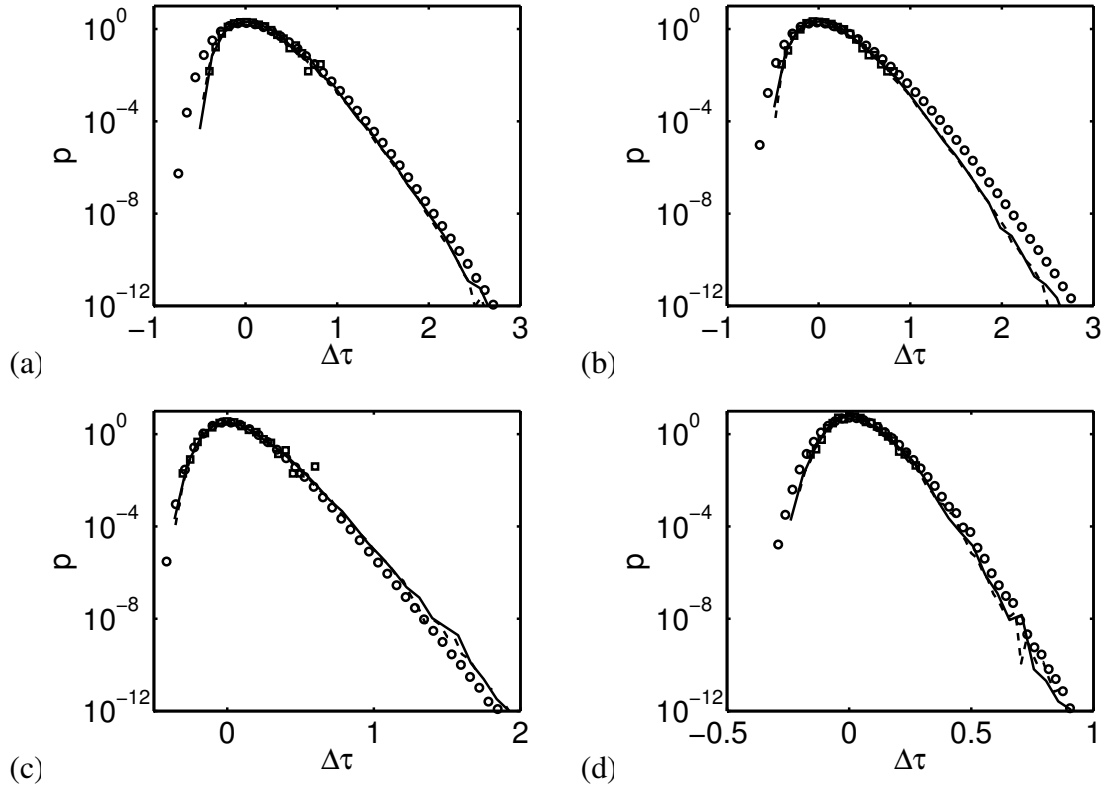
Figure 4: Comparison of pdfs obtained from simulations of Eqn. 1 (PDE) and Eqns. 11 (ODEs). Dashed (solid) lines represent 4×10^3 importance-sampled PDE simulations using mean-biasing based on fast (average) ODE dynamics, circles represent 4×10^5 importance-sampled ODE simulations, and squares represent 1×10^3 unbiased PDE simulations, in the following cases: (a) $\hat{d} = 0$ (NLS), (b) $\hat{d} = 2$, (c) $\hat{d} = 6$, and (d) $\hat{d} = 10$.
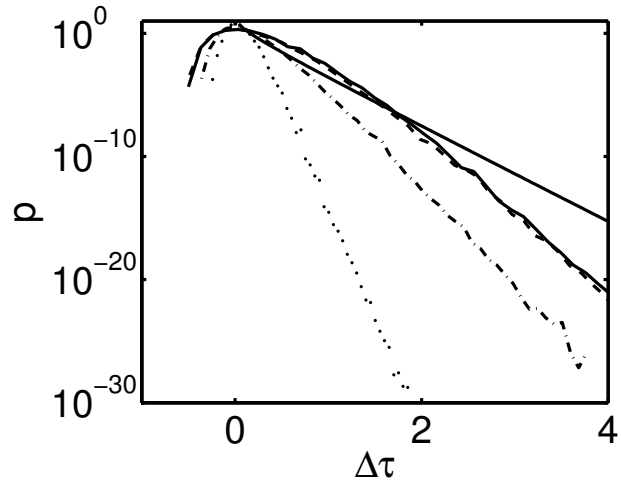
Figure 5: Effect of increasing map strength on soliton broadening caused by random dispersion fluctuations in Eqn. 1. Solid line is $\hat{d} = 0$ (NLS), dashed line is $\hat{d} = 2$, dashed-dotted line is $\hat{d} = 6$, dotted line is $\hat{d} = 10$. The solid straight line shows exact width broadening in the linear Schrödinger equation for a Gaussian pulse with $T_{\text{FWHM}} = 2.06$ and arbitrary amplitude.