SIAM J. Applied Dynamical Systems
Vol. 19, No. 2, pp. 1472–1495

Slip Rates and Slip Modes in an Actively Mode-Locked Laser*

Nathan L. Sanford†, Graham M. Donovan‡, and William L. Kath†

Abstract. Pulses in an actively mode-locked laser can occasionally slip relative to the timing signal, leading to fluctuations in the pulse repetition rate. Such events happen rarely, however, making it infeasible to use traditional methods to determine the slip rate. Here, in a model of a soliton-based mode-locked laser, quantification of the error rate is obtained using importance-sampled Monte Carlo simulations guided by soliton perturbation theory. Position slips are studied in two distinct cases: an overdamped regime where they are primarily direct and an underdamped regime where they typically involve oscillations. Quantification of the slip rate is shown to be more straightforward in the overdamped regime, but dynamic importance sampling is found to be necessary to accurately and efficiently capture error rates. Dynamic importance sampling is shown to be more difficult to implement in the underdamped regime due to the existence of multiple routes by which errors can occur.

Key words. rare event simulation, Monte Carlo, importance sampling, solitons

AMS subject classifications. 35Q51, 35Q55, 65C05, 65C20, 78A40

DOI. 10.1137/19M1297014

1. Introduction.

Mode-locked lasers are a technology with a rich history of application and study [22, 27, 28]. In recent years, the breadth of applications employing such systems has grown dramatically, and simultaneously the performance of the underlying lasers has greatly improved. Optical communications is an oft-cited application [19], but other uses include optical frequency metrology [8], optical clocks [7, 9, 42, 45], spectroscopy [41], generation of high harmonics [25], measurement of fundamental constants [14], and optical storage rings [20].

In many of the above applications, the performance of the system is limited by noise. One convenient method for determining the effects due to noise is via Monte Carlo (MC) simulation. Due to the low error rates present in high-accuracy mode-locked systems, however, traditional MC simulations are intractable, as an unreasonably large number of samples would be needed to determine error probabilities. An extension known as *importance-sampled* Monte Carlo (ISMC) simulation, however, is capable of capturing very low probability events and has been used previously to study optical systems [4, 5, 11, 15, 24, 29, 30, 35, 36, 44].

Importance sampling involves replacing the probability distribution used to draw random samples representing noise during simulation with a *biasing distribution* [37]. The goal is to

*Received by the editors November 15, 2019; accepted for publication (in revised form) by L. Billings March 20, 2020; published electronically June 4, 2020.
https://doi.org/10.1137/19M1297014
Funding: This work was supported in part by the National Science Foundation, DMS-1211912.
†Department of Applied Mathematics, McCormick School of Engineering, Northwestern University, Evanston, IL 60201-3125 (nathansanford2013@u.northwestern.edu, kath@northwestern.edu).
‡Department of Mathematics, University of Auckland, 1142, Auckland, New Zealand (g.donovan@auckland.ac.nz).

pick distributions which give rise to rare events of interest with much higher frequency than they would occur normally, correcting for the biasing so that correct error probabilities can be obtained. In soliton-based systems, soliton perturbation theory and rare event simulation provide a framework in which biasing distributions for nonlinear optical systems can be determined [17, 34]. Soliton perturbation theory allows one to define an ODE system that approximately describes the evolution of soliton parameters in the presence of deterministic and stochastic perturbations of the governing nonlinear Schrödinger (NLS) equation. The theory of rare events then allows one to formulate a constrained optimization problem within this ODE framework whose solutions provide approximate biasing distributions that can be used to guide numerical solutions of the full problem.

Here, ISMC simulations using biasing distributions obtained in this way are used to describe "position slip" errors and determine the rates at which these slips occur in a mode-locked laser model involving an active feedback mode-locking mechanism [10, 32]. The active feedback mechanism can lead to two fundamentally different modes: a nonoscillatory overdamped regime and an oscillatory underdamped locking regime. Position slip errors, or errors of pulse position slippage relative to the mode-locking, are found to occur in both cases but with qualitatively and quantitatively different behaviors. We will show that the importance sampling methods needed to capture error rates in the two cases are somewhat different, and that more sophistication is needed in the underdamped regime to deal with the more complicated paths arising due to the oscillations. We believe that this problem illustrates a general issue that may occur in systems where multiple routes to rare events may be present.

2. Model formulation.

We model an actively mode-locked fiber laser as an optical cavity oscillator including an amplifier, filter, polarization rotator (and polarizer), and phase modulator [10, 32]. This system is represented schematically in Figure 1. Propagation through the optical fiber is described by the dimensionless nonlinear Schrödinger (NLS) equation for the single polarization optical field envelope u [34]:

$$\frac{\partial u}{\partial z} - \frac{i}{2}\frac{\partial^2 u}{\partial t^2} - i|u|^2 u = F(z,t), \tag{2.1}$$

with the right-hand side representing the perturbative effects of the added physical elements used to provide the active mode-locking and stabilize the optical pulses. Here the dimensionless time has been scaled by a characteristic time T_0 (usually comparable to the width of a pulse, here assumed to be on the order of 1 ps), and evolution is with respect to a dimensionless distance z, here scaled by the dispersion length $L = T_0^2/\beta$, where β is the second-order dispersion parameter. The field envelope has also been scaled by a characteristic amplitude [34].

The mode-locking elements are (linear) gain from the amplifier and polarization rotator, filtering, nonlinear gain (or loss) from the polarization rotator, and phase modulation [27]. A simple version of the filtering due to frequency dependent gain (or loss) can be described by the perturbing term $a\,\partial^2 u/\partial t^2$, where a is the filtering strength. A simplified model of excess linear and nonlinear gain can be described by terms of the form [27]

$$c_1 u + c_2|u|^2 u + c_3|u|^4 u.$$

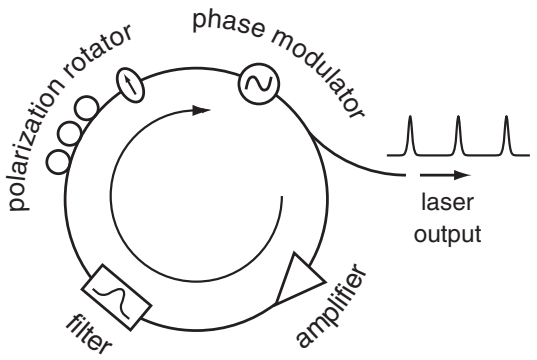

Figure 1. *A schematic of a fiber ring laser where an optical pulse passes through an amplifier, filter, polarization rotator, and phase modulator on each pass around the ring.*

Typically, c_1 will be negative, so really this term models a small excess linear loss. In addition, a simplified model of the active phase modulation is a term of the form $ib\cos(\omega t)u$ [32], where b is the modulation strength and the external modulation frequency is ω. In what follows, we will assume $\omega = 2\pi/25$, i.e., a 25 ps modulation period. It should be noted that, in practice, all of the above effects are discrete or lumped, i.e., the perturbations are each applied to the pulse once per pass through the loop. Since the net change to the pulse per pass is small, it is permissible to replace the lumped effects with their averaged continuous versions as above [26].

In addition, the linear gain provided by the amplifier is also accompanied by amplified spontaneous emission (ASE) noise. This additional perturbation is modeled as zero-mean, delta-correlated Gaussian white noise that is added to the pulse as it passes through the amplifier on every round trip through the fiber ring [34] and takes the form

$$\sum_{n=0}^{N} f_n(t)\delta(z-n), \text{ where} \tag{2.2a}$$

$$\langle f_i(t)\rangle = 0 \text{ and} \tag{2.2b}$$

$$\langle f_i(t) f_j^*(t')\rangle = \sigma^2\delta(t-t')\delta_{ij}\,. \tag{2.2c}$$

(As written, the noise bandwidth is infinite, but in reality it is merely much larger than that of the soliton; the bandwidth is finite in numerical simulations, of course [34].) It is easiest to implement this stochastic perturbation in discrete, random jumps. For simplicity, we have assumed that the noise is added once per dispersion length; if the fiber ring is shorter than this, we can combine the noise from multiple passes into a single larger perturbation once per dispersion length, as long as the total remains small. The noise variance is then $\sigma^2 = [(G-1)^2\eta_{sp}\hbar\omega_0 T_0\gamma]/[G\ln G|\beta|] = 2.667\times 10^{-5}$ [34], where we have used $G = 40$ (or 16 dB) as the amplifier gain needed to compensate loss due to the components in the loop, $\eta_{sp} = 2.0$ is the ASE (excess noise) factor, $\hbar$ is Planck's constant, $\omega_0 = 1.22\times 10^{15}$ 1/s is the carrier frequency, $T_0 = 1$ ps, and $\beta = 0.2$ ps^2/km. Thus, the full perturbative term F on the

right-hand side of (2.1) is

$$F(z,t) = a\frac{\partial^2 u}{\partial t^2} + ib\cos(\omega t)u + c_1 u + c_2|u|^2 u + c_3|u|^4 u + \sum_{n=0}^{N} f_n(t)\delta(z-n). \tag{2.3}$$

When $F \equiv 0$, the NLS admits the well-known soliton solution

$$u_s(z,t) = E\,\mathrm{sech}(E(t-T))\exp(i\Omega(t-T) + i\phi), \tag{2.4}$$

where the soliton parameters E, Ω, T, and ϕ represent, respectively, the amplitude, frequency, position, and phase of the soliton. This solution is exact for (2.1) with $F \equiv 0$ when

$$E(z) = E_0, \qquad \Omega(z) = \Omega_0,$$
$$T(z) = T_0 + \Omega z, \qquad \phi(z) = \frac{1}{2}(E^2 + \Omega^2)z + \phi_0,$$

where E_0, Ω_0, T_0, and ϕ_0 are the constant initial values of the soliton parameters. Additionally, we sometimes use the shorthand notation $u_s(z,t) = u_0(z,t)\exp(i\Theta(z,t))$, where

$$u_0(z,t) = E\,\mathrm{sech}(E(t-T)) \quad \text{and} \quad \Theta(z,t) = \Omega(t-T) + \phi\,.$$

2.1. Soliton perturbation theory and representative parameter regimes. We first consider (2.1) with the deterministic perturbations in (2.3) (i.e., without the noise). Assuming the initial solution is a soliton and the perturbative terms are small in magnitude, then the leading-order effect is to cause the soliton parameters to change slowly. Using soliton perturbation theory [2, 21], the evolution equations for the soliton parameters are (ignoring the phase, as this model is phase-insensitive)

$$\frac{dE}{dz} = \mathrm{Re}\int_{-\infty}^{\infty} u^* F\,\mathrm{d}t, \tag{2.5a}$$

$$\frac{d\Omega}{dz} = \frac{1}{E}\mathrm{Re}\int_{-\infty}^{\infty} (iu_t^* - \Omega u^*)F\,\mathrm{d}t, \tag{2.5b}$$

$$\frac{dT}{dz} = \Omega + \frac{1}{E}\mathrm{Re}\int_{-\infty}^{\infty} (t-T)u^* F\,\mathrm{d}t. \tag{2.5c}$$

Using (2.4) for u to evaluate the integrals to leading order, we obtain

$$\frac{dE}{dz} = (2c_1 - 2a\Omega^2)E + \left(\frac{4}{3}c_2 - \frac{2}{3}a\right)E^3 + \frac{16}{15}c_3 E^5, \tag{2.6}$$

$$\frac{d\Omega}{dz} = -\frac{4}{3}aE^2\Omega - \frac{\omega^2 b\pi}{2E}\mathrm{csch}\left(\frac{\pi\omega}{2E}\right)\sin(\omega T), \tag{2.7}$$

$$\frac{dT}{dz} = \Omega. \tag{2.8}$$

First of all, it is seen that for $c_1 < 0$ (excess loss), $E = 0$ is stable. Strictly speaking, soliton perturbation theory is not applicable in this limit, but a linear stability analysis shows

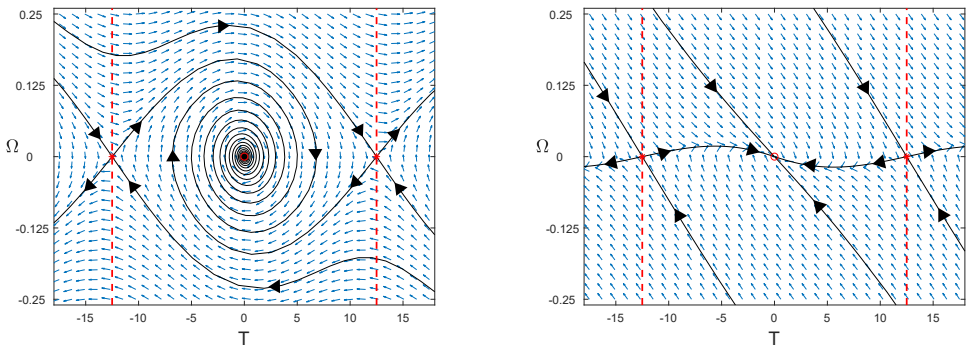

Figure 2. *The continuous dynamics of the soliton ODE system (2.7) and (2.8) are shown in (T, Ω) phase space. Left: the underdamped state with a stable spiral fixed point at $(0, 0)$ and saddles $(\pm\pi/\omega, 0)$ (at bit-slot boundaries, depicted with red-dashed vertical lines, as defined by the modulation frequency). For a pulse to move out of its bit-slot, it must escape the basin of attraction of a stable spiral and cross the stable manifold of the saddle point. Right: the overdamped state where the stable fixed point at $(0, 0)$ is now a nodal sink rather than a spiral. The parameters for the filtering, modulation, nonlinear gain strengths, and steady-state amplitude E_s in each are the representative values listed in the text.*

that $u = 0$ is stable due to the linear excess damping (the c_1 term) and filtering. For $c_2 > 0$ and $c_3 < 0$, there will be a positive stable solution for E when $\Omega = 0$, which we shall denote by E_s. Because the filtering coefficient a is typically small, the pulse energy will be more or less independent of Ω if it is not too large. Since there are two stable pulse energies ($E = 0$ and $E = E_s$), it is possible for added noise to induce transitions between the two, leading to either pulse dropouts or the spontaneous creation of new pulses. For the parameters we will use, however, the probability of either event will be very small [10].

For $E = E_s$, the dynamics of Ω and T can be underdamped or overdamped [10, 32]. In both cases, there are stable steady states at $(E, \Omega, T) = (E_s, 0, 2n\pi/\omega)$ and unstable saddles at $(E, \Omega, T) = (E_s, 0, (2n + 1)\pi/\omega)$ where $n \in \mathbb{Z}$, but the structure of the phase space is different in the two cases. Phase planes for Ω and T with E constant are shown in Figure 2. In the underdamped case, decay to the nonzero equilibrium state is oscillatory, while in the overdamped case a soliton displaced from equilibrium will be restored smoothly to the steady state. The relative strengths of filtering and modulation determine the type of dynamics, with greater filtering leading to overdamping [32]. We will use two representative parameter sets. For the underdamped regime, we will assume $a = 0.002$, $b = 0.01$, $c_1 = -0.01$, $c_2 = 0.034$, $c_3 = -0.02$ with a stable amplitude of $E_s = 1.177$. For the overdamped regime, we assume $a = 0.015$, $b = 0.002$, $c_1 = -0.01$, $c_2 = 0.04$, $c_3 = -0.02$ with a stable amplitude of $E_s = 1.150$.

One effect of the noise is that it adds jitter to the pulse parameters, broadening the laser's linewidth [1]. A more dynamically interesting event is for a noise-induced *position slip* error to occur [32]. Here the noise induces the pulse position to transition from one stable equilibrium to another. Since the goal of mode-locking is to keep a pulse synchronized to the active mode-locking signal, noise-induced sliding of a pulse's position is undesirable, as it leads to a random drift of the mode-locked laser's repetition rate. The rate at which these position slips

occur, of course, is directly related to the rate of this random drift. Because the parameters being considered are such that pulse energy dropouts are very improbable relative to position slips, in what follows we will assume the amplitude to be fixed at $E = E_s$ and only concern ourselves with position slips. We monitored the pulse energy in simulations to verify that this assumption holds.

3. Large deviation theory. When considered as the sole perturbation, the effect of the ASE noise is to cause random walks in the soliton parameters. Since any values of these parameters produce perfectly valid solutions of the NLS equation [34], it is unable to resist such random changes. The random walks that result in the absence of the deterministic perturbations described in section 2 eventually lead to large deviations in the soliton parameters even if the noise added at a single amplifier is small. The purpose of the mode-locking terms, of course, is to limit the growth of perturbations in the pulse parameters. Even in this case, however, there is a small but nonzero probability for the pulse to experience a large deviation.

To address the question of precisely how likely a position slip is to occur in the presence of mode-locking, we pose a stochastic exit problem. Specifically, we wish to determine the most likely way for a pulse starting at equilibrium ($\Omega = 0$, $T = 0$) to escape the effective potential well imposed by the mode-locking and exit a bit-slot by undergoing a position drift of $\pi/\omega = 12.5$. (By symmetry, an exit to $-\pi/\omega = -12.5$ is equally probable.) To answer this question, we first will derive optimized exit paths using the approximating system of ODEs for soliton parameter evolution, (2.6)–(2.8).

In a manner similar to the derivation of (2.6)–(2.8), we use soliton perturbation theory to determine the appropriate stochastic terms that must be added to the soliton parameter ODEs in the presence of noise [34]. Taking F in (2.1) to be the Gaussian white noise terms (ignoring the deterministic perturbations for the moment and neglecting changes in E), we find at leading order

$$\frac{d\Omega}{dz} = \mathrm{Re}\int_{-\infty}^{\infty} \underline{v}_\Omega^* e^{-i\Theta} F\, dt, \tag{3.1a}$$

$$\frac{dT}{dz} = \Omega + \mathrm{Re}\int_{-\infty}^{\infty} \underline{v}_T^* e^{-i\Theta} F\, dt, \tag{3.1b}$$

where the adjoint modes of the linearized NLS operator $\underline{v}_\Omega$ and $\underline{v}_T$ are given by

$$\underline{v}_\Omega = -\frac{i}{E}\frac{\partial u_0}{\partial t}, \tag{3.2a}$$

$$\underline{v}_T = \frac{1}{E}(t - T)u_0. \tag{3.2b}$$

Since the noise is added as a jump in u when a pulse reaches an amplifier, integrating (3.1) in z across an amplifier gives the corresponding jumps in the soliton parameters,

$$\Delta\Omega_n = \mathrm{Re}\int_{-\infty}^{\infty} \underline{v}_\Omega^* e^{-i\Theta} f_n\, dt, \tag{3.3}$$

$$\Delta T_n = \mathrm{Re}\int_{-\infty}^{\infty} \underline{v}_T^* e^{-i\Theta} f_n\, dt. \tag{3.4}$$

When the noise f_n is given by (2.2), $\Delta\Omega_n$ and ΔT_n are zero-mean Gaussian random variables with variances [34]

$$\mathbf{E}[\Delta\Omega_n^2] = \sigma_\Omega^2 = \sigma^2 E/3, \tag{3.5}$$

$$\mathbf{E}[\Delta T_n^2] = \sigma_T^2 = \pi^2\sigma^2/(12E^3)\,. \tag{3.6}$$

3.1. Derivation of optimal exit paths. Equations (2.7) and (2.8) with added random jumps given by (3.3) and (3.4) form an approximate version of the soliton dynamics in the full model, (2.1) with (2.3). This system is, letting $A = -4aE^2/3$ and $B = -\omega^2 b\pi\mathrm{csch}(\pi\omega/2E)/(2E)$, explicitly given by

$$\frac{d\Omega}{dz} = A\Omega + B\sin(\omega T) + \sum_{n=0}^{N}\Delta\Omega_n\delta(z-n), \tag{3.7a}$$

$$\frac{dT}{dz} = \Omega + \sum_{n=0}^{N}\Delta T_n\delta(z-n). \tag{3.7b}$$

In particular, these equations approximately describe the dynamics of a position slip error and allow us to derive exit paths for such an event. The most probable series of parameter kicks $\Delta\Omega_n$ and ΔT_n are those which have the lowest combined weighted L_2 norm, as the underlying noise is Gaussian. Let the functions $\eta_\Omega(z)$ and $\eta_T(z)$ be continuous functions that stand in for the discrete parameter kicks (this continuum approximation is justified, as the length scale on which the deterministic dynamics work in the system is much longer than the amplifier spacing, here $\Delta z = 1$). Then the constrained optimization problem needed to be solved for an optimal exit path is to minimize [33, 34, 39]

$$\min_{\eta_\Omega,\eta_T} S = \min_{\eta_\Omega,\eta_T}\frac{2}{\sigma^2}\int_0^{z_L} C\eta_\Omega^2 + D\eta_T^2\,dz \tag{3.8}$$

subject to

$$\frac{d\Omega}{dz} = A\Omega + B\sin(\omega T) + \eta_\Omega, \tag{3.9a}$$

$$\frac{dT}{dz} = \Omega + \eta_T, \tag{3.9b}$$

where $C = 1/\sigma_\Omega^2 = 3/(E\sigma^2)$ and $D = 1/\sigma_T^2 = 12E^3/(\pi^2\sigma^2)$. We want the final soliton position to be $T(z_L) = \hat{T}$ (the one that gives a position slip), so we have three boundary conditions: $T(0) = T_0$, $\Omega(0) = \Omega_0$, and $T(z_L) = \hat{T}$. Here the idea is to the minimize the probability associated with a particular set of deviations. This is equivalent to finding the Freidlin–Wentzell least action path [17] (as elucidated in more detail in section 4) and related methods [3, 16, 38]. Following standard variational calculus methods for optimization with differential side constraints, we construct a constrained functional using Lagrange multipliers [18]

$$\int_0^{z_L}\left\{C\eta_\Omega^2 + D\eta_T^2 + \lambda_1(z)\left[\frac{d\Omega}{dz} - A\Omega - B\sin(\omega T) - \eta_\Omega\right] + \lambda_2(z)\left[\frac{dT}{dz} - \Omega - \eta_T\right]\right\}dz. \tag{3.10}$$

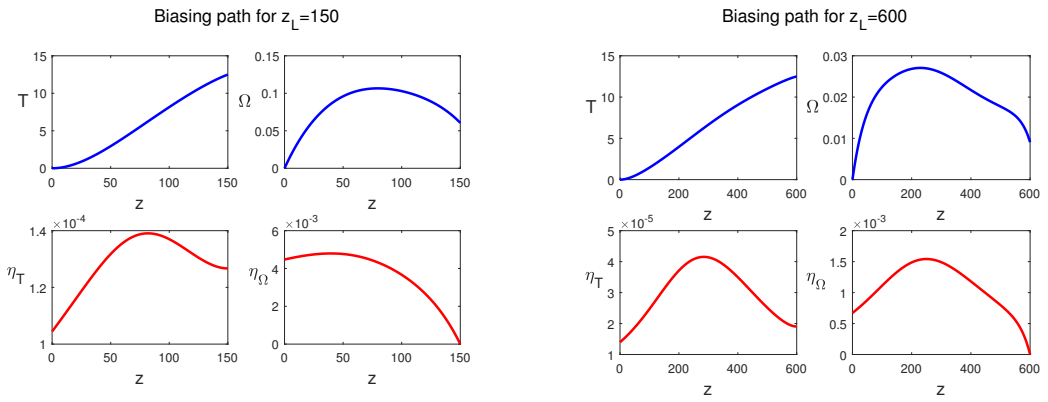

Figure 3. *Two representative solutions for the exit path in the overdamped parameter regime showing four-part solutions with different propagation distances, z_L. The lower panels show the biasing path coefficients for the frequency and position modes, and the upper panels show the expected frequency and position trajectories under biasing. The solutions were computed numerically using the MATLAB boundary value problem solver BVP4C.*

Taking variations with respect to the unknown functions $\eta_\Omega, \eta_T, \Omega$, and T, we find

$$\delta\eta_\Omega \to 2C\eta_\Omega - \lambda_1 = 0, \tag{3.11a}$$

$$\delta\eta_T \to 2D\eta_T - \lambda_2 = 0, \tag{3.11b}$$

$$\delta\Omega \to -A\lambda_1 - \lambda_2 - \frac{d\lambda_1}{dz} = 0, \tag{3.11c}$$

$$\delta T \to -B\omega\lambda_1 \cos(\omega T) - \frac{d\lambda_2}{dz} = 0\,. \tag{3.11d}$$

We use (3.11a) and (3.11b) to specify the Lagrange multipliers $\lambda_1 = 2C\eta_\Omega$ and $\lambda_2 = 2D\eta_T$ giving

$$\frac{d\eta_\Omega}{dz} = -A\eta_\Omega - \frac{D}{C}\eta_T, \tag{3.12a}$$

$$\frac{d\eta_T}{dz} = -\frac{B\omega C}{D}\eta_\Omega \cos(\omega T)\,, \tag{3.12b}$$

which along with (3.9a) and (3.9b) form a boundary value problem with four boundary conditions: $T(0) = T_0$, $\Omega(0) = \Omega_0$, $T(z_L) = \hat{T}$, and $\eta_\Omega(z_L) = 0$. The last condition is a natural boundary condition arising from the free boundary for $\Omega(z_L)$ [18].

3.2. Exit path behaviors. The solution of the boundary value problem given by (3.9) and (3.12) with $T_0 = 0$, $\Omega_0 = 0$, and $\hat{T} = 12.5$ gives an optimal exit path after propagation through $\lfloor z_L \rfloor$ amplifiers. This boundary value problem is solved numerically with the MATLAB BVP4C package. Representative numerical solutions are shown in Figures 3 and 4 for the underdamped and overdamped regimes, respectively. The exit paths for the overdamped regime have a similar character for all system lengths, as the optimal way to resist

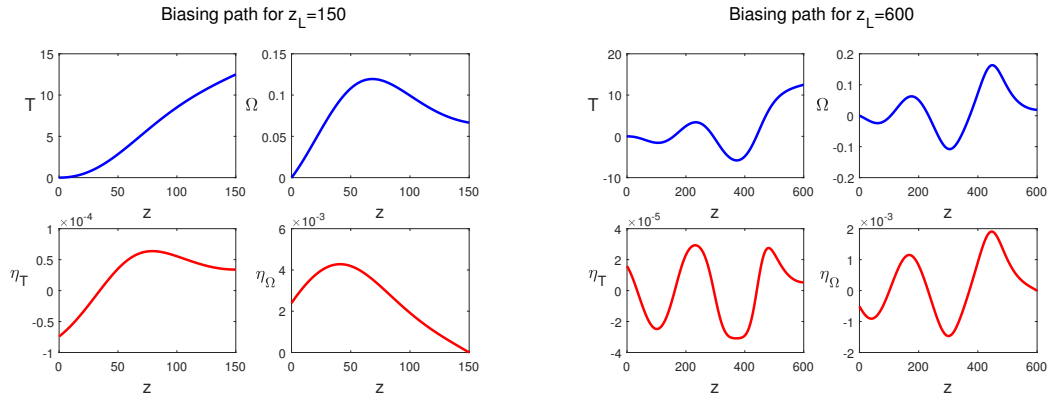

Figure 4. *Two representative solutions for the exit path in the underdamped regime showing four-part solutions with different propagation distances, z_L. The lower panels show the biasing path coefficients, and the upper panels show the expected soliton parameter trajectories under biasing. The solutions were computed numerically using the MATLAB boundary value problem solver BVP4C. For a long distance, such as the solution shown on the right, the biasing path is oscillatory, but for a short distance (left) it is not.*

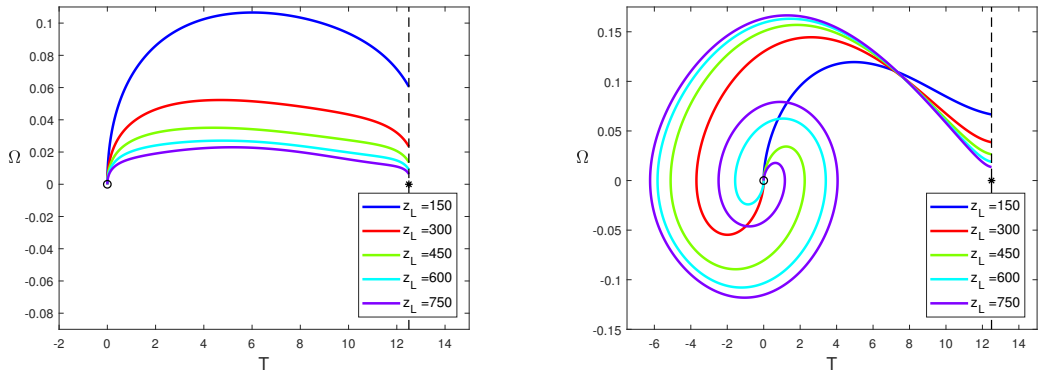

Figure 5. *Computed biasing path trajectories visualized in (T, Ω) phase space for different system lengths in the overdamped and underdamped regimes. Left: the overdamped regime. All paths monotonically approach exit and do not change qualitatively as the system length increases. For longer propagation distances, less frequency driving is necessary to force an exit. Right: the underdamped regime. As z_L increases, the biasing paths undergo more and more oscillations before exiting due to the oscillatory dynamics near the critical point.*

the overdamped mode-locking is to monotonically push the pulse toward its exit. However, the exit paths in the underdamped regime exhibit significant qualitative differences as the system length is increased. For a sufficiently large propagation distance, z_L, the optimal path takes advantage of the oscillatory dynamics inherent in the mode-locking and wraps around the spiral before exiting, with the number of oscillations (or "loops") dependent on the overall propagation distance. Figure 5 shows the qualitative behavior changes in the exit paths in both regimes as the system length is varied.

It is best to understand the bifurcating oscillatory paths in the underdamped regime as different modes of exit. As the system length increases, an infinite cascade of biasing paths

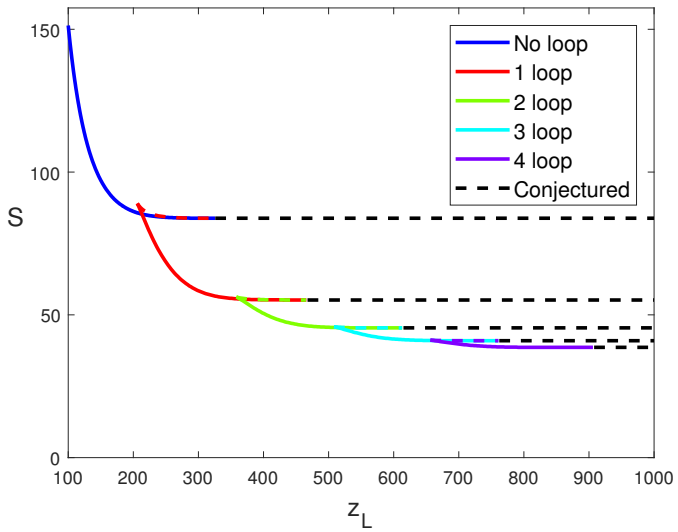

Figure 6. *Bifurcation diagram showing biasing paths in the underdamped regime with different numbers of oscillations or loops numerically continued by varying z_L and measured in terms of their action according to (3.8). The oscillatory biasing paths are conjectured to arise through a bifurcation at infinity, and each have a critical propagation distance at which they disappear (visualized here as folds in the bifurcation diagram). The black-dashed lines are conjectured extensions of computed solution families.*

occur which are each local minima in the least action problem equations (3.9) and (3.12). Continuing the paths in the right panel of Figure 5 shows that multiple paths of exit from the fixed point with different numbers of oscillations, or loops, coexist at all distances past $z_L = 206$. Figure 6 shows a bifurcation diagram continuing these paths in the propagation distance, z_L, with the path action equation (3.8) as the dependent variable. Additionally, the bifurcation diagram demonstrates that there are multiple system lengths z_L for which there are two types of paths that are roughly equal in action (e.g., see the region $205 < z_L < 220$).

4. Numerical simulations.

To quantify the probabilities of position slips happening in this system, MC simulations were performed. The full NLS equation, (2.1), was integrated numerically with different noise realizations many times, and statistics were computed based on the final results. Each sample in the MC simulations was solved numerically using the split-step Fourier method [40] with 256 Fourier modes, a computational window in time t of width 80, and a propagation stepsize of $dz = 0.05$. The initial condition was a soliton with the stable amplitude for the parameter regime being simulated, as given in section 2.1.

Position slip probabilities due to the mode-locking can be quite small, so the MC simulations were augmented with importance sampling (ISMC). In these simulations, the underlying probability distribution in the problem (Gaussian white noise) is replaced with a biasing distribution from which samples are drawn. Then the correct probabilities are computed using the likelihood ratio, which is the ratio between the original and the biased probabilities used to generate that sample [34]. Here the Gaussian biasing distributions utilized means obtained from solutions to the system of equations (3.9) and (3.12) using coefficients $\underline{v}_\Omega$ and $\underline{v}_T$,

respectively, as described previously [34].

In particular, previous work has shown that the noise should be mean shifted at each amplifier by an amount that is proportional to the adjoint mode associated with the quantity being biased (multiplied by $e^{i\Theta}$, so that it is in phase with the soliton), rather than the forward mode, a somewhat counterintuitive result [34]. This is equivalent to saying that the Freidlin–Wentzell least action minimizer for a prescribed soliton change after passage through one amplifier is in the direction of the adjoint modes of the linearized NLS equation [34]. In this case, the relevant modes that should be biased at each amplifier are the position and frequency adjoint modes, the latter being included since frequency perturbations become position perturbations upon propagation. Explicitly, the mean shift at the nth amplifier due to the biasing is

$$f_n(t) = \left(\eta_\Omega(z_n)\frac{\underline{v}_\Omega}{||\underline{v}_\Omega||^2} + \eta_T(z_n)\frac{\underline{v}_T}{||\underline{v}_T||^2}\right)e^{i\Theta}, \tag{4.1}$$

where η_Ω and η_T are the biasing coefficients that give the magnitude of the shift in the direction of each mode (where the modes and total phase of the soliton are computed at the amplifier). These biasing coefficients are the same as the functions used in (3.9a) and (3.9b). With this formulation, the functional minimized in the optimization problem (3.8) to maximize the probability of an exit path is the L_2 norm of (4.1) (note that there is no cross term since the modes are orthogonal with respect to the inner product associated with the linearized NLS operator) [34]. The normalization constants are chosen so that the mean shift at each amplifier equation (4.1) agrees with (3.9a) and (3.9b) after evaluating the jump conditions (3.3) and (3.4). The constants C and D in (3.8) are then seen to be related to the adjoint mode norms

$$\frac{\sigma^2}{2}C = ||\underline{v}_\Omega||^{-2} = \frac{3}{2E}, \tag{4.2a}$$

$$\frac{\sigma^2}{2}D = ||\underline{v}_T||^{-2} = \frac{6E^3}{\pi^2} \tag{4.2b}$$

and are solely dependent on $E(z) \approx E_s$ and therefore treated here as constants. (An overall multiplicative factor does not affect the minimization, of course.)

Taking the biasing coefficients resulting from the solutions of (3.9) and (3.12) in section 3.1 and incorporating them into (4.1) allows us to bias the noise to optimally produce a desired position slip. The simulation process to create a biased sample is to solve the deterministic part of (2.1) until an amplifier is encountered. The split-step solver is then stopped, and the mean shift equation (4.1) is computed using the precomputed biasing coefficients (extracted by fitting a soliton through the noisy pulse) and the soliton's current parameters (for construction of the adjoint modes and total phase). After adding the biased noise using this mean shift, the split-step solver is then restarted and the process is repeated until the simulation is terminated. Thousands of samples are typically generated, and the final pulse position is binned to estimate its probability distribution. Additionally, in order to form full probability density functions (P.D.F.s) of pulse position that resolve positions throughout a bit-slot and into neighboring bit-slots, multiple biasings were used. That is, to cover larger regions, biasing

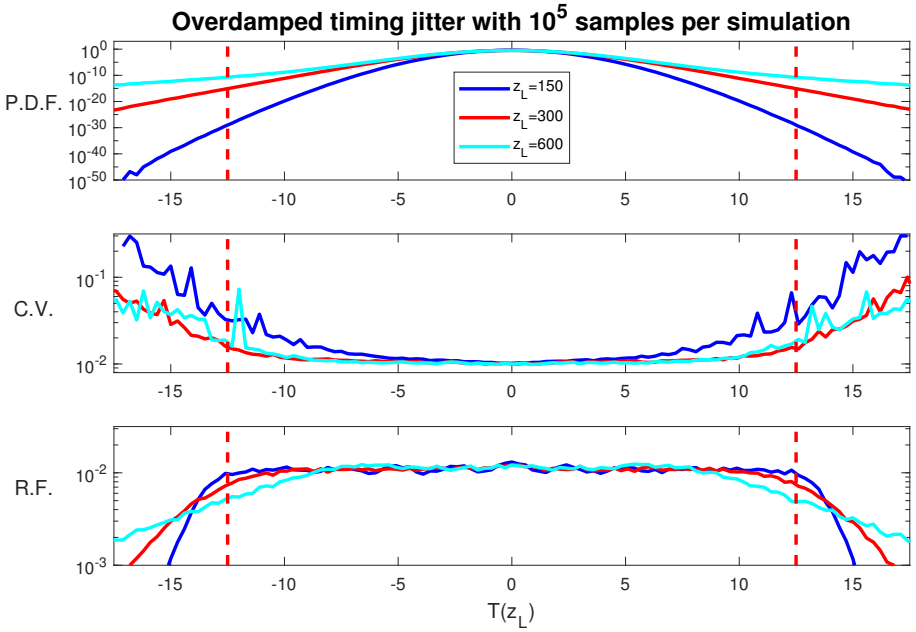

Figure 7. *ISMC simulations of soliton position in the overdamped regime for three system lengths. Top panel: the probability density function (P.D.F.) of soliton position T after propagation through z_L amplifiers. Eleven position targets were included throughout the domain (evenly spaced between $T(z_L) = -12.5$ and $T(z_L) = 12.5$) with $100,000$ total samples. Middle panel: the coefficient of variation (C.V.) indicates simulation convergence: a smaller and smoother C.V. indicates better convergence. We see that convergence is generally good for moderate propagation distances and worse for very short or long distances. Bottom panel: the raw relative frequency histogram of sample location for each simulation. Fewer samples reach the bit-slot boundary for the longer-distance simulations.*

paths with different position targets $\hat{T} = T(z_L)$ were included in multiply importance-sampled Monte Carlo (MISMC) simulations and then combined using the *balance heuristic* [44], which weights contributions from each biasing distribution by the likelihood that each particular sample came from that distribution. Thus, distributions producing low numbers of samples in specific regions are downweighted when others producing larger numbers of samples are available.

4.1. Importance sampling in the overdamped regime. We performed ISMC to assess pulse position probabilities in the laser model using the overdamped parameters given in section 2.1. P.D.F.s of the soliton's position after propagation through varying values of z_L amplifiers are included in Figure 7.

The bit-slot boundaries are indicated in red, and we see that the probability of a large position shift is greatly increased as the system length grows. Included in these figures is the *coefficient of variation* (C.V.), or the measure of the intrabin standard deviation divided by the bin's overall probability, which provides one way of assessing convergence of the simulations, even when probabilities are small. In all simulations, the samples converge well for regions at or near equilibrium but not necessarily in the tails. This is disadvantageous, as the tails are precisely the regions where position slips occur. For very short distances and very long

distances, the simulations do not converge well in the tails (as in the left and right panels of Figure 7). However, for intermediate distances, the simulations converge well. For short distances, the probabilities that we are attempting to capture are extremely low, so it is natural that convergence is slower, but for long distances the poor convergence requires additional explanation since the probabilities are larger.

The poor convergence for long system lengths can be understood by examining the bottom panel of Figure 7. The relative frequency histogram indicates the spread of samples in the standard ISMC simulations, which were conducted with 11 position targets spread evenly throughout the bit-slot and at the left and right boundaries. The samples are evenly spread throughout the interior of the bit-slot for all distances, but fewer samples are clustered around the boundaries as the system length increases. For $z_L = 600$ (light blue), too many simulations either get ahead of or behind the biasing, and end up either exiting early or getting pushed to the center of the neighboring bit-slot, or they get pushed back into the current bit-slot (a situation which is shown in Figure 8).

This issue can be ameliorated by using *dynamic importance sampling* [12]. This method, alternately known as targeted ISMC [15] or state-dependent ISMC [6], recomputes the biasing path midsimulation while a sample is being drawn. A demonstration of this scheme is shown in Figure 8. Equations (3.9) and (3.12) are resolved during the simulation by changing the left-hand boundary conditions to be the soliton's current frequency and position, as shown in the bottom left panel of Figure 8. The biasing path expects the pulse to be at the position of the blue pulse, but due to accumulated noise the pulse is actually that indicated in green. Therefore, we correct the biasing path by solving the biasing boundary value problem using the frequency and position of the soliton shown with blue dashes, which is the best-fit soliton to the noisy pulse, and use the resulting biasing path throughout the rest of the simulation or until the path is dynamically recalculated again. The criterion used to decide when to recalculate the pulse can be varied; one can recalculate the path after passage through a fixed number of amplifiers or when the deviation between the expected and computed positions is larger than some prescribed tolerance, for example. Here we choose to retarget periodically with a specific number of amplifiers between recalculations in order to precisely control how many times the boundary value problem needs to be solved since computing its solution is expensive relative to the computational costs of the Fourier split-step solver. The convergence benefits of dynamic ISMC typically come from more tightly clustering samples around a desired target [13, 15]. Here dynamically recalculating the optimal path improves the number of samples that finish in the vicinity of the bit-slot boundary. This can be seen for $z_L = 600$ simulations in Figure 9; using dynamic ISMC eliminates the fluctuations in the C.V. near $T(z_L) = \pm 12.5$.

Overall, performing ISMC in the overdamped region is mostly straightforward, as the biasing paths are typically similar in character and solving (and resolving) the biasing boundary value problem is simple. Accumulated deviations from biasing paths cause poorer convergence for exit probabilities at longer distances, but this can be counteracted by the use of dynamic ISMC. Short simulation distances have very low exit probabilities and require a large number of samples to be resolved irrespective of the amount of recalculation. Dynamic ISMC typically provides notable convergence benefits with a relatively low amount of recalculation. Performing many path recalculations provides diminishing returns in convergence, however,

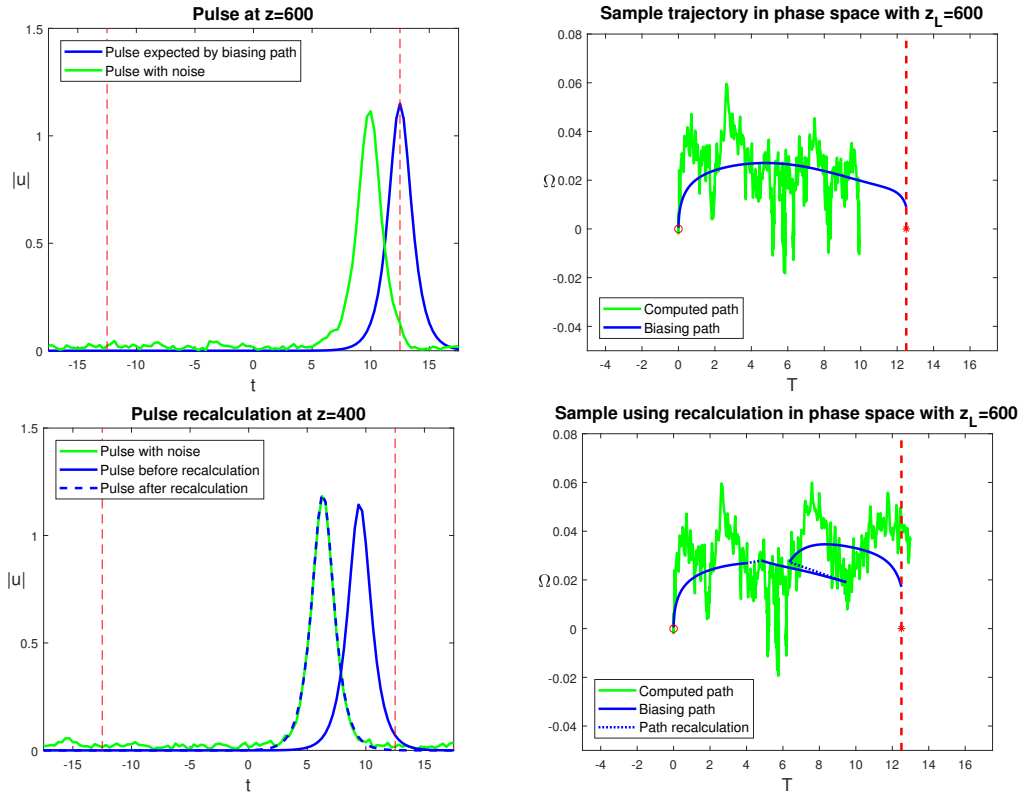

Figure 8. *A single sample trajectory with $z_L = 600$ showing the effects of dynamic path recalculation on pulse position. Top panels: a pulse and its trajectory computed using ISMC without path recalculation. The pulse lags behind the optimal sampling path (blue in the right-top panel) and approaches the saddle from a direction which causes it to be forced back to the left, away from the exit target of $\hat{T} = 12.5$. Bottom panels: the same noise realization as in the top panels is used in a sample drawn with recalculation of the path every* 200 *amplifiers. Bottom left: the biasing path recalculation at $z = 400$ showing the computed pulse, the pulse originally expected by the biasing path, and the pulse used for biasing after the path is recalculated. Bottom right: the sample trajectory in phase space again showing the parameter random walk and the expected biasing path, which is now piecewise-smooth due to the periodic recalculation. The pulse now exits the specified interval and ends up much closer to the intended position target.*

and greatly increases the computational cost. Performing simulations to track the probability of exit directly as a function of system length z_L confirms these characteristics. These simulations are important, as they directly address the formulation of the position slip as an exit problem, and are shown in Figure 10, which go out to $z_L = 600$ with 1,000,000 total samples. The figure shows the overall exit probability, i.e., the probability that $T \geq \hat{T}$ vs. z_L. For an accurate exit probability estimate, the P.D.F. for $T(z_L)$ must be computed accurately at and just beyond the exit point. The nondynamic ISMC simulations (blue) produce smooth estimates for intermediate distances and experience variance fluctuations for longer distances. Dynamic ISMC with one path recalculation greatly improves the convergence, eliminating the fluctuations and generally lowering the C.V. with the same number of samples. Recalculating the path twice slightly lowers the C.V. level relative to one recalculation, but recalculating

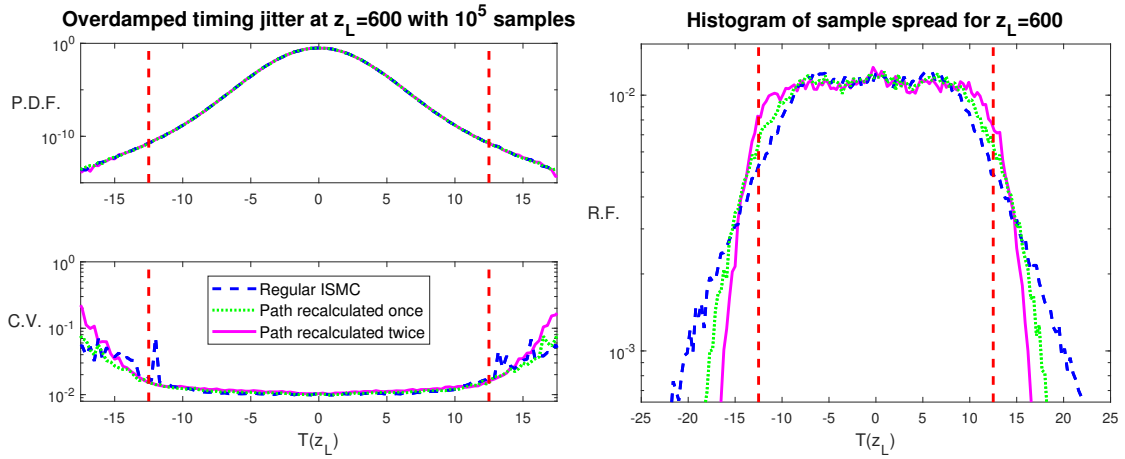

Figure 9. *ISMC simulations for the overdamped case showing different amounts of dynamic recalculation, and their effects on convergence, for $z_L = 600$ simulations. In addition to the blue, which has no recalculation, curves showing simulations where the path was recalculated once (at $z = 300$) or twice (at $z = 200$ and $z = 400$) are shown. Dynamic ISMC concentrates more samples in the region of the bit-slot boundary and improves the convergence there. The improvement in convergence between recalculating the path twice versus once is small.*

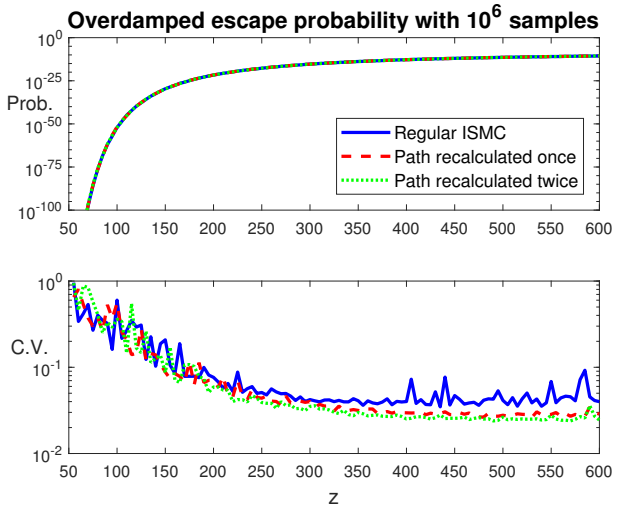

Figure 10. *ISMC simulations of soliton exit simulations in the overdamped case with varying amounts of dynamic recalculation. Top: the P.D.F. of pulse exit probability as a function of distance. Bottom: the C.V. shows improved convergence for the dynamic simulations compared with the regular ISMC, with additional but saturating benefit as more recalculation is used.*

further does not provide any noticeable additional benefit.

4.2. The underdamped regime and the problem of multiple paths. ISMC simulation in the underdamped regime is more difficult due to the oscillatory nature of the biasing paths. The numerical solution of the biasing problem equations (3.9) and (3.12) involves making an initial guess and then iterating to find a solution. While this works well in the

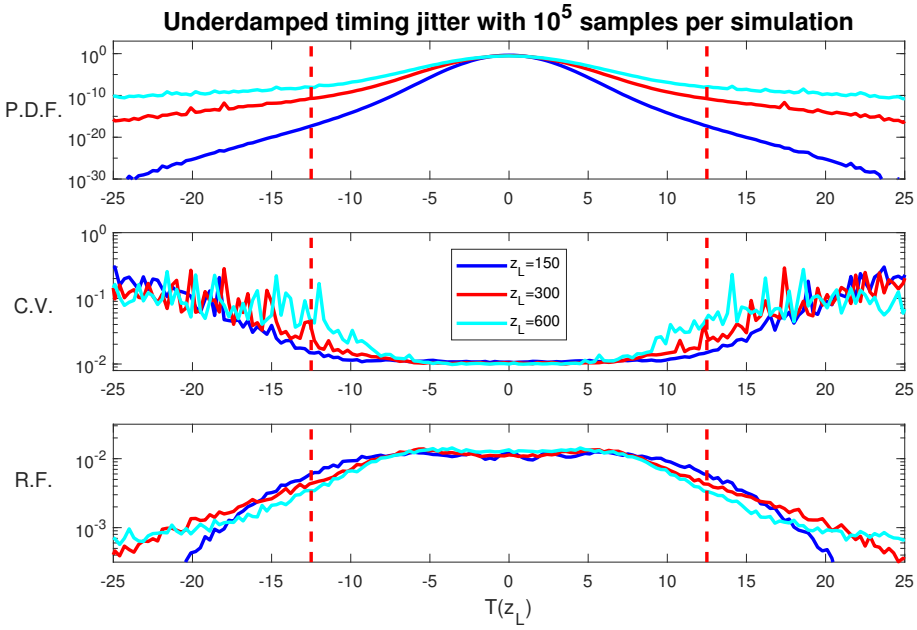

Figure 11. *ISMC simulations of soliton position in the underdamped regime at three system lengths. Top panel: the P.D.F. of soliton position T after propagation through z_L amplifiers. Eleven position targets were included throughout the domain (evenly spaced between $T(z_L) = -12.5$ and $T(z_L) = 12.5$) with 100,000 total samples. Middle panel: the C.V. provides an estimate of simulation convergence. Bottom panel: the raw relative frequency histogram of sample location for each simulation.*

overdamped case, in the underdamped regime a generic initial guess (for example, a constant or linear function) and use of the MATLAB boundary value solver BVP4C do not always converge to a proper solution at every distance. When optimal biasing solutions can be found, simulations for the pulse position P.D.F. can be completed; examples are shown in Figure 11. As before, the general trend is that ISMC simulations converge better for shorter than for longer distances. The distances with good convergence are much shorter than before, however; by $z_L = 300$, the C.V. near the bit-slot boundaries exhibits significant fluctuations and we see that progressively fewer samples fall in the proper range as system length increases, as shown in the bottom panel of Figure 11.

The reasons behind this behavior can be explored by considering Figures 6 and 12. In particular, the convergence of the exit probability simulations is good up until $z_L \approx 200$, but past that point the convergence degrades. Furthermore, multiple locally optimal paths with similar actions appear soon after this, at $z_L \approx 206$. Multiple importance sampling and other similar variants [37] predict that all events with similar probabilities must be included in ISMC simulations, and omitting one or more similarly likely paths has been shown to lead to convergence problems for the results [39]. This issue is observed here if one performs ISMC simulations for the position P.D.F. at $z_L = 210$, since this distance has three distinct paths with similar action values, as shown in Figure 13. Simulations were performed to find the P.D.F. using the least action path, the two paths with the smaller actions, and all three paths. One of the results was obtained using only the path with the smallest action (the

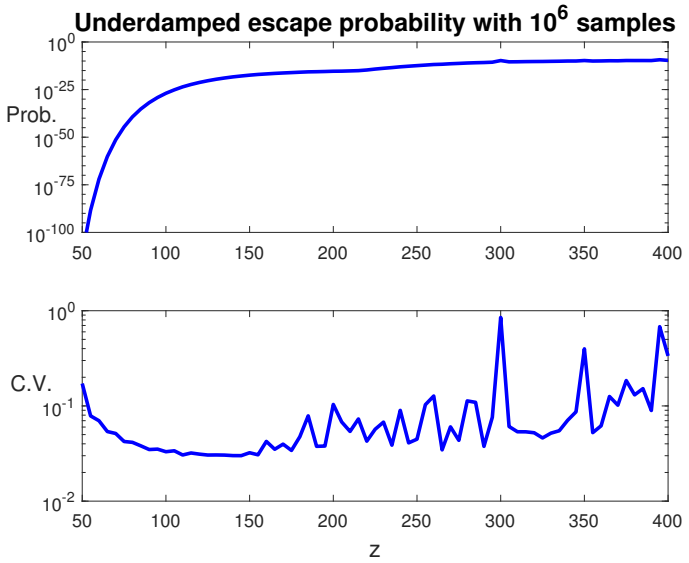

Figure 12. *ISMC simulations of soliton exit simulations in the underdamped regime. Top: the P.D.F. of pulse exit probability as a function of distance. Bottom: the C.V. shows good convergence for only the shortest system lengths.*

largest probability), and it is seen that this can substantially misestimate the P.D.F. near the target. The simulation with one path converged to a significantly lower estimate for the probabilities, while the simulations with two and three paths converged to the same values. None of the three simulations converged quickly (as measured by the C.V.), however, with the single path simulations having the worst convergence.

Applying dynamic ISMC in the underdamped case is also more difficult than before, again due to problems solving the biasing boundary value problem. Performing dynamic simulations for the full P.D.F. (i.e., with multiple position targets) for even intermediate distances does not succeed, as the boundary value solver fails to converge to a usable path sufficiently frequently. The problem is not so severe for simulations of only the region around the exit point (i.e., a single position target) and can be shown to be effective in some cases (e.g., Figure 14), but the boundary value solver failure rate is an impediment for relatively long distances (such as $z_L = 600$) and there are many distances where dynamic ISMC appears to not improve convergence.

With additional effort, dynamic MISMC simulations can be used to improve the results for the P.D.F. at distances just past the point at which multiple paths appear, e.g., $z_L = 210$. Note that combining MISMC with dynamic recalculation requires simultaneously recalculating, using the same intermediate position, both the path currently being used to generate samples *and* all of the other paths since dynamically updated versions of the latter are needed for proper implementation of the balance heuristic [5, 37, 44]. The results at $z_L = 210$ using one, two, or three paths, taken in order from highest to lowest probability, are shown in Figure 15. The results again show that convergence is best when all paths are included and that using just one path is not sufficient; the one-path simulation still converges to a substantially lower

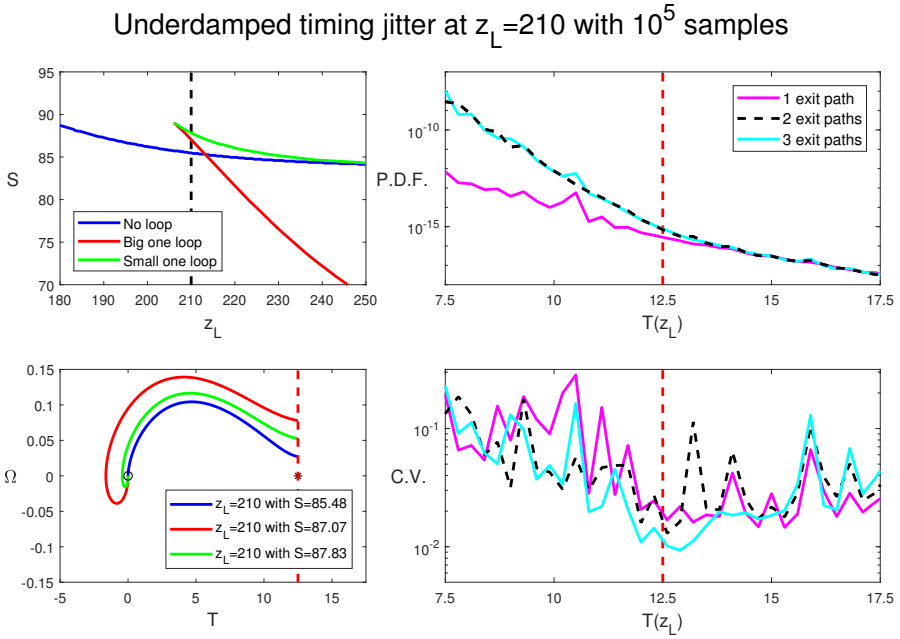

Figure 13. *MISMC in the first region of multiple paths. Left top: a zoomed version of the bifurcation diagram. Left bottom: the three paths and their actions. Right: MISMC results using the three paths. The blue curve uses only the global minimum curve, the red uses the two paths with least action, and the green uses all three. The P.D.F. of pulse position shows that using multiple paths provides a substantial correction to the probability in the tail, though the convergence in the C.V. is not very good, regardless of how many paths are used.*

probability, and the C.V. is markedly worse. The C.V. for the two-path simulation is also not improved relative to the nondynamic simulation (shown in black), while only the three-path simulation shows the improved convergence expected with path recalculation. It is also worth noting that in this case, once the remaining distance shrinks sufficiently, the three different paths all merge and become identical.

Not all simulation distances past the onset of the oscillatory paths require the use of multiple paths in the simulation, however. Examination of Figure 6 indicates that between roughly $250 < z_L < 350$ the path with one oscillation dominates the exit probability, for instance.

Dynamic ISMC is still difficult to implement at longer distances, however, and the fraction of failed samples due to nonconvergence in the boundary value problem solver is a persistent issue, particularly for $z_L > 300$. In a nonconvergent sample, the boundary value problem solver, using the remaining portion of the previously computed biasing path as an initial guess, fails to converge to a prescribed tolerance. The convergence rate of the boundary value problem solver can be considerably improved, however, by updating the optimal path using a series of intermediate steps between the previously computed path and the current location in (T, Ω) phase space if the path recalculation initially fails (implementing a simple homotopy algorithm in this case). While this approach is computationally expensive, the rate at which

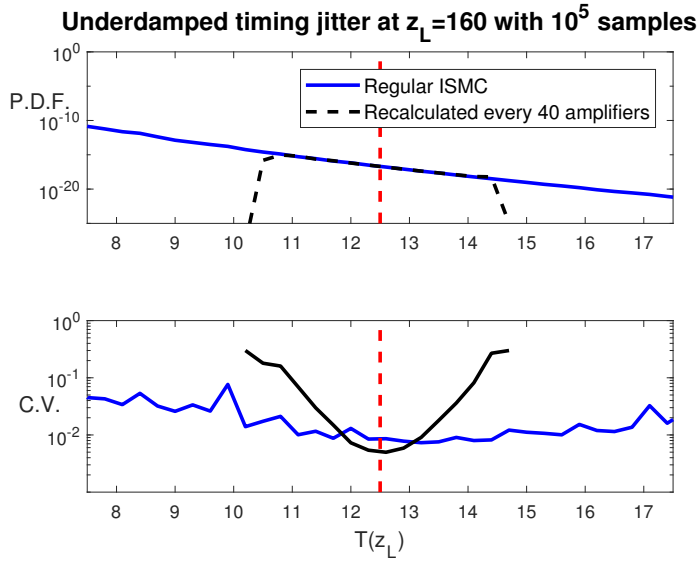

Figure 14. *ISMC simulations of soliton position in the underdamped regime at $z_L = 160$ comparing tail-only simulations using recalculation every 40 amplifiers with simulations without recalculation. Top: the P.D.F. of soliton position T after propagation through z_L amplifiers. Bottom: the C.V. reports simulation convergence, showing that convergence is improved by very tight targeting at the bit-slot boundary.*

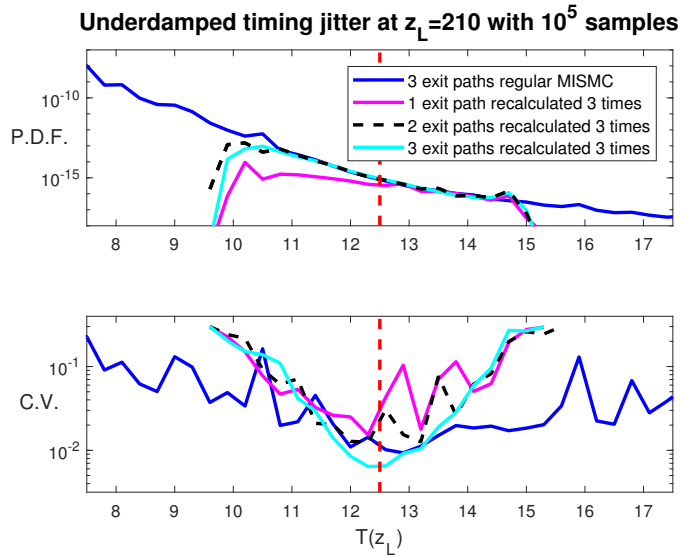

Figure 15. *Dynamic MISMC with three paths. Dynamic MISMC is implemented starting from each of the three paths from Figure 13 with the three-path simulation without recalculation shown in black for comparison. For each dynamic simulation, three-path recalculations were performed for each sample. The one-path simulation still converges to a lower probability than the two- or three-path simulations, while only the three-path simulation shows improved convergence relative to nondynamic MISMC.*

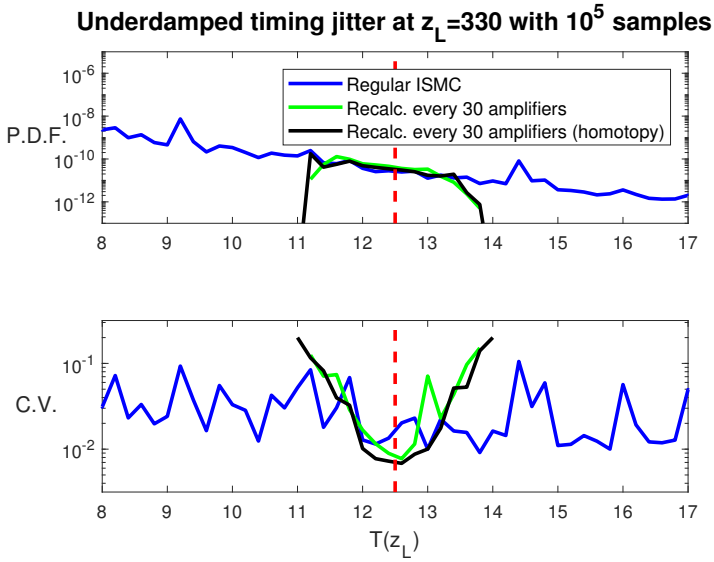

Figure 16. *Dynamic ISMC computed for $z_L = 330$. Top: Regular ISMC, dynamic ISMC, and homotopy-assisted dynamic ISMC simulations were performed with 10 path recalculations per sample for the dynamic simulations. This level of recalculation results in a very narrow spread of samples around the intended target at the bit-slot boundary. Bottom: The C.V. compares the three methods. Dynamic ISMC at this distance results in 50.3% of samples failing due to BVP4C nonconvergence, while the homotopy-assisted method fails in only 3.5% of samples (failed samples were discarded for the purposes of computing the ISMC estimates). This failure rate limits the effectiveness of dynamic ISMC, but the much lower failure rate of the homotopy-assisted method allows the convergence of dynamic ISMC to be significantly improved.*

samples experience a failed recalculation can be cut by an order of magnitude, as shown in Figure 16. The lowered failure rate makes dynamic ISMC effective in this longer propagation regime.

It is possible for even more demanding situations to arise, however. Figure 17 shows a possible trajectory constructed by examining solutions of the exit problem in specific (T, Ω, z_L) regions. In this situation, the starting point has a unique exit trajectory, but as the solution evolves and is pushed away from the optimal path by noise, a second path becomes possible (the dashed curves in the figure). Eventually, the first path disappears and only the second path remains. Performing dynamic ISMC simulations for cases of this type presents a new level of difficulty, of course, as multiple importance sampling requires enumerating the different paths and keeping track of the likelihood ratios for each of them. A general method capable of tracking such situations properly would require combining a continuation/bifurcation method with dynamic MISMC simulations. Such a combined method is beyond the scope of the present work.

5. Discussion.

We have shown that multiple importance sampling techniques based on soliton perturbation theory can be extended to be a useful tool for the study of rare events in mode-locked laser systems, which are of particular interest due to recent technological developments in optical frequency metrology that depend on high-performance laser sources. This rare event simulation method allows fast and accurate simulation of these rare events

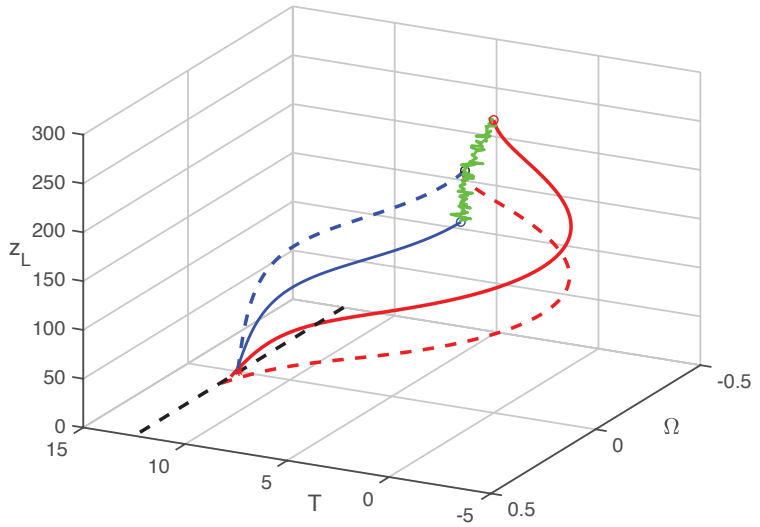

Figure 17. *A situation where the number of exit paths increases midsample. A single exit path, with one oscillation (solid red), exists at the outset of the sample with $z_L = 330$, but the sample trajectory (green) propagates into a region of (T_0, Ω_0, z_L) phase space where a nonoscillatory path is most likely (solid blue). The point at which the two paths are equally likely (dashed curves) is also depicted in between these two points. This situation can be anticipated by considering numerically computed solutions of the exit problem.*

and their statistics, which are difficult to compute with traditional MC methods. While the method does depend on the mathematical structure of the equations and being able to use soliton perturbation theory to derive appropriate approximate biasing schemes, many laser systems do in fact use soliton or near-soliton pulses.

We have also shown that dynamic importance sampling is a useful addition to multiple importance sampling techniques in situations where system dynamics make the tails of probability distributions especially difficult to sample. In the laser model, we considered two parameter regimes where dynamic importance sampling was used to target a saddle point in order to determine the probability of a position slip. In the overdamped regime, computing biasing paths was relatively simple and dynamic ISMC was straightforward to implement. In the underdamped regime, computing biasing paths was considerably more difficult due to the oscillatory character of the biasing paths. In order to sample the tails, we had to both target the saddle and escape from an oscillatory well, leading to biasing paths with differing numbers of oscillations coexisting at various propagation lengths. We implemented dynamic ISMC both in regions where multiple paths have roughly equal probability and at relatively long distances by using a homotopy method to recompute optimal paths. This approach is computationally intensive, but it is a tractable resolution to the inherently difficult problem of performing dynamic ISMC when multiple paths exist in phase space. We anticipate that the methods used here to resolve these difficulties can be a guide to others when rare events arising from multiple paths are possible.

The approach we have used here relies on solving a boundary value problem in order to find optimal biasing paths for the ISMC simulations. This boundary value problem comes

from the minimization of an approximate action functional (from (3.8)) using calculus of variations, guided by the projection of the deterministic and stochastic perturbations into the NLS soliton modes. It may be possible, of course, that other more general methods, such as the geometric minimum action method (GMAM), could be used to minimize the action directly, without appeal to the specific structure of solitons [23, 31, 43]. Such methods have been shown to be robust over a large class of problems (see [32]); their disadvantage is that they are inherently much higher dimensional than lower-dimensional approximate methods employing soliton perturbation theory and are therefore more computationally expensive. This additional computational expense would be magnified significantly, of course, if one were to attempt to use such methods in combination with dynamic ISMC or when multiple optimal paths are present.

REFERENCES

[1] M. J. Ablowitz, B. Ilan, and S. T. Cundiff, *Noise-induced linewidth in frequency combs*, Optim. Lett., 31 (2006), pp. 1875–1877.

[2] M. J. Ablowitz and H. Segur, *Solitons and the Inverse Scattering Transform*, SIAM Stud. Appl. Math. 4, SIAM, Philadelphia, 1981.

[3] M. Bauver, E. Forgoston, and L. Billings, *Computing the optimal path in stochastic dynamical systems*, Chaos, 26 (2016), 083101.

[4] A. Bilenca, A. Desjardins, B. E. Bouma, and G. J. Tearney, *Multicanonical Monte-Carlo simulations of light propagation in biological media*, Opt. Express, 13 (2005), pp. 9822–9833.

[5] G. Biondini, W. L. Kath, and C. R. Menyuk, *Importance sampling for polarization-mode dispersion: Techniques and applications*, J. Light. Technol., 22 (2004), pp. 1201–1215.

[6] J. Blanchet and H. Lam, *State-dependent importance sampling for rare-event simulation: An overview and recent advances*, Surv. Oper. Res. Manag. Sci., 17 (2012), pp. 38–59.

[7] B. J. Bloom, T. L. Nicholson, J. R. Williams, S. L. Campbell, M. Bishof, X. Zhang, W. Zhang, S. L. Bromley, and J. Ye, *An optical lattice clock with accuracy and stability at the* 10^{-18} *level*, Nature, 506 (2014), pp. 71–75.

[8] S. T. Cundiff and J. Ye, *Colloquium: Femtosecond optical frequency combs*, Rev. Modern Phys., 75 (2003), pp. 325–342.

[9] S. A. Diddams, J. C. Bergquist, S. R. Jefferts, and C. W. Oates, *Standards of time and frequency at the outset of the* 21*st century*, Science, 306 (2004), pp. 1318–1324.

[10] G. M. Donovan and W. L. Kath, *Rare event simulation of the performance of an actively mode-locked fiber laser model*, in Proceedings of the 2007 Conference on Lasers & Electro-Optics/Quantum Electronics and Laser Science, 2007, pp. 1953–1954.

[11] G. M. Donovan and W. L. Kath, *An iterative stochastic method for simulating large deviations and rare events*, SIAM J. Appl. Math., 71 (2011), pp. 903–924, https://doi.org/10.1137/100789373.

[12] P. Dupuis and H. Wang, *Importance sampling, large deviations, and differential games*, Stoch. Stoch. Rep., 76 (2004), pp. 481–508.

[13] P. Dupuis and H. Wang, *Dynamic importance sampling for uniformly recurrent Markov chains*, Ann. Appl. Probab., 15 (2005), pp. 1–38.

[14] M. Fischer, N. Kolachevsky, M. Zimmermann, R. Holzwarth, T. Udem, T. W. Hänsch, M. Abgrall, J. Grünert, I. Maksimovic, S. Bize, H. Marion, F. P. D. Santos, P. Lemonde, G. Santarelli, P. Laurent, A. Clairon, C. Salomon, M. Haas, U. D. Jentschura, and C. H. Keitel, *New limits on the drift of fundamental constants from laboratory measurements*, Phys. Rev. Lett., 92 (2004), 230802.

[15] S. L. Fogal, G. Biondini, and W. L. Kath, *Targeted importance sampling for first-order polarization mode dispersion*, in Proceedings of the Conference on Lasers and Electro-Optics (CLEO '03), 2003, pp. 851–853.

[16] E. Forgoston and R. O. Moore, *A primer on noise-induced transitions in applied dynamical systems*, SIAM Rev., 60 (2018), pp. 969–1009, https://doi.org/10.1137/17M1142028.
[17] M. I. Freidlin, J. Szücs, and A. D. Wentzell, *Random Perturbations of Dynamical Systems*, Grundlehren Math. Wiss. 260, Springer, Berlin, Heidelberg, 2012.
[18] I. M. Gelfand and S. V. Fomin, *Calculus of Variations*, Prentice-Hall, Englewood Cliffs, NJ, 1963.
[19] J. M. Gill, *Lasers: A 40-year perspective*, IEEE J. Sel. Top. Quantum Electron., 6 (2000), pp. 1111–1115.
[20] K. L. Hall, J. D. Moores, K. A. Rauschenbach, W. S. Wong, E. P. Ippen, and H. A. Haus, *All-optical storage of a 1.25 kb packet at 10 Gb/s*, IEEE Photon. Technol. Lett., 7 (1995), pp. 1093–1095.
[21] A. Hasegawa and Y. Kodama, *Solitons in Optical Communications*, Oxford University Press, New York, 1995.
[22] H. A. Haus, *Mode-locking of lasers*, IEEE J. Sel. Top. Quantum Electron., 6 (2000), pp. 1173–1185.
[23] M. Heymann and E. Vanden-Eijnden, *The geometric minimum action method: A least action principle on the space of curves*, Comm. Pure Appl. Math., 61 (2008), pp. 1052–1117.
[24] R. Holzlöhner and C. R. Menyuk, *Use of multicanonical Monte Carlo simulations to obtain accurate bit error rates in optical communications systems*, Opt. Letters, 28 (2003), pp. 1894–1896.
[25] H. Kapteyn, O. Cohen, I. Christov, and M. Murnane, *Harnessing attosecond science in the quest for coherent x-rays*, Science, 317 (2007), pp. 775–778.
[26] W. L. Kath, A. Mecozzi, P. Kumar, and C. G. Goedde, *Long-term storage of a soliton bit stream using phase-sensitive amplification: Effects of soliton-soliton interactions and quantum noise*, Opt. Commun., 157 (1998), pp. 310–326.
[27] J. N. Kutz, *Mode-locked soliton lasers*, SIAM Rev., 48 (2006), pp. 629–678, https://doi.org/10.1137/S0036144504446357.
[28] W. E. Lamb, *Theory of an optical maser*, Phys. Rev., 134 (1964), pp. A1429–A1450.
[29] J. Li, E. Spiller, and G. Biondini, *Noise-induced perturbations of dispersion-managed solitons*, Phys. Rev. A, 75 (2007), 053818.
[30] A. O. Lima, I. T. Lima, and C. R. Menyuk, *Error estimation in multicanonical Monte Carlo simulations with applications to polarization-mode-dispersion emulators*, J. Light. Technol., 23 (2005), pp. 3781–3789.
[31] B. S. Lindley and I. B. Schwartz, *An iterative action minimizing method for computing optimal paths in stochastic dynamical systems*, Phys. D, 255 (2013), pp. 22–30.
[32] R. O. Moore, *Trade-off between linewidth and slip rate in a mode-locked laser model*, Opt. Lett., 39 (2014), pp. 3042–3045.
[33] R. O. Moore, G. Biondini, and W. L. Kath, *Importance sampling for noise-induced amplitude and timing jitter in soliton transmission systems*, in Nonlinear Physics: Theory and Experiment, II (Gallipoli, 2002), World Scientific, River Edge, NJ, 2003, pp. 383–390.
[34] R. O. Moore, G. Biondini, and W. L. Kath, *A method to compute statistics of large, noise-induced perturbations of nonlinear Schrödinger solitons*, SIAM Rev., 50 (2008), pp. 523–549, https://doi.org/10.1137/080722977.
[35] R. O. Moore, T. Schäfer, and C. K. R. T. Jones, *Soliton broadening under random dispersion fluctuations: Importance sampling based on low-dimensional reductions*, Opt. Commun., 256 (2005), pp. 439–450.
[36] I. Nasieva, A. Kaliazin, and S. K. Turitsyn, *Multicanonical Monte Carlo modelling of BER penalty in transmission systems with optical regeneration*, Opt. Commun., 262 (2006), pp. 246–249.
[37] A. Owen and Y. Zhou, *Safe and effective importance sampling*, J. Amer. Statist. Assoc., 95 (2000), pp. 135–143.
[38] I. B. Schwartz, E. Forgoston, S. Bianco, and L. B. Shaw, *Converging towards the optimal path to extinction*, J. R. Soc. Interface, 8 (2011), pp. 1699–1707.
[39] E. T. Spiller and W. L. Kath, *A method for determining most probable errors in nonlinear lightwave systems*, SIAM J. Appl. Dyn. Syst., 7 (2008), pp. 868–894, https://doi.org/10.1137/070708123.
[40] T. R. Taha and M. J. Ablowitz, *Analytical and numerical aspects of certain nonlinear evolution equations. II. Numerical, nonlinear Schrödinger equation*, J. Comput. Phys., 55 (1984), pp. 203–230.
[41] M. J. Thorpe, D. Balslev-Clausen, M. S. Kirchner, and J. Ye, *Cavity-enhanced optical frequency comb spectroscopy: Application to human breath analysis*, Opt. Express, 16 (2008), pp. 2387–2397.
[42] T. Udem, R. Holzwarth, and T. W. Hänsch, *Optical frequency metrology*, Nature, 416 (2002), pp. 233–

237.
[43] E. Vanden-Eijnden and M. Heymann, *The geometric minimum action method for computing minimum energy paths*, J. Chem. Phys., 128 (2008), 061103.
[44] E. Veach, *Robust Monte Carlo Methods for Light Transport Simulation*, Ph.D. thesis, Stanford University, Stanford, CA, 1997.
[45] J. Ye and S. Cundiff, *Femtosecond Optical Frequency Comb: Technology, Principles, Operation and Application*, Springer, New York, 2005.