Morphological Computation in the Nonlinear Schrödinger Wavefield: Solitonic Dynamics Enable Noise-Tolerant, Self-Organizing Stability Inaccessible to Linear Systems

Anderson M. Rodriguez

Abstract

Physical reservoir computing exploits the transient dynamics of nonlinear media to process information with minimal training overhead. Recent work has demonstrated that wave-based systems governed by the nonlinear Schrödinger equation (NLSE) can serve as computational reservoirs, but the condition under which nonlinearity becomes functionally beneficial for robustness has yet to be isolated. Here we construct a classification task that encodes information exclusively in phase gradients with the expectation that linear dynamics diffuse this information. We compare the focusing NLSE against its linear counterpart (LSE) across a sweep of phase jitter, evaluating performance over multiple independent trials. Both reservoirs achieve near-perfect accuracy at low noise, but their behaviors diverge as jitter increases. At moderate noise, the LSE retains a slight advantage, indicating that nonlinearity imposes a computational cost. Beyond a crossover threshold, this reverses: in the extreme-noise regime, the NLSE achieves statistically significant superiority with large effect sizes, while the LSE decays more than twice as fast. We provide quantitative evidence that this robustness arises from nonlinear self-stabilization occurring past a threshold we measure. Here, the NLSE's dynamics spontaneously convert disordered phase differences into stable amplitude structures which we interpret as a realization of morphological computation relevant to the Artificial Life objective of identifying noise-tolerant computational substrates. These findings suggest general design principles for robust physical computing architectures and lifelike information processing in nonlinear media.

1 Introduction

Reservoir computing (RC) is a computational paradigm (Jaeger & Haas, 2004; Maass et al., 2002) where the reservoir is typically a fixed, high-dimensional dynamical system which is driven by an input signal, and only a linear (or otherwise simple) readout layer is trained to perform a task. Echo state networks (Jaeger, 2001) and liquid state machines (Maass et al., 2002) are canonical examples in the digital setting, while photonic (Brunner et al., 2013; Vandoorne et al., 2014) and other analog (Marcucci et al., 2020, 2023) implementations realize the same idea in physical media.[1] The core mechanism is the exploitation of nonlinear, high-dimensional dynamics to map input distinctions into linearly separable states, while retaining robustness to noise and perturbations (Jaeger & Haas, 2004).

While noise tolerance is inherent to such systems, it is not understood what functional necessity, the nonlinearity of the computational reservoir plays in robustness. Many of these systems exploit physical phenomena including nonlinear optical effects in waveguides, fibers, or integrated photonic circuits (Vandoorne et al., 2014). Nonlinear dispersive wave structures naturally generate rich spatiotemporal structure from simple initial conditions (Marcucci et al., 2020), and the nonlinear Schrödinger equation (NLSE) (Ablowitz & Segur, 1981; Zakharov & Shabat, 1972), for instance, governs light propagation in Kerr-nonlinear media (Kivshar & Agrawal, 2003), and supports solitons and soliton-like structures (Filippov, 2010) whose interactions can be sensitive to phase, amplitude, and spatial arrangement.

Recent work has established that wave systems governed by the nonlinear Schrödinger equation (NLSE[2]) can serve as computational reservoirs, both in theory (Marcucci et al., 2020) and in physical implementations via hydrodynamic substrates (Marcucci et al., 2023; Pierro et al., 2023) as well as discrete oscillator networks (Borlenghi et al., 2018; Silva et al., 2021)[3].

While the reservoir computing literature identifies nonlinearity as a prerequisite for rich dynamics (Lee et al., 2024), the conditions under which nonlinearity itself becomes functionally beneficial– i.e., its contribution to robustness–is, to the author's best knowledge, yet to be isolated and quantified by direct comparison to an appropriate linear system. This work addresses a fundamental question: what specific computational advantages do NLSE dynamics provide over the linear dispersive dynamics? Demonstrating computational expressivity of the NLSE wavefields contributes to the broader Artificial Life (ALife) program of identifying general substrates that support information processing, self-organization, and adaptive computation (Bedau, 2003; Gershenson, 2025). In this sense, solitonic reservoirs occupy a conceptual niche alongside chemical, morphological, and photonic computing systems (Lee et al., 2024), offering a mechanism by which coherent structures can emerge, interact, and compute without rigid, top-down architectural constraints.

To investigate, we construct a classification task where conventional static methods fail by design. We then directly compare the performance of a focusing NLSE reservoir to a dispersive linear Schrödinger Equation (LSE[4]) reservoir across a sweep of increasing input phase jitter (η). The observation that the LSE reservoir achieves perfect accuracy for low jitter ($\eta \leq 1.0$) confirms that the task is solvable via linear, time-dependent feature expansion. The central challenge, and the focus of this work, is computational stability under noise.

Prior NLSE reservoir work has demonstrated universality and physical implementations, but

---

[1]For a comprehensive review of reservoir computing methods and their physical implementations, see (Lukoševičius & Jaeger, 2009), and the extensive reviews in (Nakajima, 2020; Tanaka et al., 2019).

[2]Here, we use the focusing nonlinear Schrödinger equation with nonlinearity strength $g = 1$.

[3]See 1 at the end of the *Introduction*.

[4]Throughout the text we refer to the Schrödinger equation ($g = 0$) as the "LSE" as clarifying convention to emphasize the direct comparison with the NLSE.

has not isolated the functional advantage of nonlinearity via direct comparison to linear dynamics. Our contribution:

- First direct NLSE vs. LSE comparison under controlled noise

- Quantifies robustness margin (faster decay in LSE)

- Identifies threshold-driven mechanism via curvature analysis

- Relevance to the ALife search for general physical substrates capable of robust information processing and self-organized computation.

These findings are relevant beyond typical usage of physical or simulated variants of reservoir computing in that they demonstrate how substrate physics can provide computational robustness without architectural complexity. From an Artificial Life perspective, we ask a specific question: under what conditions do interference-driven nonlinear wavefields provide noise-tolerant information processing that could support lifelike, self-organizing computation, as opposed to merely amplifying signals? By directly comparing NLSE and LSE reservoirs under controlled noise, and showing a clear threshold at which nonlinear self-organization becomes beneficial, we identify a concrete computational advantage of a particular physical substrate. This transition threshold marks the onset of a regime we designate **Threshold-Driven Nonlinear Self-Stabilization** (TDNS), in which the nonlinear term preserves separability after the linear dynamics have failed.

This paper is structured as follows: We begin by defining a baseline static classifier[5] (Maass et al., 2002) that operates on initial, low-dimensional features and demonstrate that it achieves modest accuracy across the jitter sweep. We then show that while the LSE reservoir[6] experiences significant performance degradation as noise increases, the NLSE reservoir maintains a decisive advantage in the extreme-jitter regime ($\eta \geq 1.73$). Notably, at moderate noise levels, the LSE slightly outperforms the NLSE, revealing a computational cost to nonlinearity that is only recovered when noise severity triggers self-stabilizing dynamics. Quantitatively, the LSE exhibits a performance decay rate that is faster than the NLSE, with the NLSE achieving a peak superiority margin at later times.

The findings of this work show that the nonlinear dynamics–established here by phase-gradient driven flow fields measurable as interference patterns–are necessary for computational robustness in noisy environments. Importantly, our results do not imply that nonlinearity is universally necessary, but rather that it becomes functionally effective only once the noise regime exceeds the point where linear dynamics no longer maintain stable representations.

We demonstrate the underlying mechanism is TDNS, observable as **nonlinear self-stabilization occurring past a threshold**—a phenomenon where, beyond a critical nonlinearity, the accuracy as a function of the nonlinearity (represented by parameter weight g) exhibits positive curvature. This is interpreted as the nonlinear wavefield interference structures spontaneously converting noisy self-interactions into nonlinear dynamics that resist further degradation—a form of "computational self-focusing" intrinsic to the physics of the stable amplitude structures. We show that this property grants the NLSE access to a specific, high-noise regime that is inaccessible to the linear dynamic baseline.

---

[5]By "baseline static classifier," we refer to a linear classifier trained on a small set of global descriptors of the initial wavefield (such as mean phase statistics, mass, and gradient energy), with no access to features generated by dynamical evolution; "static" denotes the use of fixed, low-dimensional descriptors extracted at $t = 0$. See Section 3.1 for additional details.

[6]The linear readout was calibrated on a fixed subset of reservoir states and then evaluated on held-out states generated under independent noise realizations, ensuring that classification performance reflects physical structure produced by the dynamics rather than fitting to noise.

Table 1: Comparison of NLSE-based reservoir computing approaches

Work	Physics	Input Encoding	Task	Focus
Marcucci'20	1D NLSE	Time-varying $\psi(t)$	Universal approx.	Theory/Universality
Marcucci'23	KdV (Shallow Water)	Bias wave/signal	XNOR gate	Hardware/Boolean Logic
Pierro'23	KdV (Simulated)	$\psi(t)$	Optimization	Parameter tuning
Silva'21	Toda lattice (1D)	Time-varying forces	Pattern recog.	Discrete chain dynamics
Borlenghi'18	Discrete NLSE	Driven network	Pattern recog.	Network Dynamics
This work	**2D NLSE**	**Static Phase $\psi(x, y, 0)$**	**Phase/Jitter Class.**	**Functional Robustness (NLSE vs. LSE)**

2 Model and Reservoir Dynamics

2.1 Nonlinear Schrödinger dynamics

We consider a dimensionless focusing nonlinear Schrödinger equation on a two-dimensional square domain,

$$i\frac{\partial\psi}{\partial t} = -\frac{1}{2}\nabla^2\psi - g|\psi|^2\psi, \tag{1}$$

where $\psi(x, y, t) \in \mathbb{C}$ is the complex wavefield, $\nabla^2 = \partial_x^2 + \partial_y^2$ is the Laplacian, and $g > 0$ controls the focusing nonlinearity.

Although the $2D$ focusing NLSE is susceptible to wave collapse (Fibich, 2015), in our parameter regime and integration times the localized structures remain stable while demonstrating soliton-like behavior, which we deem sufficient for reservoir function and which is visually demonstrated in Figure 1.

In the simulations below we set the 'nonlinearity variable weight' $g = 1$ and work in units where the mass and time scales are dimensionless.

Equation (1) conserves several quantities in the continuous limit, including:

$$M[\psi] = \int |\psi|^2\, dx\, dy, \quad \text{(mass / norm)} \tag{2}$$

$$H[\psi] = \int \left(\frac{1}{2}|\nabla\psi|^2 - \frac{1}{2}|\psi|^4\right) dx\, dy, \quad \text{(Hamiltonian)}. \tag{3}$$

We discretize Eq. (1) on a 64×64 grid using a standard split-step Fourier method (Ablowitz & Segur, 1981) with periodic boundary conditions (see Appendix B for details). The method alternates linear propagation in Fourier space with nonlinear phase kicks in real space, preserving mass exactly; energy conservation is subject to small splitting error but remains accurate to numerical precision over the integration times used here.

2.2 Sech Bump Initial Conditions

To simulate reservoir dynamics we establish soliton-like initial conditions consisting of two localized sech bumps.

For each class $c \in \{0, 1, 2, 3\}$, we construct two localized sech bump centers—let $\mathbf{c}_j = (c_{j,x}, c_{j,y})$ denote the center of the j-th bump.

We initialize the field with a superposition of two radially localized sech profiles:

$$\psi_0(x, y) = \sum_{j=1}^{2} \text{sech}\left(\sqrt{(x - c_j^x)^2 + (y - c_j^y)^2}\right), \tag{4}$$

normalized to unit mass and multiplied by a plane-wave phase $\exp(i\mathbf{k}\cdot\mathbf{r})$ to impose a global phase gradient (wavevector). The two sech bumps are then superposed:

$$\psi_0^{(c)} = \psi_{\text{bump},1} + \psi_{\text{bump},2}. \tag{5}$$

The classes differ only in the global phase gradient wavevector applied to the same two-bump initial condition. This encodes class information into the initial velocity field, which determines the distinct interference and dynamical patterns generated during NLSE evolution. Encoding class information solely in the phase, while holding the amplitude envelope fixed, provides a stringent test of a physical reservoir because the system must convert fragile phase structure into observable spatial morphology during evolution.

2.3   Phase jitter as input noise

To test robustness, we inject random phase noise into the superposed initial conditions. After constructing $\psi_0^{(c)}$ as described above, we apply pointwise phase perturbations:

$$\psi_0^{(c,s)}(x,y) = \psi_0^{(c)}(x,y)\,\exp(i\,\delta\Phi(x,y))\,, \tag{6}$$

where $\delta\Phi(x,y)$ is sampled i.i.d. from a uniform distribution in $[-\eta, \eta]$ radians. This form of jitter preserves the gross spatial structure of the initial two-bump field while perturbing its phase gradients, providing a controlled test of robustness to phase-localized distortions. In addition to uniform phase jitter, a small amplitude perturbation (10% of η, Gaussian) was applied. This perturbation was included to prevent an artificially noise-free amplitude envelope under large phase jitter; it affects both reservoirs identically and does not introduce class-discriminative information.
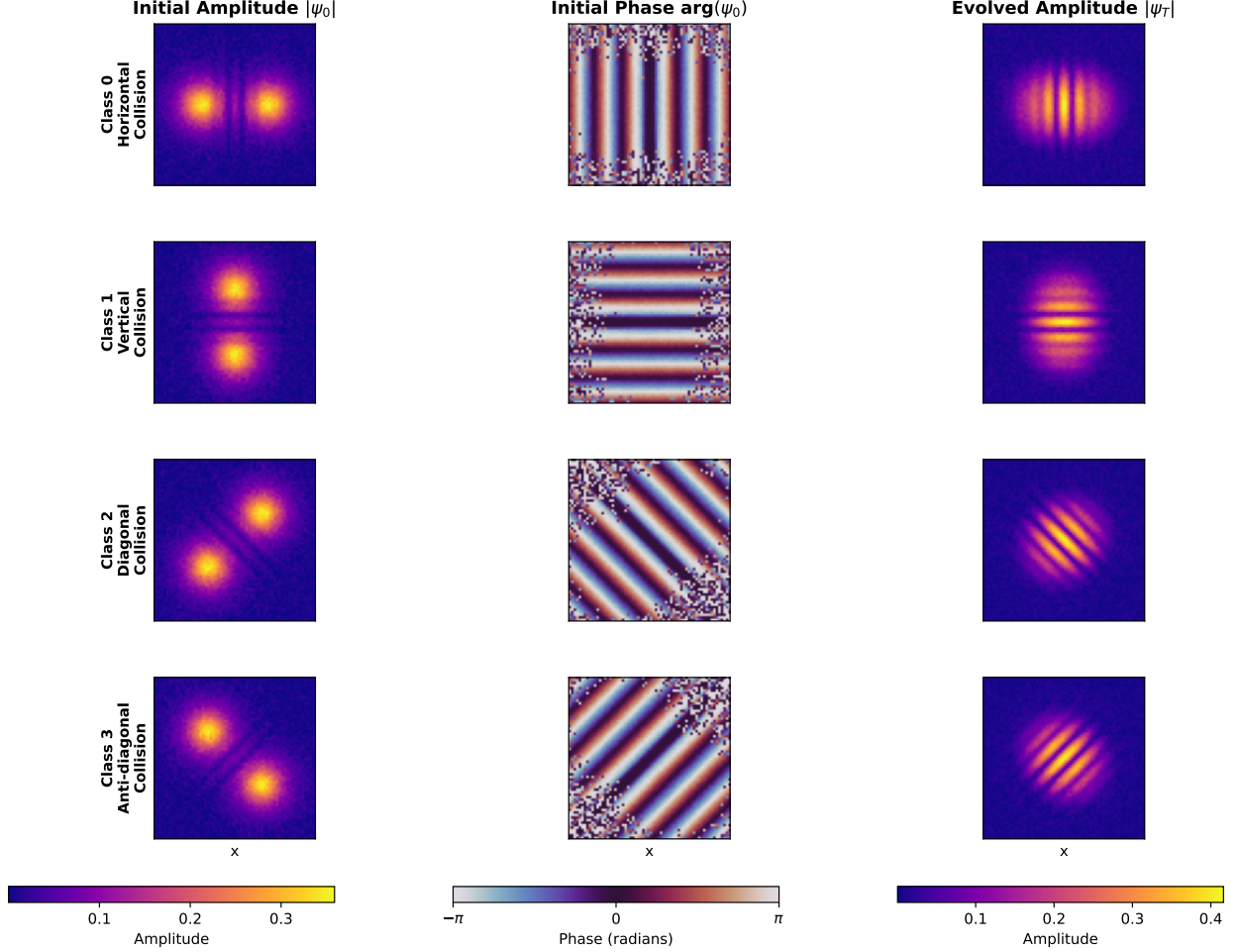
Figure 1: **Sech bump initial conditions encode class information through phase structure:** This figure illustrates representative dynamics on a reduced domain($L = 12$, $T = 0.6$); quantitative experiments use $L = 20$ with $T = 1.0$ for the main robustness analysis and $T = 2.0$ for the g-sweep (Appendix C). Each row shows one of four classes, differing only applied phase gradient (plane-wave wavevector). **Left:** Initial amplitude $|\psi_0|$ is nearly identical across classes—two localized bumps positioned to generate interference. **Center:** Initial phase $\arg(\psi_0)$ encodes class identity through plane-wave momentum imprints (horizontal, vertical, diagonal, anti-diagonal patterns plus jitter). **Right:** Evolved amplitude $|\psi_T|$ after NLSE propagation shows distinct patterns resulting from the phase-dependent interference geometry of the initial conditions. Despite shared initial amplitude profiles, the phase-controlled NLSE wave dynamics produce linearly separable amplitude signatures at time T, enabling perfect classification. Jitter level $\eta = 0.1$.

3 Feature Extraction and Classification

We compare two different ways to build feature vectors for classification: a control baseline that uses only simple global descriptors of the initial condition, and a reservoir readout that uses features derived from the evolved wavefield at time T.

We generate $N_{\text{class}} = 100$ samples per class, for a total of 400 samples in each experiment (for each jitter level). A random train–test split with 70% training and 30% testing is used throughout.

3.1 Control features from the initial condition (Static Baseline)

The control setting is designed to answer the question: How well can a classifier perform if it is only allowed simple, low-dimensional global features of the input, without access to the evolved field? To establish this necessary baseline, we define the **Static Baseline** classifier.

From each jittered initial condition $\psi_0^{(c,s)}$, we construct a 4-dimensional feature vector:

$$\boldsymbol{\phi}_0 = \big[\langle \cos\varphi \rangle, \langle \sin\varphi \rangle, M[\psi_0], K[\psi_0]\big], \tag{7}$$

where $\varphi(x,y) = \arg(\psi_0(x,y))$ is the phase, $\langle f \rangle$ denotes the spatial average, $M[\psi_0]$ is the total mass (norm), and $K[\psi_0]$ is the total gradient energy. The first two components encode the global phase alignment, while the latter two reflect total energy and mass (see Appendix B for full discretized definitions).

This feature includes no spatial information and no features of the evolved dynamics. The Static Baseline can only distinguish classes to the extent that they induce differences in these global descriptors, rigorously confirming that the input phase patterns are not linearly separable prior to the system's dynamic evolution.

3.2 Reservoir features from the evolved wavefield

In the reservoir setting, we evolve each initial condition under the NLSE dynamics to the final time T and extract features from the evolved wavefield $\psi_T^{(c,s)}$. Thus the reservoir features must already be linearly separable for classification to succeed.

Let $A_T(x,y) = |\psi_T(x,y)|$ denote the amplitude field.

To form the readout feature vector, we first downsample the A_T field from its original 64×64 grid resolution to 16×16 by subsampling (strided slicing), selecting every fourth pixel in both dimensions. This yields 256 spatial features. To this, we append the same four global descriptors (mass, gradient energy, and mean sine/cosine of phase), computed at time T, forming a final feature vector

$$\boldsymbol{\phi}_T \in \mathbb{R}^{260}. \tag{8}$$

The key point is that $\boldsymbol{\phi}_T$ is a high-dimensional, nonlinear embedding of the initial phase pattern, which is induced entirely by the wavefield dynamics. We train a simple linear classifier (logistic regression) on $\boldsymbol{\phi}_T$, thus leaving the reservoir dynamics and feature mapping fixed to perform the necessary nonlinear transformation.

3.3 Classifier and evaluation protocol

For both the Static Baseline and the reservoir feature sets, we use a multinomial logistic regression classifier with ℓ_2 regularization, as implemented in `scikit-learn` (Pedregosa et al., 2011). This choice keeps the readout as simple and linear as possible, in accordance with standard practice in

physical reservoir computing. The purpose is to ensure that all nonlinear processing capacity is attributed to the wavefield dynamics, not the final classifier.

The dataset for each jitter level consists of $N_\text{total} = 400$ samples ($N_\text{class} = 100$ per class). We split the data into 280 training and 120 test points (a 70/30 stratified split) and held this split fixed across all runs for reproducibility. Reported accuracies throughout this paper are **test accuracies**, which quantify the generalization performance of the trained classifier on unseen data.

4 Experimental Protocol

Both the NLSE and LSE reservoirs operate on features extracted from the evolved wavefield ψ_T, isolating the computational effect of the nonlinear term. To explore the differences in reservoir computational capabilities we run a series of experiments in which we vary the phase jitter η across 15 levels (from 0.10 to 2.00), keeping all other NLSE and discretization parameters fixed, with evolution time $T = 1.0$ for the main robustness analysis (parameter sensitivity to T and g is explored in Appendix C).

The jitter strength η was varied from 0.10 (low-noise baseline) to 2.00 (maximum stress level tested; performance is significantly impaired but above chance) using 15 linearly spaced points. The upper bound of $\eta = 2.00$ was selected to define the physical failure limit of the reservoir, as this value corresponds to a phase uncertainty of approximately $\pm 115°$ (compared to a full 2π phase cycle).

For each value of η, the experimental procedure is:

1. For each statistical trial $s \in \{1, \ldots, 10\}$ (seeded independently):

   (a) Generate $N_{\mathrm{class}} = 100$ jittered initial conditions $\psi_0^{(c,s)}$ for each class using the unique random state s.

   (b) Compute Static Baseline features $\phi_0^{(c,s)}$.

   (c) Evolve each $\psi_0^{(c,s)}$ independently under LSE ($g = 0$) and NLSE ($g = 1$).

   (d) Compute $\phi_T^{(c,s)}$ reservoir features.

2. Concatenate the resulting features into three datasets: $X_{\mathrm{Static}} \in \mathbb{R}^{400 \times 4}$, $X_{\mathrm{LSE}} \in \mathbb{R}^{400 \times 260}$, and $X_{\mathrm{NLSE}} \in \mathbb{R}^{400 \times 260}$, with corresponding label vector $\boldsymbol{y} \in \{0, 1, 2, 3\}^{400}$.

3. Split the datasets into 280 training and 120 test points (a 70/30 stratified split).

4. Train a logistic regression classifier on each of the three feature sets and evaluate each on the held-out test set.

5. Record the test accuracies. The final results are reported as the **mean test accuracy** μ and **standard deviation** σ calculated across the $N = 10$ independent statistical trials.

The full sweep covers 15 iterations of phase jitter perturbations from $\eta = 0.10$ to $\eta = 2.00$ (see Table 2).

As shown in Figure 1, while the initial amplitude profiles are nearly identical across classes, the phase-driven interference dynamics produce distinct spatiotemporal patterns that are readily distinguishable by the reservoir readout.

5 Statistical Methodology

5.1 Accuracy comparisons and effect sizes

For each jitter level we ran $N = 10$ statistically independent trials per reservoir (distinct random seeds). To quantify the divergence between the NLSE and LSE reservoirs in the high-noise region, we restricted multiple-comparison correction to the seven highest jitter values, where robustness is of primary interest,

$$\eta \in \{1.1857, 1.3214, 1.4571, 1.5929, 1.7286, 1.8643, 2.0000\}.$$

For each η in this set, we computed a two-sample t-statistic for the difference in mean test accuracy (NLSE–LSE) using the empirical standard deviations over seeds (degrees of freedom df $= 18$), and obtained two-tailed p-values from the Student t-distribution. Because NLSE and LSE trials share random seeds, the design is paired; however, we report independent-samples t-tests samples statistics without exploiting this pairing as our goal is to quantify a robust regime-level separation rather than optimize inference under the paired structure. A paired test would be a natural alternative and may have higher power. Two-tailed tests were chosen because we sought to detect significant differences in either direction—both NLSE superiority at extreme noise and potential LSE advantages at moderate noise. This bi-directional approach helped to enable identification of the crossover threshold where the performance relationship between the NLSE and LSE reverses.

To control the family-wise error rate across these seven planned comparisons in the high-noise regime, we used a Bonferroni correction (Bonferroni, 1936) with a nominal $\alpha = 0.01$, yielding

$$\alpha_{\mathrm{corr}} = \frac{0.01}{7} \approx 0.0014.$$

This conservative choice is appropriate here because the number of pairwise contrasts is small and our primary goal is to avoid false positives when claiming a robustness advantage in the most challenging (high-noise) conditions. Significance was assessed by comparing the raw two-tailed p-values for each contrast to the Bonferroni-adjusted threshold $\alpha_{\mathrm{corr}} = 0.0014$.

Effect sizes are reported as Cohen's d (Cohen, 1988), computed using the pooled standard deviation

$$s_{\mathrm{pooled}} = \sqrt{\frac{\sigma^2_{\mathrm{NLSE}} + \sigma^2_{\mathrm{LSE}}}{2}},$$

where σ_{NLSE} and σ_{LSE} are the sample standard deviations across the $N = 10$ seeds.

In the extreme-jitter regime ($\eta \geq 1.7286$), the NLSE significantly outperforms the LSE under the Bonferroni-corrected threshold. Specifically:

$\eta = 1.7286 :\quad \Delta\mathrm{Acc} = +0.0708,\ t(18) = 4.85,\ p = 1.29 \times 10^{-4},\ d = +2.17\ (< \alpha_{\mathrm{corr}});$

$\eta = 1.8643 :\quad \Delta\mathrm{Acc} = +0.1108,\ t(18) = 7.25,\ p = 1.0 \times 10^{-6},\ d = +3.24\ (< \alpha_{\mathrm{corr}});$

$\eta = 2.0000 :\quad \Delta\mathrm{Acc} = +0.1317,\ t(18) = 6.31,\ p = 6.0 \times 10^{-6},\ d = +2.82\ (< \alpha_{\mathrm{corr}}).$

These correspond to large effects and establish a robustness margin for the nonlinear reservoir at high noise.

For completeness, we note that in the moderate-noise subset the linear reservoir retains an advantage:

$\eta = 1.1857 :\quad \Delta\mathrm{Acc} = -0.0467,\ t(18) = -6.10,\ p = 9.0 \times 10^{-6},\ d = -2.73\ (< \alpha_{\mathrm{corr}});$

$\eta = 1.3214 :\quad \Delta\mathrm{Acc} = -0.0575,\ t(18) = -6.44,\ p = 5.0 \times 10^{-6},\ d = -2.88\ (< \alpha_{\mathrm{corr}});$

$\eta = 1.4571 :\quad \Delta\mathrm{Acc} = -0.0400,\ t(18) = -3.66,\ p = 1.791 \times 10^{-3},\ d = -1.64\ (\text{n.s. after Bonferroni});$

$\eta = 1.5929 :\quad \Delta\mathrm{Acc} = +0.0050,\ t(18) = 0.39,\ p = 0.703,\ d = +0.17\ (\text{n.s.}).$

Together, these results quantify the crossover: a slight linear advantage at moderate noise, followed by a statistically reliable and substantial nonlinear advantage once the jitter is severe enough to trigger self-stabilizing dynamics where, at the highest jitter level ($\eta = 2.0$), the NLSE's advantage of $+13.17$ percentage points corresponds to a large effect size (Cohen's $d \approx 2.8$), confirming that the robustness margin is both statistically significant and substantial.

5.2 Decay rate analysis in the high-noise regime

To complement the comparison, we quantified the overall difference in performance degradation between the NLSE and LSE reservoirs by estimating the decay rate (β) of their accuracy curves in the critical high-noise regime.

The decay rate β is defined as the slope of a linear regression of mean test accuracy (μ) versus jitter strength (η),

$$\mu(\eta) = \beta\eta + C,$$

fit over the regime

$$1.05 \leq \eta \leq 2.00,$$

corresponding to indices $\eta_{\text{idx}} = 7$ to 14 (inclusive) in the full sweep, where the LSE performance begins its monotonic decline (the onset of degradation). This yielded

- **LSE reservoir ($g = 0.0$):** $\beta_{\text{LSE}} \approx -0.3903$,

- **NLSE reservoir ($g = 1.0$):** $\beta_{\text{NLSE}} \approx -0.1755$.

We summarize the relative steepness as

$$\text{Relative Steepness} = \left(\frac{\beta_{\text{LSE}}}{\beta_{\text{NLSE}}} - 1 \right) \times 100\% \approx \mathbf{122.4\%},$$

indicating that the LSE accuracy decays approximately 122.4% faster than the NLSE accuracy in this high-jitter regime. Decay rates were estimated by linear regression on mean test accuracy as a function of jitter η over the specified regime. This regression-based measure provides a global, complementary quantification of the NLSE's superior robustness margin beyond the pointwise statistical comparisons.

6 Results

6.1 Initial Validation: The Nontrivial Classification Task

Before introducing jitter, we validate that the synthetic class definition is not trivially separable by low-dimensional statistics. The four classes are defined to share the same initial amplitude envelope and differ only in phase structure (global phase-gradient orientation). The objective is to confirm that the classification task requires the computational work of the wavefield's dynamics.

At $\eta = 0.1$ (no jitter), the Static Baseline classifier (four global features) achieves a test accuracy of 0.5133. In contrast, the LSE and NLSE reservoir classifiers (utilizing the evolved amplitude field) both achieve 100% accuracy. The Static Baseline struggles to resolve all four classes, entirely confusing at least one pair. This confirms that (a) the classes are not trivial to separate using simple global features of the initial condition, and (b) the evolved wavefield does indeed encode the class-specific information in a linearly separable way (Figure 1). As shown in Figure 1, while the initial amplitude profiles are nearly identical across classes, the phase-driven interference dynamics produce distinct interference patterns that are readily distinguishable by the reservoir readout.

6.2 Robustness Sweep: Quantifying the Informational Margin

We varied the phase jitter η across 15 levels from 0.10 to 2.00 and repeated the training and evaluation protocol for the three feature sets: the static control, the LSE reservoir ($g = 0.0$), and the NLSE reservoir ($g = 1.0$). Table 2 summarizes the resulting test accuracies.

Table 2: **Test accuracy (Mean $\pm$ Std. Dev.) as a function of phase jitter η.** Results are computed over $N = 10$ independent statistical seeds. Both reservoirs achieve near-perfect accuracy at low noise. A crossover occurs near $\eta \approx 1.6$: LSE maintains advantage at moderate noise ($\eta \approx 1.2$–1.5) before NLSE achieves significant superiority at extreme jitter ($\eta \geq 1.73$). For each jitter value, the table reports the mean test classification accuracy ($\mu \pm \sigma$) across 10 independently simulated reservoir realizations evaluated on the same fixed train/test split.

Jitter η	Static Control	LSE Reservoir ($g = 0.0$)	NLSE Reservoir ($g = 1.0$)
0.1000	0.513 ± 0.054	1.000 ± 0.000	0.996 ± 0.006
0.2357	0.496 ± 0.041	1.000 ± 0.000	0.993 ± 0.007
0.3714	0.473 ± 0.041	0.993 ± 0.008	0.993 ± 0.006
0.5071	0.398 ± 0.033	0.999 ± 0.003	0.988 ± 0.010
0.6429	0.360 ± 0.043	0.999 ± 0.003	0.988 ± 0.013
0.7786	0.348 ± 0.044	1.000 ± 0.000	0.977 ± 0.009
0.9143	0.343 ± 0.044	0.999 ± 0.003	0.972 ± 0.013
1.0500	0.358 ± 0.047	0.998 ± 0.003	0.954 ± 0.019
1.1857	0.366 ± 0.049	0.994 ± 0.005	0.948 ± 0.024
1.3214	0.358 ± 0.033	0.978 ± 0.010	0.921 ± 0.026
1.4571	0.358 ± 0.036	0.947 ± 0.021	0.907 ± 0.027
1.5929	0.358 ± 0.039	0.891 ± 0.031	0.896 ± 0.026
1.7286	0.353 ± 0.040	0.808 ± 0.041	0.878 ± 0.021
1.8643	0.340 ± 0.053	0.727 ± 0.037	0.838 ± 0.031
2.0000	0.326 ± 0.055	$\mathbf{0.635 \pm 0.036}$	$\mathbf{0.767 \pm 0.055}$

Notably, at low jitter ($\eta \leq 0.78$), the LSE reservoir achieves marginally higher accuracy than the NLSE (e.g., 100.0% vs 99.6% at $\eta = 0.10$). This suggests the nonlinear term introduces a small computational cost when noise levels are insufficient to trigger the self-stabilization mechanism. The NLSE advantage emerges only when the noise regime is severe enough to destabilize linear dynamics.

The results show a performance divergence between the NLSE and LSE reservoirs past the threshold $\eta_{\text{break}} \approx 1.6$. Figure 2 illustrates this margin of superiority. The LSE reservoir moves toward failure, while the NLSE reservoir maintains computational utility deep into the high-noise regime. Specifically, at the extreme jitter limit ($\eta = 2.00$), the NLSE achieves $76.7\% \pm 5.5\%$ accuracy, significantly outperforming the LSE, which achieves $63.5\% \pm 3.6\%$.
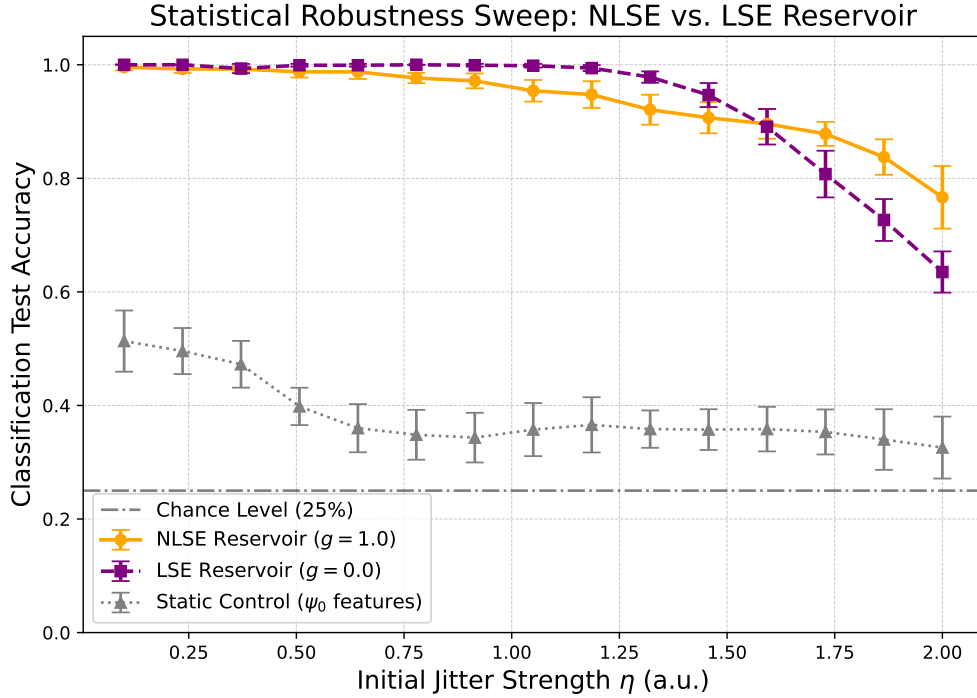

Figure 2: **NLSE (Nonlinear) vs. LSE (Linear) Reservoir Robustness to Input Jitter (η).** The curves represent the **mean test accuracy (μ)** computed across $N = 10$ independent statistical trials. Vertical bars indicate the **standard deviation (σ)**. Both systems perform well at low noise ($\eta \leq 1.05$) but LSE shows a flawless 1.00 performance in the low-noise regime while the NLSE demonstrates a slight performance cost (0.996). The NLSE (orange solid line) maintains a statistically significant performance advantage over the LSE (blue dashed line) in the **high-noise regime ($\eta > 1.6$)**, where the LSE performance begins to collapse in measured accuracy (black dot-dash line).

It is important to note that as the jitter increases, the LSE reservoir rapidly degrades toward a 63.5% accuracy level. In contrast, the NLSE reservoir maintains a consistent performance advantage, particularly in the range $\eta > 1.6$, demonstrating its ability to sustain computational utility by actively maintaining class separability above this baseline.

To quantify the functional advantage of the nonlinearity, we analyzed performance in the extreme-noise region ($\eta \geq 1.73$), where NLSE superiority is statistically significant.
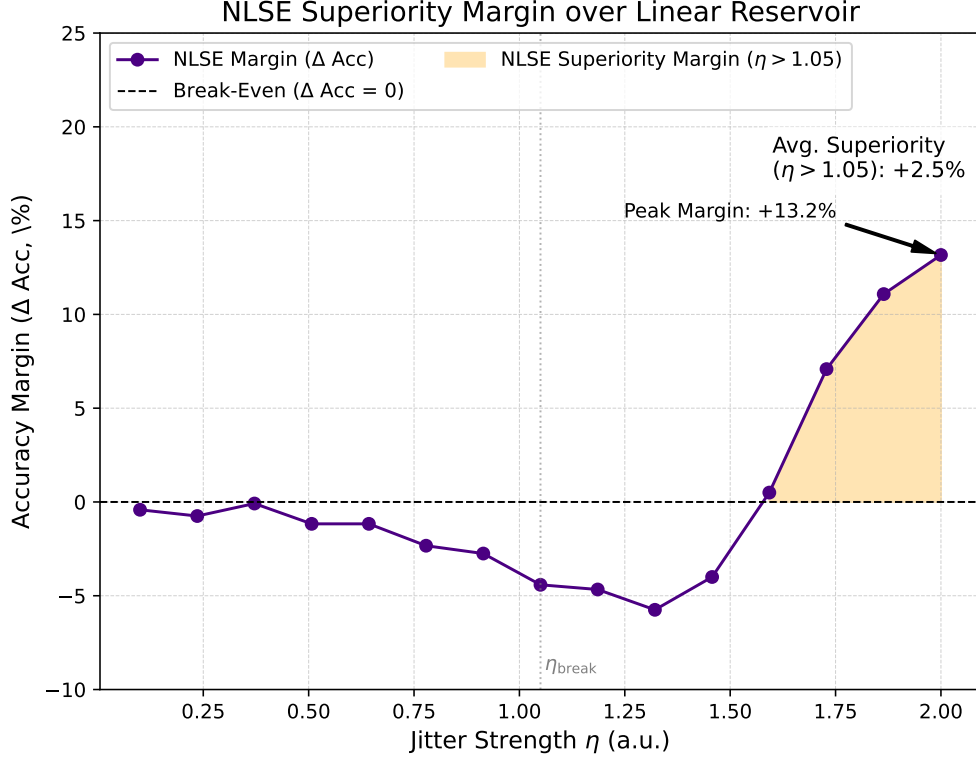
Figure 3: **NLSE Superiority Margin over the Linear Reservoir.** The curve displays the difference in mean test accuracy, $\Delta\mathrm{Acc} = \mathrm{Acc_{NLSE}} - \mathrm{Acc_{LSE}}$, as a function of phase jitter η. The gray dashed line marks the threshold crossover $\eta_{\mathrm{cross}} \approx 1.6$. In the moderate-noise regime ($\eta \approx 1.2$–1.5), the LSE maintains a slight advantage. At extreme jitter ($\eta \geq 1.73$), the NLSE achieves decisive superiority, with a peak margin at $\eta = 2.00$.

Peak Margin: The maximum performance divergence (Max $\Delta\mathrm{Acc}$) occurs at the highest tested jitter, $\eta = \mathbf{2.00}$, where the NLSE sustains a superiority margin of $\mathbf{13.17}$ percentage points over the LSE.

Statistical Significance: To assess statistical reliability, we computed t-statistics for the difference in mean accuracy between NLSE and LSE with Bonferroni correction for 7 comparisons (adjusted $\alpha = 0.0014$). In the moderate-noise regime ($\eta \approx 1.2$–1.5), the LSE maintains a small but significant advantage, indicating that the nonlinear term imposes a computational cost when noise is insufficient to trigger self-stabilization. At extreme jitter ($\eta \geq 1.73$), this reverses: the NLSE achieves significant superiority (all $|t| > 4.8$, $p < 0.001$), with Cohen's $d \approx 2.8$ at $\eta = 2.0$. The crossover at $\eta \approx 1.6$ marks the threshold where nonlinear self-organization becomes beneficial.

Utility Threshold: We define the critical jitter for utility (η_c) as the maximum noise level at which the mean accuracy remains above $\mathbf{70}\%$. The NLSE extends computational utility to $\eta_{\mathbf{c}} = \mathbf{2.00}$ units, whereas the LSE's utility collapses earlier at $\eta_{\mathbf{c}} = \mathbf{1.8643}$. The NLSE thus extends the operational regime by $\mathbf{0.1357}\ \eta$ units.

Decay Rate Analysis: We quantified the robustness by calculating the rate of performance loss across the degradation regime ($\eta \geq 1.05$). A linear regression fit reveals that the LSE reservoir decays at a rate ($\beta_{\mathrm{LSE}} \approx -\mathbf{0.3903}$) that is $\mathbf{122.4}\%$ faster than the NLSE decay rate ($\beta_{\mathrm{NLSE}} \approx -\mathbf{0.1755}$). This is quantitative evidence for the superior noise tolerance of the nonlinear

physics.

In summary, the sustained performance advantage of the NLSE, driven by the nonlinear term $g|\psi|^2\psi$, demonstrates that the nonlinear dynamics established by the interference of the initial wavefield components are crucial for maintaining computational stability under high perturbation levels. The fundamental mechanism behind this stability is further explored in Section C through second derivative analysis.

7 Discussion

This finding exemplifies morphological computation in the sense introduced by Pfeifer et al. (Pfeifer et al., 2007): the physical dynamics of the medium itself perform information processing that simplifies the downstream readout task. In contrast to classical morphological computation examples where a robot body or chemical medium performs physical preprocessing (Hauser & Hughes, 2024; Hauser et al., 2012; Lee et al., 2024; Pfeifer et al., 2007), the NLSE wavefield's intrinsic self-focusing dynamics of the phase-structured initial superposition act as the "morphology"—not a static body, but a dynamic substrate whose nonlinear evolution spontaneously converts fragile phase-encoded information into robust amplitude structure, offloading computational work from the classifier to the physics. This principle aligns with recent demonstrations of task-adaptive physical reservoir computing in magnetic systems Lee et al., 2024, where phase-tunable spin-wave spectra reconfigure reservoir properties to maintain performance across diverse tasks, underscoring the broader role of substrate physics in achieving noise robustness in physical reservoirs.

7.1 Nonlinearity Provides Computational Robustness

The experiments show that while linear wave dynamics (LSE, $g = 0.0$) can initially separate the phase-encoded classes, the nonlinear dynamics (NLSE, $g = 1.0$) are functionally necessary to preserve that separability under high-noise conditions. The divergence in performance past the crossover $\eta_{\text{cross}} \approx 1.6$ demonstrates that the energy-conserving nonlinear dynamics of the NLSE create a more stable, non-contracting representation of the input.

The crossover behavior observed at $\eta \approx 1.6$—where NLSE transitions from slight deficit to decisive advantage—provides direct evidence that the nonlinear term is not universally beneficial, but rather activates a specific self-stabilization mechanism that becomes necessary only when linear dynamics can no longer maintain feature separability.

The LSE's rapid failure past $\eta_{\text{break}} \approx 1.6$ demonstrates that this regime is indeed too noisy for a purely dispersive wave system to retain feature separability; linear dispersion diffuses the class information, which is quantitatively demonstrated by the LSE's **122.4%** faster decay rate ($\beta_{\text{LSE}} \approx -0.3903$). We refer to the NLSE's capacity to maintain separability in this regime as **Threshold-Driven Nonlinear Self-Stabilization**: once noise exceeds the level where linear dynamics fail, the nonlinear term acts as a computational filter, allowing the system to persistently map noisy input manifolds into a high-dimensional, but non-collapsing, feature space. This property grants the NLSE reservoir access to a computational regime that is closed off to linear wave physics.

The functional superiority of the NLSE reservoir in the high-noise regime is consistent with morphological computation principles (Hauser & Hughes, 2024; Hauser et al., 2012). The NLSE's dynamics—governed by the competition between dispersion and Kerr nonlinearity—is not merely a feature, but a computational mechanism that actively sculpts the wavefield into stable, class-specific amplitude patterns, resisting the entropic mixing caused by high jitter. Furthermore, the positive time derivative ($\text{d(Acc)}/\text{dT} > 0$) confirms this stability is sustained and beneficial throughout the evolution time T, fulfilling a core tenet of the ALife search for minimal, noise-tolerant computational substrates(Bedau, 2003).

7.2 Isolating the Functional Advantage of NLSE Dynamics

Prior NLSE reservoir computing research has focused on three themes: (i) universal approximation and nonlinearity thresholds (Marcucci et al., 2020), (ii) physical implementations in hydrodynamic (Marcucci et al., 2023; Pierro et al., 2023) and discrete oscillator (Borlenghi et al., 2018; Silva et al., 2021) substrates, and (iii) optimization of encoding and readout parameters.

Our contribution is complementary: rather than demonstrating universality or building hardware, we identify what NLSE dynamics compute best: robustness against noise. Our methodological choice centers on functional isolation: By constructing a direct comparison between the NLSE and its linear counterpart (LSE) under increasing phase jitter, we isolate the functional advantage of nonlinear self-focusing dynamics. This intra-system comparison ($g = 1.0$ vs. $g = 0.0$) is critical because it attributes the observed computational stability to the emergent physics of the $g|\psi|^2\psi$ term, rather than confounding factors inherent in engineered systems (e.g., specific network topologies or time-delay feedback loops). This approach provides a foundational understanding of the medium itself, which is a core requirement for the ALife goal of identifying self-organizing computational substrates.

Specifically, we demonstrate that nonlinear dynamics, established autonomously by interference-driven self-focusing of the localized wavepackets, are necessary to maintain computational utility in the high-noise regime ($\eta > 1.6$). A parameter sweep over $g \in [0, 1]$ (Section C) revealed that this robustness is not a continuous, increasing function of the nonlinearity, but is rather governed by a nonlinear self-stabilization mechanism occurring past a threshold. This quantitatively demonstrated by the positive second derivative ($\mathrm{d}^2(\mathrm{Acc})/\mathrm{d}g^2 > 0$) observed in the critical g regime. To our knowledge, prior NLSE reservoir computing work has not explicitly isolated this robustness mechanism, with existing frameworks either targeting function approximation or requiring physically engineered bias elements (e.g., stationary KdV solitons (Marcucci et al., 2023)) or external parameter tuning (Pierro et al., 2023).

The spontaneous transformation of noisy phase geometry into a linearly separable spatial amplitude morphology is the underlying process by which the NLSE provides its margin of superiority. The linear system quickly succumbs to entropic mixing at high jitter, whereas the NLSE's self-stabilizing nature preserves the class distinctions, a critical requirement for a practical computational medium.

7.3 Morphological Mechanism: Threshold-Driven Nonlinear Self-Stabilization

The functional regime-superiority of the NLSE reservoir is a direct realization of **morphological computation**, where the physics itself performs the necessary noise-filtering transformation. The robust performance exhibited by the NLSE reservoir in the Critical Regime ($\eta = 2.0$, $T = 2.0$) is achieved through interference driving nonlinear dynamics which lead to self-focusing stabilization past the threshold of ($\eta > 1.6$). The positive mean derivative $d(\mathrm{Acc})/dT = +0.3389$ (from the time sweep) confirms that this g-driven structural stability is sustained and beneficial throughout the evolution time T, preventing the features from collapsing.

The mechanism is consistent with threshold behavior, not a continuous linear process. As demonstrated by the g-sweep (Figure 4), the performance is not a smooth, linear function of g. While the overall mean $d^2\mathrm{Acc}/dg^2$ is (-0.5660) due to saturation (overall diminishing returns), the curve exhibits a sharp concave-up region in the critical range of $g \in [0.3, 0.7]$. The positive local second derivative, peaking at $+1.2500$ at $g \approx 0.3$, confirms a **threshold-driven acceleration** in accuracy. This is the physical signature of the nonlinear self-stabilization mechanism: the g term acts as a continuous control parameter, driving the wavefield's morphology from a rapidly diffusing state to a stable, self-focused, localized amplitude pattern once the nonlinearity overcomes the critical noise threshold ($\eta > 1.6$).

The linear system, conversely, lacks this mechanism entirely. The LSE fails because its dynamics are constrained to diffusion, erasing the encoded phase gradients over the necessary evolution time T. The NLSE succeeds because the nonlinear dynamics of its morphology compute the required stability, converting disordered phase differences into classification-preserving amplitude patterns–a

configuration that is demonstrably stable and improving across the evolution time, as shown by the positive $d(\text{Acc})/dT$ result.

7.4 Why Phase-only Encoding as a Test

Encoding class information purely in the initial phase, while keeping the amplitude envelope fixed, is a stringent test for any physical reservoir:

- Many physical observables (e.g. intensity measurements) are primarily sensitive to amplitude. A system that relies only on instantaneous amplitude differences would fail this test.

- In the present system, phase controls interference between the localized sech-profiled wavepackets. The subsequent nonlinear dynamics convert these phase differences into amplitude and gradient structure, which can be observed at time T.

- The control classifier, with only four global features, does not see enough of this latent structure to perfectly recover the labels.

That the reservoir readout achieves 100% accuracy for moderate jitter levels and high robustness otherwise therefore provides strong evidence that the nonlinear wavefield is functioning as a genuine computational substrate, not merely as a trivial feature expander.

7.5 Robustness, Jitter, and Noise

The robustness sweeps shown in Table 2 demonstrate that the NLSE's self-stabilizing dynamics create a high degree of noise tolerance. This suggests that the class-specific manifolds in the evolved state space are relatively stable under small perturbations of the initial condition. Although the NLSE is a conservative Hamiltonian system and formally lacks attractors, the four classes map to distinct, well-separated regions of the reservoir state space, such that jittered initial conditions remain within their respective class boundaries. Of course, in more complex settings and with more realistic noise models, perfect accuracy may not persist. The current results should be viewed as a proof-of-concept in a controlled synthetic environment. Nonetheless, they illustrate that the nonlinear dynamics of wavefields established by self-focusing interference patterns can exhibit precisely the kind of robust, noise-tolerant separation of inputs that reservoir computing seeks to exploit.

7.6 Limitations and Future Directions

The present work has several limitations that also point toward natural extensions:

- **Synthetic Task:** The classification task is synthetic and low-dimensional (four classes). An obvious next step is to couple the NLSE reservoir to more realistic inputs, or to embed higher-dimensional data (e.g. images) into phase or amplitude patterns and test more challenging tasks.

- **Multi-time Readouts:** We only use features extracted at a single final time T. In many reservoir computing applications, one exploits the full temporal evolution (e.g. by concatenating features across time). Investigating multi-time readouts will be addressed in future work.

- **Comparison to Other Reservoirs:** While we included a simple control baseline, we did not compare the NLSE reservoir to other standard reservoir models (e.g. echo state networks

19

or random recurrent networks) on the same synthetic task. Such comparisons would clarify what is unique (or not) about interference-driven nonlinear wavefields.

This work uses synthetic four-class phase patterns as a controlled testbed; extensions to real-world data and temporal processing are natural next steps. More broadly, one could imagine extending this framework to: incorporate dissipation or driving, turning the NLSE into a driven-dissipative system closer to experimental physical reservoirs; use more biologically or physically motivated initial conditions (e.g. experimental imaging data mapped into phase and amplitude); study memory capacity and temporal processing capabilities by feeding in time-varying inputs rather than one-shot initial conditions.

8 Conclusion

This work isolated the functional advantage of nonlinear dynamics in physical reservoir computing. By constructing a direct comparison between the NLSE reservoir ($g = 1.0$) and its linear counterpart (LSE, $g = 0.0$), we established the central thesis: dynamical nonlinearity is not just a component but a key factor driving computational robustness once the noise level exceeds the linear regime's stability threshold.

The quantitative evidence supporting this functional necessity is strong. In the high-noise regime ($\eta > 1.6$), the NLSE maintains superiority, with statistically significant advantages emerging at extreme noise ($\eta \geq 1.73$). The LSE exhibits a decay rate that is 122.4% faster than the NLSE, which sustains a peak superiority margin of 13.17 percentage points. Statistical significance was confirmed via t-tests, with effects at the highest jitter levels ($\eta \geq 1.73$) surviving Bonferroni correction (all $|t| > 4.8$, p < 0.001)[7]. This divergence demonstrates the functional boundary between linear diffusion and nonlinear self-organization.

These metrics constitute the quantitative signature of the specific mechanism which we term Threshold-Driven Nonlinear Self-Stabilization (TDNS). An initially recognized implication of TDNS is that there exists a regime of physical dynamics where nonlinear evolution preserves information that linear evolution inevitably destroys. As evidenced by our second derivative analysis ($\mathrm{d}^2(\mathrm{Acc})/\mathrm{d}g^2 > 0$), the system undergoes dynamics consistent with threshold behavior where the balance between dispersion and nonlinearity causes the wavefield to accelerate into stable self-focusing localized amplitude patterns. We have likewise shown that the LSE significantly fails to maintain computational efficacy in this exact regime where the NLSE sees an acceleration followed by a shift in behavior. Taken together, this heavily suggests a nonlinear dynamical mechanism is the physical means by which the reservoir maintains informational computability by actively converting disordered phase distinctions into linearly separable, structured amplitude patterns.

The ability of the NLSE to access this high-noise regime provides a physical example of a system whose computational boundaries are defined by its capacity for self-organization, independent of complex architectural design. This finding establishes a crucial conceptual foundation for designing robust physical computing architectures and is highly relevant to understanding emergent, noise-tolerant information flow in dynamic physical and biological systems.

Further, our work provides a concrete, mechanism-level example of morphological computation in a continuous physical system. Here, the NLSE reservoir offers a tractable case of a threshold transition from a linear diffusion computational regime to a nonlinear self-stabilization computational regime which can be identified, quantified, and experimentally reproduced. These results point toward potential implementations in photonic, hydrodynamic, and other nonlinear physical substrates where threshold-driven stabilization may be exploited for robust computation.

---

[7]We used $\alpha = 0.01$ to be conservative given the small number of planned high-noise contrasts and the strength of the mechanistic claim.

Data and Code

Data supporting the findings of this work are available from the author upon reasonable request. Simulation code can be found at:

https://github.com/amr28693/solitonic_reservoir

References

Ablowitz, M. J., & Segur, H. (1981). *Solitons and the inverse scattering transform*. SIAM.

Bedau, M. A. (2003). Artificial life: Organization, adaptation and complexity from the bottom up. *Trends in Cognitive Sciences*, *7*(11), 505–512. https://doi.org/https://doi.org/10.1016/j.tics.2003.09.012

Bonferroni, C. E. (1936). *Teoria statistica delle classi e calcolo delle probabilità*. Reale Istituto Superiore di Scienze Economiche e Commerciali di Firenze.

Borlenghi, S., Boman, M., & Delin, A. (2018). Modeling reservoir computing with the discrete nonlinear schrödinger equation. *Phys. Rev. E*, *98*, 052101. https://doi.org/10.1103/PhysRevE.98.052101

Brunner, D., Soriano, M. C., Mirasso, C. R., & Fischer, I. (2013). Parallel photonic information processing at gigabyte per second data rates using transient states. *Nature Communications*, *4*(1), 1364. https://doi.org/10.1038/ncomms2368

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd). Lawrence Erlbaum Associates.

Fibich, G. (2015). *The nonlinear schrödinger equation: Singular solutions and optical collapse*. Springer. https://doi.org/10.1007/978-3-319-12748-4

Filippov, A. T. (2010). *The versatile soliton*. Birkhäuser. https://doi.org/https://doi.org/10.1007/978-0-8176-4974-6

Gershenson, C. (2025). Complexity, artificial life, and artificial intelligence. *Artificial Life*, *31*(3), 289–303. https://doi.org/https://doi.org/10.1162/artl_a_00462

Hauser, H., & Hughes, J. (2024). Morphological computation—past, present and future. *Device*, *2*(9), 100439. https://doi.org/https://doi.org/10.1016/j.device.2024.100439

Hauser, H., Ijspeert, A. J., Füchslin, R. M., Pfeifer, R., & Maass, W. (2012). The role of feedback in morphological computation with compliant bodies. *Biological Cybernetics*, *106*, 595–613. https://doi.org/https://doi.org/10.1007/s00422-012-0516-4

Jaeger, H. (2001). The" echo state" approach to analysing and training recurrent neural networks-with an erratum note'. *Bonn, Germany: German National Research Center for Information Technology GMD Technical Report*, *148*.

Jaeger, H., & Haas, H. (2004). Harnessing nonlinearity: Predicting chaotic systems and saving energy in communication. *Science*, *304*, 78–80. https://doi.org/https://doi.org/10.1126/science.1091277

Kivshar, Y. S., & Agrawal, G. P. (2003). *Optical solitons: From fibers to photonic crystals*. Academic Press.

Lee, O., Wei, T., Stenning, K. D., Gartside, J. C., Prestwood, D., Seki, S., Aqeel, A., Karube, K., Kanazawa, N., Taguchi, Y., Back, C., Tokura, Y., Branford, W. R., & Kurebayashi, H. (2024). Task-adaptive physical reservoir computing. *Nature Materials*, *23*, 79–87. https://doi.org/https://doi.org/10.1038/s41563-023-01698-8

Lukoševičius, M., & Jaeger, H. (2009). Reservoir computing approaches to recurrent neural network training. *Computer Science Review*, *3*(3), 127–149. https://doi.org/https://doi.org/10.1016/j.cosrev.2009.03.005

Maass, W., Natschläger, T., & Markram, H. (2002). Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural Computation*, *14*, 2531–2560. https://doi.org/https://doi.org/10.1162/089976602760407955

Marcucci, G., Pierangeli, D., & Conti, C. (2020). Theory of neuromorphic computing by waves: Machine learning by rogue waves, dispersive shocks, and solitons. *Physical Review Letters*, *125*(9), 093901. https://doi.org/https://doi.org/10.1103/PhysRevLett.125.093901

Marcucci, G., Rechtsman, M. C., Pichler, H., & Conti, C. (2023). A new paradigm of reservoir computing exploiting hydrodynamics. *Physics of Fluids*, *35*, 071703. https://doi.org/https://doi.org/10.1063/5.0157919

Nakajima, K. (2020). Physical reservoir computing: An introductory perspective. *Japanese Journal of Applied Physics*, *60*, 060501. https://doi.org/https://doi.org/10.35848/1347-4065/ab8d4f

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, *12*, 2825–2830. https://doi.org/https://doi.org/10.5555/1953048.2078195

Pfeifer, R., Lungarella, M., & Iida, F. (2007). Self-organization, embodiment, and biologically inspired robotics. *Science*, *318*(5853), 1088–1093. https://doi.org/https://doi.org/10.1126/science.1145803

Pierro, A., Heiney, K., Shrivastava, S., Marcucci, G., & Nichele, S. (2023). Optimization of a hydrodynamic computational reservoir through evolution. *Proceedings of the Genetic and Evolutionary Computation Conference*, *130*, 202–210. https://doi.org/https://doi.org/10.1145/3583131.3590355

Silva, N. A., Ferreira, T. D., & Guerreiro, A. (2021). Reservoir computing with solitons. *New Journal of Physics*, *23*, 023013. https://doi.org/https://doi.org/10.1088/1367-2630/abda84

Tanaka, G., Yamane, T., Héroux, J. B., Nakane, R., Kanazawa, N., Takeda, S., Numata, H., Nakano, D., & Hirose, A. (2019). Recent advances in physical reservoir computing: A review. *Neural Networks*, *115*, 100–123. https://doi.org/https://doi.org/10.1016/j.neunet.2019.03.005

Vandoorne, K., Mechet, P., Van Vaerenbergh, T., Fiers, M., Morthier, G., Verstraeten, D., Schrauwen, B., Dambre, J., & Bienstman, P. (2014). Experimental demonstration of reservoir computing on a silicon photonics chip. *Nature Communications*, *5*, 3541. https://doi.org/https://doi.org/10.1038/ncomms4541

Zakharov, V. E., & Shabat, A. B. (1972). Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media. *Soviet Physics JETP*, *34*, 62–69.

A Summary of Simulation Parameters

For completeness, we summarize the core numerical parameters used in the experiments:

- **Governing Equation:** Focusing NLSE, $i\psi_t = -\frac{1}{2}\nabla^2\psi - g|\psi|^2\psi$, with nonlinearity strength $g = 1.0$ (NLSE) or $g = 0.0$ (LSE).

- **Domain & Grid:** Square domain of side length $L = 20$ (dimensionless) with periodic boundary conditions. Grid size is $N_x = N_y = 64$ points, with uniform spacing $\Delta x = \Delta y$.

- **Time Stepping:** Split-step Fourier method with step size $\Delta t = 0.001$; $N_t = 1000$ steps ($T = 1.0$) for the main robustness analysis, and $N_t = 2000$ steps ($T = 2.0$) for the nonlinearity g-sweep reported in Appendix C. The longer integration time in the g-sweep is used to ensure that nonlinear self-stabilization dynamics have sufficient time to manifest, while the main robustness analysis uses the minimal horizon required for separability.

- **Initial Conditions:** Two localized sech bumps with class-dependent positions and phases, superposed and normalized.

- **Jitter Procedure:** Initial phase mask $\Phi_c(x, y)$ is perturbed by jitter $\delta\Phi$, where $\delta\Phi$ is drawn i.i.d. from a uniform distribution $[-\eta, \eta]$ at each grid point.

- **Samples:** $N_{\text{class}} = \mathbf{100}$ per class, resulting in **400** total samples per jitter level η.

- **Static Features (Control):** Four global descriptors of the initial wavefield ψ_0: $\langle\cos\varphi\rangle$, $\langle\sin\varphi\rangle$, mass, and gradient energy.

- **Reservoir Features:** 16×16 amplitude field created by **subsampling** (taking every 4th pixel) the 64×64 evolved wavefield ψ_T, plus the same four global descriptors at T. This yields a **260**-dimensional feature vector ($16^2 + 4$).

- **Classifier:** Multinomial logistic regression with ℓ_2 regularization; 70/30 stratified train/test split[8] with fixed random seed.

---

[8]The train/test split fixes the linear measurement apparatus; repeated runs probe whether the physical dynamics consistently generate separable structure rather than fitting noise.

B   Numerical Discretization

The two-dimensional Nonlinear Schrödinger Equation (NLSE) is solved numerically using the standard split-step Fourier method on a periodic domain. This method is widely used for wave propagation problems as this method is robust, explicit, and conserves the wavefield mass $M = \int |\psi|^2 dx dy$ to numerical precision. All simulations used a grid size of $N_x = N_y = 64$, domain length $L = 20$ (leading to $\Delta x \approx 0.3125$) and a time step of $\Delta t = 0.001$. The main robustness analysis uses $T = 1.0$; the supplementary g-sweep (Appendix C) uses $T = 2.0$ to allow maximal time for self-stabilization dynamics.

The equation of motion (Eq. 1) can be split into a linear part (dispersion) and a nonlinear part (Kerr effect):

$$i\frac{\partial \psi}{\partial t} = (\mathcal{L} + \mathcal{N})\psi, \tag{9}$$

where the linear operator is $\mathcal{L} = -\frac{1}{2}\nabla^2$ and the nonlinear operator is $\mathcal{N}(\psi) = -g|\psi|^2$.

For a time step Δt, the solution is approximated by applying the linear and nonlinear steps sequentially:

$$\psi(t + \Delta t) \approx \exp\left(\frac{\Delta t}{2}\mathcal{L}\right) \exp\left(\Delta t \mathcal{N}\right) \exp\left(\frac{\Delta t}{2}\mathcal{L}\right) \psi(t). \tag{10}$$

1. **Nonlinear Step:** The nonlinear phase is applied in real space:

$$\psi^* = \exp\left(-ig|\psi(t)|^2 \Delta t\right) \psi(t). \tag{11}$$

2. **Linear Step (Fourier):** The dispersion is applied in Fourier space, where the linear operator $\mathcal{L}$ becomes a simple multiplication:

$$\hat{\psi}(t + \Delta t) = \exp\left(-i\frac{1}{2}|\mathbf{k}|^2 \Delta t\right) \mathcal{F}(\psi^*), \tag{12}$$

where $\mathcal{F}$ denotes the 2D Fourier transform and $|\mathbf{k}|^2 = k_x^2 + k_y^2$. The field is then transformed back to real space: $\psi(t + \Delta t) = \mathcal{F}^{-1}(\hat{\psi}(t + \Delta t))$.

This method is robust, explicit, and conserves the wavefield mass $M = \int |\psi|^2 dx dy$ to numerical precision.

C Morphological Computation Mechanism: Nonlinearity-weight Sweep Demonstrates Self-Stabilization

The nonlinear interaction term $g|\psi|^2\psi$ scales the reservoir dynamics while also acting as a morphologically-assigned control parameter that tunes the system between distinct computable regimes. To evaluate this morphological mechanism, we performed a sweep over the nonlinearity strength $g \in [0.0, 1.0]$ at the high-noise setting $\eta = 2.0$, while holding the evolution time fixed at $T = 2.0$. The nonlinearity sweep was conducted at $T = 2.0$ to allow maximal time for self-stabilization dynamics to manifest. The resulting accuracy margin over the linear ($g = 0$) baseline is shown in Fig. 4.

The key prediction of morphological computation is supported by the nonlinear shape of the performance curve. If robustness were attributable to a simple gain effect, the dependence on g would be expected to be monotonic and concave-down. Instead, the curve exhibits a characteristic *sigmoidal* structure with three distinct phases:

1. **Threshold Buildup ($g < 0.3$):** Improvements in test accuracy are small and fluctuate due to the dominance of dispersive and noise-induced drift.

2. **Critical Acceleration ($0.3 \leq g \leq 0.7$):** The self-focusing nonlinearity begins to dominate dispersion and noise, and the dynamics transition into a coherent soliton-like morphology. This region exhibits rapidly increasing accuracy.

3. **Saturation ($g > 0.7$):** The system has entered a stable self-organized state; additional nonlinear strength yields diminishing returns.

To isolate the acceleration regime, we computed the discrete second derivative of accuracy with respect to nonlinearity,

$$\frac{d^2(\text{Acc})}{dg^2},$$

as detailed in Appendix E. The presence of a concave-up region provides a quantitative signature of the threshold behavior:

- The strongest positive curvature, $d^2(\text{Acc})/dg^2 = +1.2500$ at $g \approx 0.30$, confirms the sharp acceleration in accuracy as the system crosses the self-stabilization threshold.

- Beyond the critical region, the curvature becomes strongly negative, reaching -2.0556 at $g \approx 0.72$, indicating that the reservoir has entered a stable, saturated regime.
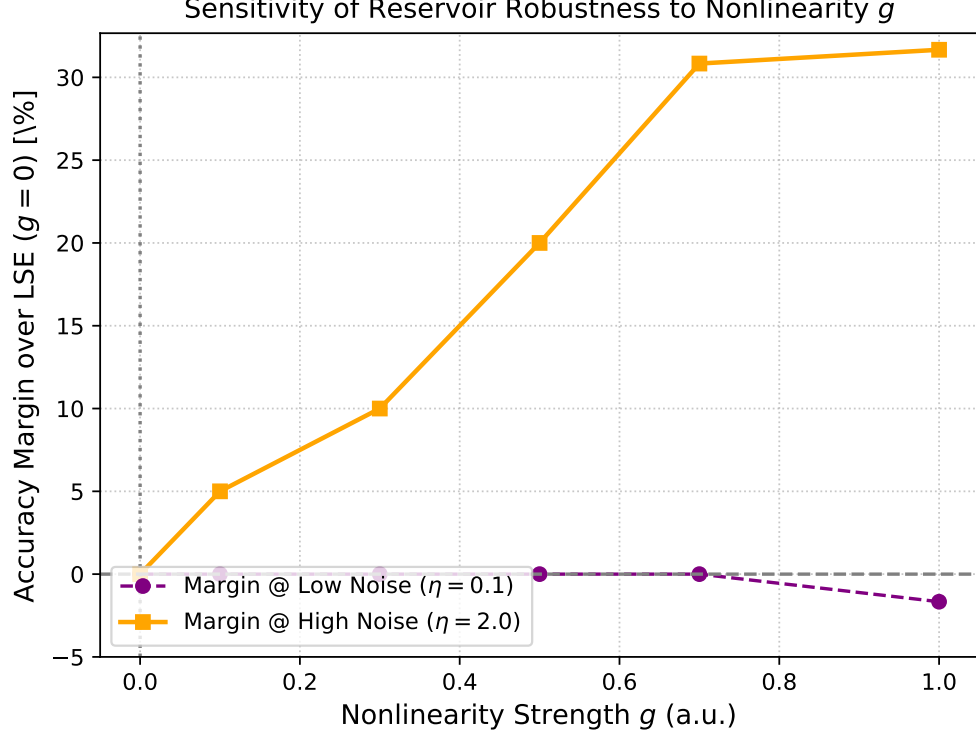
Figure 4: **Accuracy Margin over the Linear Baseline ($g = 0$) as a Function of Nonlinearity Strength g.** The plot displays the accuracy improvement (Margin) relative to the LSE ($g = 0$) for both low-noise ($\eta = 0.1$) and high-noise ($\eta = 2.0$) regimes. The data demonstrates that the computational benefit of the nonlinearity is threshold-driven, exhibiting negligible margin at low g (≈ 0.2) but a substantial, sustained increase in the critical high-noise regime ($\eta = 2.0$) as $g \rightarrow 1.0$. This non-linear growth is mathematically characterized by a region of critical acceleration where the second derivative reaches a maximum of $+\mathbf{125.00}\%/$unit $\mathbf{g^2}$, confirming the threshold profile detailed in Appendix **E** (Figure 5).

Below the threshold, the wavefield remains susceptible to entropic diffusion caused by high phase jitter ($\eta = 2.0$). Once the threshold is crossed, the nonlinear dynamics spontaneously organize into a resilient, class-preserving morphology that resists noise and maintains separability.

D   Classification Data and Full Statistical Sweep

This appendix provides a detailed description of the data used for training and testing, including feature definitions and the complete statistical sweep table.

D.1   Feature Definitions

The four global features used for the static control (ϕ_0) and appended to the reservoir feature set (ϕ_T) are defined as:

- $\langle \cos\varphi \rangle$ and $\langle \sin\varphi \rangle$: The average value of the real and imaginary components of the phase factor, $\exp(i\varphi)$, which capture global phase alignment.

- $\mathbf{M}[\psi]$ (Mass / Norm): The discrete integral $\sum_{x,y} |\psi|^2 \Delta x \Delta y$, which is conserved to numerical precision in the NLSE evolution.

- $\mathbf{K}[\psi]$ (Gradient Energy): The discrete integral $\sum_{x,y} |\nabla\psi|^2 \Delta x \Delta y$, which is related to the kinetic energy.

The reservoir feature vector ϕ_T consists of 256 downsampled amplitude values plus the 4 global features at time T, yielding a vector in $\mathbb{R}^{260}$.

D.2   Complete Statistical Robustness Sweep Data

Table 3 presents the complete set of accuracy results for all three classifiers, reported as the **mean ± standard deviation** over $N = 10$ independent statistical trials. The full dataset is provided here for reproducibility. Values represent test accuracies, reported as $\mu(\sigma)$ across $N = 10$ independently simulated reservoir realizations evaluated using the same fixed train/test split, computed separately for each jitter level η.

Table 3: **Complete Statistical Test Accuracy Sweep.** Mean test accuracy (μ) and standard deviation (σ) reported as $\mu(\sigma)$ across $N = 10$ random seeds for all 15 jitter levels. The results show a crossover regime: LSE maintains a modest advantage at moderate noise ($\eta \approx 1.2$–1.5), while NLSE achieves significant superiority at higher noise levels where $\eta \geq 1.73$.

Jitter η	Static Control		LSE Reservoir ($g = 0.0$)		NLSE Reservoir ($g = 1.0$)	
0.1000	0.513	± 0.054	1.000		0.996	± 0.006
0.2357	0.496	± 0.041	1.000		0.993	± 0.007
0.3714	0.473	± 0.041	0.993	± 0.008	0.993	± 0.006
0.5071	0.398	± 0.033	0.999	± 0.003	0.988	± 0.010
0.6429	0.360	± 0.043	0.999	± 0.003	0.988	± 0.013
0.7786	0.348	± 0.044	1.000		0.977	± 0.009
0.9143	0.343	± 0.044	0.999	± 0.003	0.972	± 0.013
1.0500	0.358	± 0.047	0.998	± 0.003	0.954	± 0.019
1.1857	0.366	± 0.049	0.994	± 0.005	0.948	± 0.024
1.3214	0.358	± 0.033	0.978	± 0.010	0.921	± 0.026
1.4571	0.358	± 0.036	0.947	± 0.021	0.907	± 0.027
1.5929	0.358	± 0.039	0.891	± 0.031	0.896	± 0.026
1.7286	0.353	± 0.040	0.808	± 0.041	0.878	± 0.021
1.8643	0.340	± 0.053	0.727	± 0.037	0.838	± 0.031
2.0000	0.326	± 0.055	0.635	± 0.036	0.767	± 0.055

E   Derivative Analysis for Critical Regimes

To quantify the dynamic behavior of the NLSE reservoir in the high-noise regime ($\eta = 2.0$), we performed discrete derivative analyses on the performance curves. These measurements allow us to characterize how the reservoir's separability and morphological stability evolve with respect to both evolution time and nonlinearity strength.
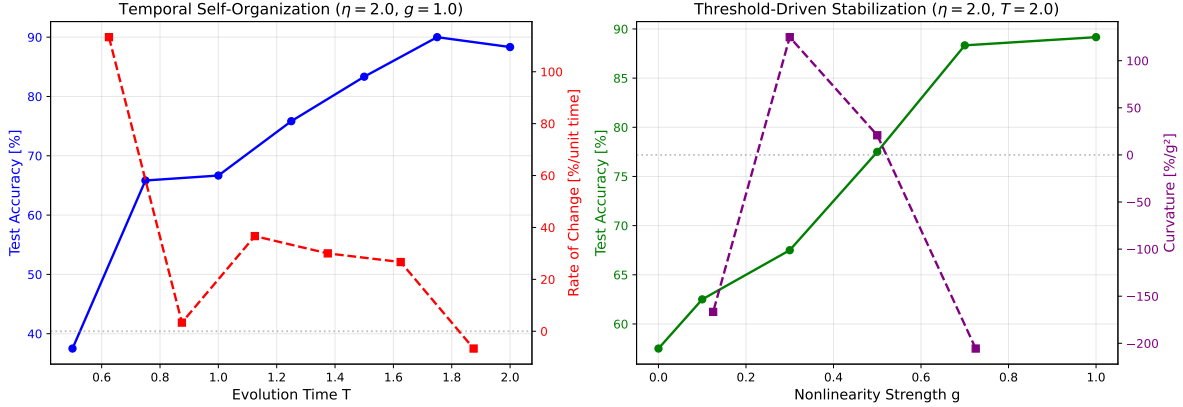

Figure 5: Derivative analysis for the NLSE reservoir in the high-noise regime ($\eta = 2.0$). Left: temporal evolution showing accuracy and its first derivative $d(\text{Acc})/dT$. Right: nonlinearity sweep showing accuracy and curvature $d^2(\text{Acc})/dg^2$, revealing the emergence of acceleration and saturation regimes.

E.1   Evolution Time (T) Derivative Analysis

The first derivative analysis with respect to time is performed on the NLSE accuracy curve at the critical noise level ($\eta = 2.0$). We use a forward finite difference method to estimate the derivative at midpoints between successive data points. The results, summarized in Table 4, show that the mean derivative $d(\text{Acc})/dT$ is positive. This quantitatively rules out that the reservoir is failing due to insufficient or excessive evolution time, confirming that the self-organizing dynamics remain beneficial throughout the evolution interval explored over the course of this work.

The first derivative of accuracy with respect to evolution time,

$$\frac{d(\text{Acc})}{dT},$$

captures the rate at which separability improves (or degrades) as the NLSE state evolves. A positive slope indicates that the evolving morphology continues to enhance class separability over time.

E.2   Nonlinearity Strength (g) Second Derivative Analysis

To characterize the nonlinear contribution to the morphological dynamics we analyze the curvature of the performance with respect to the nonlinearity parameter g at $\eta = 2.0$. The second derivative $d^2(\text{Acc})/dg^2$ serves as an indicator of the transition behavior: positive values specify an acceleration region (concave-up) consistent with threshold-like behavior.

Table 4: **First derivative of accuracy,** $d(\mathbf{Acc})/dT$, **for the NLSE reservoir** $(\eta = 2.0,\ g = 1.0)$. Positive values indicate that the self-organized morphological state maintains or improves separability over evolution time.

Time Step Midpoint T	$d(\mathrm{Acc})/dT$
0.62	1.1333
0.88	0.0333
1.12	0.3667
1.38	0.3000
1.62	0.2667
1.88	-0.0667
Mean:	+0.3389

Computational Details The g-sweep was conducted at six values: $g \in \{1.0, 0.7, 0.5, 0.3, 0.1, 0.0\}$ (note: these are evaluated in descending order), with spacing varying between $\Delta g = 0.1$ and $\Delta g = 0.3$ to capture the transition region. The second derivative was computed via sequential finite differences on the non-uniform grid: first, the first derivative $d(\mathrm{Acc})/dg$ was estimated at midpoints using $[\mathrm{Acc}(g_{i+1}) - \mathrm{Acc}(g_i)]/[g_{i+1} - g_i]$; second, this procedure was repeated on the resulting values to obtain $d^2(\mathrm{Acc})/dg^2$ at the second-level midpoints. No smoothing or filter pre-processing was applied. The qualitative sign pattern—negative curvature at the boundaries ($g \approx 0.1$ and $g \approx 0.7$) with positive curvature in the critical acceleration region ($g \approx 0.3$)—is robust to the specific grid choice: the positive peak remains localized to $g \in [0.2, 0.5]$ regardless of whether adjacent grid points are included or excluded from the analysis.

As shown in Table 5, the local acceleration peaking at $g \approx 0.3$ provides quantitative evidence for threshold-like behavior in the nonlinear stabilization regime.

The second derivative of accuracy with respect to nonlinearity,

$$\frac{d^2(\mathrm{Acc})}{dg^2},$$

describes the concavity of the performance curve and is used to detect acceleration regions (concave-up) and saturation regions (concave-down). Concave-up behavior in the transition range supports the interpretation of a threshold-like nonlinear stabilization effect.

Table 5: **Second derivative of accuracy,** $d^2(\mathbf{Acc})/dg^2$, **for the** g**-sweep** $(\eta = 2.0,\ T = 2.0)$. Positive values around $g \approx 0.3$–0.5 indicate a concave-up acceleration regime consistent with a nonlinear transition in performance.

Midpoint g	$d^2(\mathrm{Acc})/dg^2$
0.72	-2.0556
0.50	+0.2083
0.30	+1.2500
0.12	-1.6667
Mean:	-0.5660 (indicates an overall saturation trend)