Proceedings Track

Residual-Guided Active Learning of Multi-Component Soliton Dynamics with Physics-Informed Neural Networks

Editors: List of editors' names

Abstract

Physics-Informed Neural Networks (PINNs) provide a flexible framework for solving nonlinear partial differential equations (PDEs), but their efficiency is often limited by static collocation strategies that distribute points uniformly across the domain. This work develops and evaluates a residual-driven active learning approach for the three-component Manakov system, a coupled nonlinear Schrödinger-type model characterized by strongly inhomogeneous solution geometry. The method incrementally augments the training set by identifying regions of high PDE residual and concentrating new collocation points in these areas, thereby adapting the sampling distribution to the evolving dynamics of the solution. Numerical experiments show that this targeted placement reduces global loss and component-wise error relative to uniform baselines, achieving comparable accuracy with substantially fewer collocation points. The approach requires no modification to the network architecture or optimization procedure, ensuring that the observed improvements arise solely from the sampling policy. These results demonstrate that residual-driven collocation can improve sample efficiency and solution fidelity in multi-component dispersive systems, suggesting a scalable path for applying PINNs to more complex and stiff nonlinear wave equations. Limitations regarding computational overhead, reproducibility, and the lack of a formal error analysis are also discussed, highlighting directions for future investigation.

Keywords: Physics-Informed Neural Networks (PINNs), Residual-driven collocation, Active learning, Manakov system, Nonlinear Schrödinger equation, Sample efficiency, Dispersive wave dynamics

1. Introduction

A central insight across geometric deep learning and scientific machine learning is that the effectiveness of neural solvers is tightly coupled to the alignment between their training signal and the underlying *geometry* and *dynamics* of the target system. Physics-Informed Neural Networks (PINNs) (Raissi et al., 2019a) instantiate this principle by embedding differential constraints directly into the loss function, enabling networks to approximate continuous solution maps that remain faithful to governing partial differential equations (PDEs) even in the absence of dense supervision. This physics-driven supervision has established PINNs as a versatile framework for solving nonlinear PDEs across diverse scientific and engineering domains, with demonstrated applications in bio-engineering (Kissas et al., 2020; Sahli Costabal et al., 2020; Herrero Martin et al., 2022), fluid dynamics (Raissi et al., 2020, 2019b), stochastic differential equations (Zhang et al., 2020), nano-optics and metamaterials (Chen et al., 2020; Kollmann et al., 2020), photonic device design (Jiang et al., 2021), and epidemic modeling (Malinzi et al., 2022; Shaier et al., 2022; Kharazmi et al., 2021).

Multi-component nonlinear Schrödinger (NLS) systems represent a particularly stringent test case for PINNs(Jaganathan et al., 2023; Wang and Yan, 2021; Pu et al., 2021;

Zhou and Yan, 2021a,a,b; Li and Li, 2021; Meiyazhagan et al., 2022). In the defocusing Manakov model, the three-component variant produces rich soliton dynamics with sharp, spatially inhomogeneous curvature, intricate phase structures, and strong stiffness induced by inter-component coupling. Such regimes present challenges not only for classical numerical solvers but also for PINNs, whose success depends critically on the placement of collocation points(Lau et al., 2024; Celaya et al., 2025). Uniform random sampling of the input domain tends to oversample flat, low-curvature "background" regions while undersampling sharp soliton interfaces or "valley" structures where the solution exhibits strong gradients and nontrivial phase behavior. This mismatch inflates training costs, hampers residual minimization, and constrains the achievable fidelity(Arthurs and King, 2021).

This work critically examines the role of sample placement in PINN training for the three-component defocusing Manakov system, a model that has been previously studied in the literature (Jaganathan et al., 2023). Our analysis proceeds using a fixed network architecture and a purely physics-informed loss. We compared two regimes: a static baseline that deploys all 20,000 collocation points sample using LHS(Stein, 1987) from the outset, and a residual-driven *active learning* strategy that starts with mere 1000 collocation points and appends 200 high-residual points every 3,000 epochs. By construction, the latter concentrates training effort in regions where the model exhibits the largest PDE violations, thereby improving sample efficiency without modifying either the network structure or the optimization framework.

All experiments employ an identical feedforward architecture and are optimized using Adam with fixed hyperparameters. In addition to imbalanced initializations, we also evaluate balanced configurations and their active counterparts, enabling a controlled comparison across regimes. This design ensures that any observed performance differences arise solely from sampling policy rather than architectural or optimization strategies(Münzer and Bard, 2022; Magar and Barati Farimani, 2023). The central question we address is how the *location* of collocation points mediates the *quantity* required to achieve high-fidelity reconstructions of three-component soliton solutions.

By isolating the effect of adaptive sampling, this study advances the understanding of how training signal geometry interacts with PDE stiffness in multi-component nonlinear wave systems. Beyond benchmarking static versus residual-adaptive collocation, our results demonstrate that targeted sampling along high-curvature soliton interfaces not only reduces sample complexity but also improves robustness across distinct initializations. These findings highlight the importance of aligning data placement with the underlying PDE geometry in scientific machine learning, offering practical guidance for deploying PINNs on high-dimensional, strongly coupled dynamical systems(Yan et al., 2023; Gao et al., 2023; Gao and Wang, 2025; Bakthavatchalam et al., 2024). The mathematical formulation of the problem is presented in Section 2, and the sampling policies are described in Section 3.

2. Problem Formulation and PINN Objective

We consider the three-component defocusing Manakov system, a vector-valued generalization of the nonlinear Schrödinger (NLS) equation with cubic self- and cross-phase modulation(Qin et al., 2021; Jaganathan et al., 2023). The governing PDE reads

$$i\, q_{j,t}(x,t) + \tfrac{1}{2}\, q_{j,xx}(x,t) - \left(\sum_{k=1}^{\max(j)} |q_k(x,t)|^2 \right) q_j(x,t) = 0, \qquad j = 1,2,3 \tag{1}$$

where $q_j : \Omega \to \mathbb{C}$ denotes the j-th field component, and the spatiotemporal domain is $\Omega := \Omega_x \times \Omega_t \subset \mathbb{R}^2$, with $x \in \Omega_x \subset \mathbb{R}$ and $t \in \Omega_t \subset \mathbb{R}_+$. Equation (1) describes the evolution of three coupled wave envelopes, each interacting through a shared nonlinear potential determined by the total intensity $\|\mathbf{q}\|^2 = \sum_{k=1}^{3} |q_k|^2$.

PINN approximation. We approximate the solution $q = (q_1, q_2, q_3)$ with a shared real-valued feedforward neural network $f_\theta : \mathbb{R}^2 \to \mathbb{R}^6$, parameterized by weights θ. The output encodes the real and imaginary parts of each component:

$$f_\theta(x,t) \;=\; \big(\Re q_1(x,t),\, \Im q_1(x,t),\, \Re q_2(x,t),\, \Im q_2(x,t),\, \Re q_3(x,t),\, \Im q_3(x,t)\big). \tag{2}$$

The neural surrogate is trained to minimize a physics-informed objective that enforces consistency with the PDE, initial conditions (IC), and boundary conditions (BC).

Physics-informed loss functional. The total objective is given by

$$\mathcal{L}(\theta) \;=\; \lambda_{\text{PDE}} \sum_{(x,t)\in\Omega_{\text{int}}} \big\|\mathcal{N}_\theta(x,t)\big\|_2^2 + \lambda_{\text{IC}} \sum_{x\in\Omega_x} \big\|f_\theta(x,0) - q^\star(x,0)\big\|_2^2 + \lambda_{\text{BC}} \sum_{t\in\Omega_t} \big\|\mathcal{B}[f_\theta](t)\big\|_2^2, \tag{3}$$

where:

- $\Omega_{\text{int}} \subset \Omega$ is the interior collocation set.
- $q^\star(x,0)$ specifies the prescribed initial condition at $t = 0$.
- $\mathcal{B}[\cdot]$ encodes boundary constraints in space.
- $\mathcal{N}_\theta(x,t)$ is the PDE residual of (1) with q replaced by f_θ:

$$\mathcal{N}_\theta^{(j)}(x,t) := i\, q_{j,t}(x,t) \;+\; \tfrac{1}{2}\, q_{j,xx}(x,t) \;-\; \left(\sum_{k=1}^{3} |q_k(x,t)|^2 \right) q_j(x,t), \qquad j = 1,2,3, \tag{4}$$

The loss (3) encodes no explicit conservation laws (e.g. mass, momentum, Hamiltonian) or symmetry penalties. Thus, the only inductive bias stems from minimization of the PDE and auxiliary data constraints, ensuring a clean setting for isolating the impact of sampling strategies.

3. Sampling Policies and Active Learning

Let $\mathcal{S}_0 \subset \Omega$ denote the initial collocation set and $\mathcal{S}_t$ the state of the set after t active updates. We compare two sampling regimes under a fixed overall budget.

Baseline policy. The static baseline consumes the full allocation of collocation points at initialization:

$$\text{Baseline:} \qquad \mathcal{S}_0 = \mathcal{S}_{\text{full}}, \quad \big|\mathcal{S}_{\text{full}}\big| = 20{,}000, \quad \mathcal{S}_t = \mathcal{S}_0 \quad \forall t \geq 0. \tag{5}$$

Residual-driven active policy. In contrast, the adaptive policy augments the collocation set incrementally based on the PDE residual defined in Eq. (4). At fixed intervals of $T = 3000$ epochs, residuals $\|\mathcal{N}_\theta(x,t)\|_2$ are evaluated on a dense candidate grid. The $K = 200$ points with the largest violations are selected:

$$\text{Active update (every } T \text{ epochs):} \quad \mathcal{U}_t = \mathrm{TopK}\Big(\{(x,t') \in \Omega : \ \|\mathcal{N}_\theta(x,t')\|_2\}\Big), \quad |\mathcal{U}_t| = 200. \tag{6}$$

$$\mathcal{S}_{t+1} = \mathcal{S}_t \cup \mathcal{U}_t. \tag{7}$$

This procedure concentrates collocation mass in regions of high residual—often corresponding to sharp curvature, soliton overlap, phase interference zones, or strong cross-component coupling—while avoiding oversampling of flat zones.

4. Experiments and Results

In this section, we evaluate the effectiveness of incorporating residual-driven active learning and coefficient symmetry constraints into the Physics-Informed Neural Network (PINN) framework for solving three-component Manakov systems. Our analysis focuses on three dimensions of performance: (i) data efficiency, (ii) convergence dynamics, and (iii) accuracy of the recovered spatiotemporal fields.

4.1. Experimental Setup

Model architecture. All experiments employ an identical PINN architecture to avoid confounding effects from model capacity. The network consists of four hidden layers, each of width 50, with tanh nonlinearities. Inputs are $(x,t) \in \mathbb{R}^2$; outputs are the six real-valued fields corresponding to the complex solution vector (q_1, q_2, q_3). Training is performed with the Adam optimizer ($\eta = 10^{-3}$, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-8}$) and no weight decay. These settings are consistent across all experiments.

Coefficient initialization. We distinguish between two initialization strategies for the nonlinear coupling coefficients of the Manakov model:

- **Balanced initialization:** all three components are initialized with equal nonlinear interaction strength, enforcing permutation symmetry in the coupling terms.
- **Imbalanced initialization:** the three components are initialized with unequal coupling coefficients, breaking permutation symmetry at initialization.

This distinction allows us to isolate the effect of coefficient balance on solution accuracy and convergence.

Residual-driven active learning. For each initialization strategy, we compare static collocation sampling with residual-driven refinement. In the active learning setting, training begins with 1,000 collocation points and every $T = 3000$ epochs, $K = 200$ points with the highest PDE residuals are added to the training set. This process continues until a maximum of 300,000 epochs, yielding up to 20,000 collocation points. In contrast, static models are trained on a fixed uniform set of 20,000 collocation points for 100,000 epochs.

Evaluation metrics. At fixed checkpoints (every 100 epochs), we record:

1. the total loss $\mathcal{L}(\theta)$ and its decomposition into PDE residual, initial condition (IC), and boundary condition (BC) contributions,
2. the size of the collocation set $|\mathcal{S}_t|$,
3. the elapsed wall-clock training time,
4. and the relative L^2 error of each component (q_1, q_2, q_3) with respect to a high-resolution spectral reference solution.

Configuration	Elapsed	#points	L^2 error	q_1	q_2	q_3
Imbalanced (static)	3.10×10^3	20000	1.34×10^{-4}	3.81×10^{-2}	5.68×10^{-2}	7.19×10^{-2}
Imbalanced (active)	5.14×10^3	6800	7.72×10^{-6}	3.37×10^{-2}	5.07×10^{-2}	8.37×10^{-2}
	1.61×10^4	17600	1.75×10^{-6}	3.28×10^{-2}	4.83×10^{-2}	8.30×10^{-2}
Balanced (static)	2.15×10^3	20000	3.16×10^{-4}	3.39×10^{-2}	5.19×10^{-2}	2.88×10^{-2}
	2.60×10^3	20000	1.30×10^{-4}	2.07×10^{-2}	3.70×10^{-2}	2.21×10^{-2}
Balanced (active)	4.55×10^3	3000	4.92×10^{-5}	1.27×10^{-2}	2.63×10^{-2}	2.10×10^{-2}
	1.35×10^4	20000	2.11×10^{-6}	2.54×10^{-3}	2.22×10^{-3}	3.28×10^{-3}

Table 1: Quantitative comparison of PINN performance on the three-component Manakov system. Results are shown for balanced vs. imbalanced coefficient initialization, each with static or residual-driven active sampling. Reported values include elapsed training time (in seconds elapsed), number of collocation points, overall relative L^2 error, and component-wise errors for (q_1, q_2, q_3).

4.2. Quantitative Results

Table 1 highlights the effect of both coefficient initialization and active learning. For static sampling, balanced initialization yields lower L^2 error compared to imbalanced initialization, indicating that symmetry in coupling accelerates convergence. Residual-driven active learning provides a further improvement of two orders of magnitude, achieving errors as low as 10^{-5} with only one-third of the data. Notably, the balanced-active configuration attains the best performance, reducing error to 2.11×10^{-6} with 20,000 points, and achieving comparable accuracy even with 3,000 adaptively chosen points.

4.3. Qualitative Results

Figures 2 compare PINN-predicted densities $|q_j(x,t)|^2$ against reference solutions for each component under the imbalanced-active configuration. Refer Appendix A for full list of the figures comparing performances for all the experiments. In Figure 2 panels (a,b) show predicted vs. exact solutions, (c) reports the pointwise squared error, and (d–f) plot temporal slices at $t = 0.5$, $t = 2.01$, and $t = 3.5$. The close alignment between predicted and reference dynamics confirms that the PINN accurately resolves double-valley dark soliton structures.

The maximum pointwise error is on the order of 10^{-2}, consistent with the quantitative benchmarks.

It is worth noting at this juncture that the L^2-norm of the discrepancy for the scalar NLS equation with symmetric potentials is typically of the order 10^{-3}–10^{-2} (Meiyazhagan et al., 2022), while for the two-component coupled NLS equation it is of the order 10^{-2} (Mo et al., 2022; Pu and Chen, 2022). These benchmarks provide context for evaluating our results: the close alignment between the predicted and reference dynamics in the three-component Manakov system confirms that the PINN framework is able to accurately resolve double-valley dark soliton structures. The maximum pointwise error remains on the order of 10^{-2}, consistent with the established quantitative baselines. Taken together, this indicates that PINNs can achieve comparable levels of accuracy even in higher-component, more nonlinear coupled NLS systems.

Figure 1 further illustrates the difference between uniform collocation sampling and residual-driven active refinement. Active learning concentrates samples in regions of steep gradients and near soliton valleys, thereby reducing wasted sampling in smooth domains. This geometric adaptivity explains why active models reach high accuracy with significantly fewer training points.

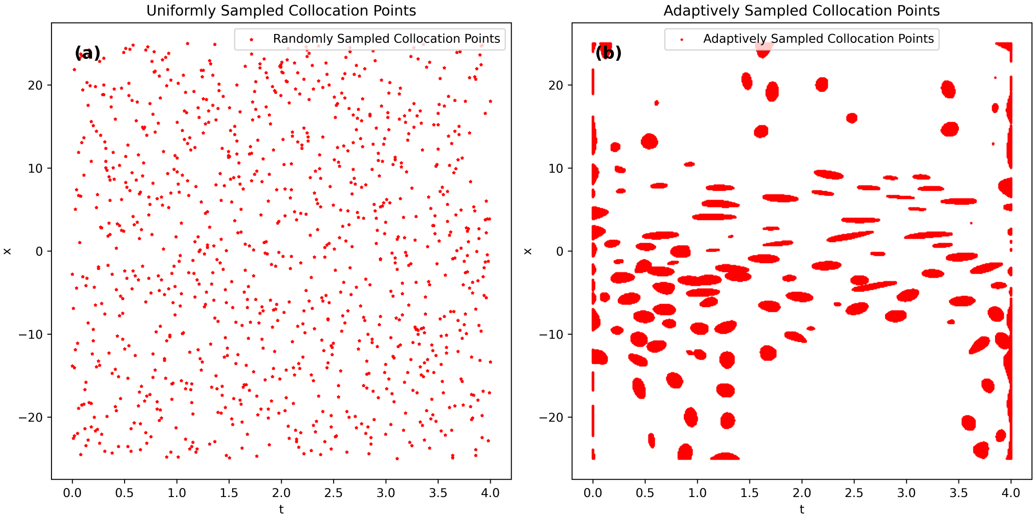

Figure 1: Collocation strategies for the Manakov system. (a) Uniform random sampling spreads points evenly across the domain. (b) Residual-driven active sampling concentrates points in regions of high PDE residual, such as soliton interaction zones and temporal boundaries.

5. Discussion

The experiments establish two central determinants of PINN performance for three-component Manakov systems: (i) the placement of collocation points, and (ii) the symmetry of coefficient initialization. Here we synthesize the broader implications.

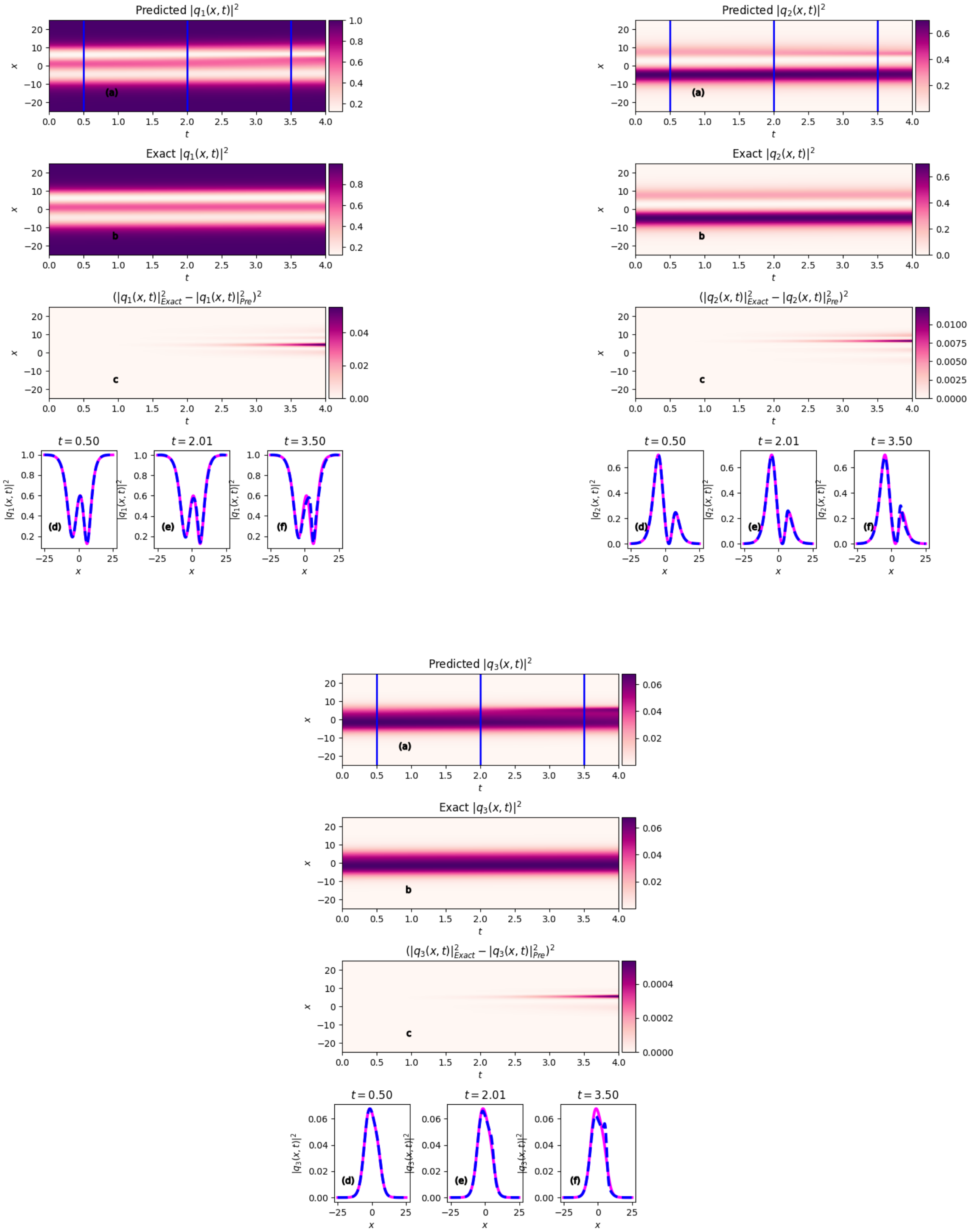

Figure 2: Performance of the PINN under imbalanced initialization with residual-driven active sampling. Shown are predicted spatiotemporal intensities $|q_1(x,t)|^2$, $|q_2(x,t)|^2$, and $|q_3(x,t)|^2$, together with their exact counterparts and error fields.

Adaptive collocation and efficiency. In the three-component defocusing Manakov system studied here, accuracy depends not only on the number of collocation points but also on their alignment with regions of large PDE residual. Residual-driven augmentation concentrates samples near soliton valleys and high-gradient interfaces while reducing redundancy in smooth regions. Empirically, this yields order-of-magnitude improvements in residual loss at comparable or smaller budgets, without modifying the network or objective. These findings highlight adaptive sampling as a practical strategy in this setting, though broader applicability to other PDE classes remains an open question.

Role of initialization balance. Balanced initialization of nonlinear coupling coefficients further reduces error and accelerates convergence. This finding indicates that enforcing structural symmetries at the level of model parametrization can interact constructively with adaptive sampling policies. In multi-component dispersive systems, such balance appears to regularize the optimization landscape, guiding the network toward physically consistent solutions with lower sample budgets.

Broader scientific context. The interplay between initialization symmetry and adaptive sampling highlights a general principle: PINN effectiveness depends not only on the architecture and optimizer, but critically on the geometry of the training distribution. For nonlinear Schrödinger-type systems, where localized nonlinearities dominate the dynamics, this insight suggests that carefully designed sampling policies are as important as network design. More broadly, these results support the view that successful scientific machine learning requires joint consideration of physical invariants, data geometry, and numerical efficiency, rather than relying on raw overparameterization.

Limitations and future work. While the present study focuses on the three-component Manakov model, the methodology generalizes to higher-component or higher-dimensional dispersive systems. Open questions remain regarding the theoretical characterization of sampling–symmetry interactions, as well as their extension to stochastic forcing or non-integrable couplings. Addressing these may further clarify how physical structure and adaptive learning can be co-optimized in PINN frameworks.

6. Conclusion

This study examined the impact of collocation point placement and coefficient initialization symmetry on the performance of Physics-Informed Neural Networks applied to the three-component Manakov system. Two consistent findings emerge. First, residual-driven active learning yields substantial gains in sample efficiency by concentrating collocation points in regions where the PDE residual is largest. Second, balanced initialization of nonlinear coupling coefficients reduces training error and interacts constructively with adaptive sampling, producing the lowest observed losses and component-wise errors.

Together, these results indicate that efficiency in PINNs is governed not solely by network, physical loss constraints or optimizer choice, but by the geometry of the training distribution and its interaction with problem symmetries. The methodology presented here does not alter model structure or loss function, suggesting that the benefits are broadly applicable across dispersive wave systems governed by nonlinear Schrödinger-type equations.

Systematic studies of variance across extension to higher-dimensional or non-integrable regimes are natural directions for future work. Within these bounds, the evidence supports residual-driven collocation as a principled mechanism for improving efficiency and fidelity in physics-informed training.

References

Christopher J. Arthurs and Andrew P. King. Active training of physics-informed neural networks to aggregate and interpolate parametric solutions to the navier-stokes equations. *Journal of Computational Physics*, 438:110364, August 2021. ISSN 0021-9991. doi: 10.1016/j.jcp.2021.110364. URL `http://dx.doi.org/10.1016/j.jcp.2021.110364`.

Tamil Arasan Bakthavatchalam, Selvakumar Murugan, Murugesan Vadivel, Meiyazhagan Jaganathan, Gopinath Balu, and Malaikannan Sankarasubbu. Active learning for probabilistic machine learning-based modeling of dynamical systems. In Walter Lacarbonara, editor, *Advances in Nonlinear Dynamics, Volume I*, pages 399–408, Cham, 2024. Springer Nature Switzerland. ISBN 978-3-031-50631-4.

Adrian Celaya, David Fuentes, and Beatrice Riviere. Adaptive collocation point strategies for physics informed neural networks via the qr discrete empirical interpolation method, 2025. URL `https://arxiv.org/abs/2501.07700`.

Yuyao Chen, Lu Lu, George Em Karniadakis, and Luca Dal Negro. Physics-informed neural networks for inverse problems in nano-optics and metamaterials. *Optics express*, 28(8): 11618–11633, 2020.

Wenhan Gao and Chunmei Wang. Active learning based sampling for high-dimensional nonlinear partial differential equations, 2025. URL `https://arxiv.org/abs/2112.13988`.

Zhiwei Gao, Liang Yan, and Tao Zhou. Failure-informed adaptive sampling for pinns, 2023. URL `https://arxiv.org/abs/2210.00279`.

Clara Herrero Martin, Alon Oved, Rasheda A Chowdhury, Elisabeth Ullmann, Nicholas S Peters, Anil A Bharath, and Marta Varela. Ep-pinns: Cardiac electrophysiology characterisation using physics-informed neural networks. *Front. cardiovasc. med*, page 2179, 2022.

Meiyazhagan Jaganathan, Tamil Arasan Bakthavatchalam, Murugesan Vadivel, Selvakumar Murugan, Gopinath Balu, Malaikannan Sankarasubbu, Radha Ramaswamy, Vijayalakshmi Sethuraman, and Boris A Malomed. Data-driven multi-valley dark solitons of multi-component manakov model using physics-informed neural networks. *Chaos, Solitons & Fractals*, 172:113509, 2023.

Jiaqi Jiang, Mingkun Chen, and Jonathan A Fan. Deep neural networks for the evaluation and design of photonic devices. *Nature Reviews Materials*, 6(8):679–700, 2021.

Ehsan Kharazmi, Min Cai, Xiaoning Zheng, Zhen Zhang, Guang Lin, and George Em Karniadakis. Identifiability and predictability of integer-and fractional-order epidemiological models using physics-informed neural networks. *Nat. Comput. Sci.*, 1(11):744–753, 2021.

Georgios Kissas, Yibo Yang, Eileen Hwuang, Walter R Witschey, John A Detre, and Paris Perdikaris. Machine learning in cardiovascular flows modeling: Predicting arterial blood pressure from non-invasive 4d flow mri data using physics-informed neural networks. *Comput. Methods Appl. Mech. Eng.*, 358:112623, 2020.

Hunter T Kollmann, Diab W Abueidda, Seid Koric, Erman Guleryuz, and Nahil A Sobh. Deep learning for topology optimization of 2d metamaterials. *Materials & Design*, 196: 109098, 2020.

Gregory Kang Ruey Lau, Apivich Hemachandra, See-Kiong Ng, and Bryan Kian Hsiang Low. Pinnacle: Pinn adaptive collocation and experimental points selection, 2024. URL https://arxiv.org/abs/2404.07662.

Jiaheng Li and Biao Li. Solving forward and inverse problems of the nonlinear Schrödinger equation with the generalized $\mathcal{PT}$-symmetric Scarf-II potential via PINN deep learning. *Commun. Theor. Phys.*, 73(12):125001, 2021.

Rishikesh Magar and Amir Barati Farimani. Learning from mistakes: Sampling strategies to efficiently train machine learning models for material property prediction. *Computational Materials Science*, 224:112167, 2023. ISSN 0927-0256. doi: https://doi.org/10.1016/j.commatsci.2023.112167. URL https://www.sciencedirect.com/science/article/pii/S0927025623001611.

Joseph Malinzi, Simanga Gwebu, and Sandile Motsa. Determining covid-19 dynamics using physics informed neural networks. *Axioms*, 11(3):121, 2022.

J Meiyazhagan, K Manikandan, JB Sudharsan, and M Senthilvelan. Data driven soliton solution of the nonlinear schrödinger equation with certain $\mathcal{PT}$-symmetric potentials via deep learning. *Chaos*, 32(5):053115, 2022.

Yifan Mo, Liming Ling, and Delu Zeng. Data-driven vector soliton solutions of coupled nonlinear schrödinger equation using a deep learning algorithm. *Phys. Lett. A*, 421: 127739, 2022.

Marcus Münzer and Chris Bard. A curriculum-training-based strategy for distributing collocation points during physics-informed neural network training, 2022. URL https://arxiv.org/abs/2211.11396.

Jun-Cai Pu and Yong Chen. Data-driven vector localized waves and parameters discovery for manakov system using deep learning approach. *Chaos, Solitons & Fractals*, 160: 112182, 2022.

Juncai Pu, Weiqi Peng, and Yong Chen. The data-driven localized wave solutions of the derivative nonlinear schrödinger equation by using improved pinn approach. *Wave Motion*, 107:102823, 2021.

Yan-Hong Qin, Li-Chen Zhao, Zeng-Qiang Yang, and Liming Ling. Multivalley dark solitons in multicomponent bose-einstein condensates with repulsive interactions. *Physical Review E*, 104(1):014201, 2021.

Maziar Raissi, Paris Perdikaris, and George Em Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019a.

Maziar Raissi, Zhicheng Wang, Michael S Triantafyllou, and George Em Karniadakis. Deep learning of vortex-induced vibrations. *J. Fluid Mech.*, 861:119–137, 2019b.

Maziar Raissi, Alireza Yazdani, and George Em Karniadakis. Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations. *Science*, 367(6481):1026–1030, 2020.

Francisco Sahli Costabal, Yibo Yang, Paris Perdikaris, Daniel E Hurtado, and Ellen Kuhl. Physics-informed neural networks for cardiac activation mapping. *Frontiers in Physics*, 8:42, 2020.

Sagi Shaier, Maziar Raissi, and Padmanabhan Seshaiyer. Data-driven approaches for predicting spread of infectious diseases through dinns: Disease informed neural networks. *Lett. Biomath.*, 9(1):71–105, 2022.

Michael Stein. Large sample properties of simulations using latin hypercube sampling. *Technometrics*, 29(2):143–151, 1987.

Li Wang and Zhenya Yan. Data-driven rogue waves and parameter discovery in the defocusing nonlinear schrödinger equation with a potential using the pinn deep learning. *Phys. Lett. A*, 404:127408, 2021.

Junjun Yan, Xinhai Chen, Zhichao Wang, Enqiang Zhoui, and Jie Liu. St-pinn: A self-training physics-informed neural network for partial differential equations, 2023. URL `https://arxiv.org/abs/2306.09389`.

Dongkun Zhang, Ling Guo, and George Em Karniadakis. Learning in modal space: Solving time-dependent stochastic pdes using physics-informed neural networks. *SIAM Journal on Scientific Computing*, 42(2):A639–A665, 2020.

Zijian Zhou and Zhenya Yan. Deep learning neural networks for the third-order nonlinear schrodinger equation: Solitons, breathers, and rogue waves. *arXiv:2104.14809*, 2021a.

Zijian Zhou and Zhenya Yan. Solving forward and inverse problems of the logarithmic nonlinear Schrödinger equation with $\mathcal{PT}$-symmetric harmonic potential via deep learning. *Phys. Lett. A*, 387:127010, 2021b.

Appendix A. Performance comparison of Q_1, Q_2, Q_3 of three component System

This appendix shows performance comparison of Q_1, Q_2, Q_3 on static and active variants of imbalanced and balanced configurations. In our experiments, we compare imbalanced and balanced initializations under both static and active learning regimes. Figure 3 shows the performance plots for experiment configuration imbalanced initialization and static setup serves as the baseline. It is trained using 20,000 uniformly sampled collocation points. This setup achieves multi-valley soliton resolution but incurs high training cost. Figure 2 shows the performance for its active counterpart and is explained in section 4.2. Switching to a balanced initialization (Figure 4) improves both training efficiency and accuracy: the network achieves relative errors of 3.39×10^{-2}, 5.19×10^{-2}, and 2.88×10^{-2} for q_1, q_2, and q_3, respectively, using only two-thirds of the time required in the imbalanced case. Continued training further reduces these errors to 2.07×10^{-2}, 3.70×10^{-2}, and 2.21×10^{-2}, confirming that symmetry in the system aids convergence. Most significantly, incorporating active learning with balanced initialization (Figure 5) leads to considerable data efficiency: with only 3,000 adaptively selected points (6x lesser compared to baseline), the network achieves errors of 1.27×10^{-2}, 2.63×10^{-2}, and 2.10×10^{-2}. While the training duration doubles due to adaptive sampling overhead, the tradeoff favors accuracy per sample. Allowing the active model to use the full budget of 20,000 points drives errors an order of magnitude lower, reaching 2.54×10^{-3}, 2.22×10^{-3}, and 3.28×10^{-3}. These results collectively affirm that balanced configurations accelerate convergence, and that residual-based active sampling significantly improves learning efficiency while preserving accuracy.

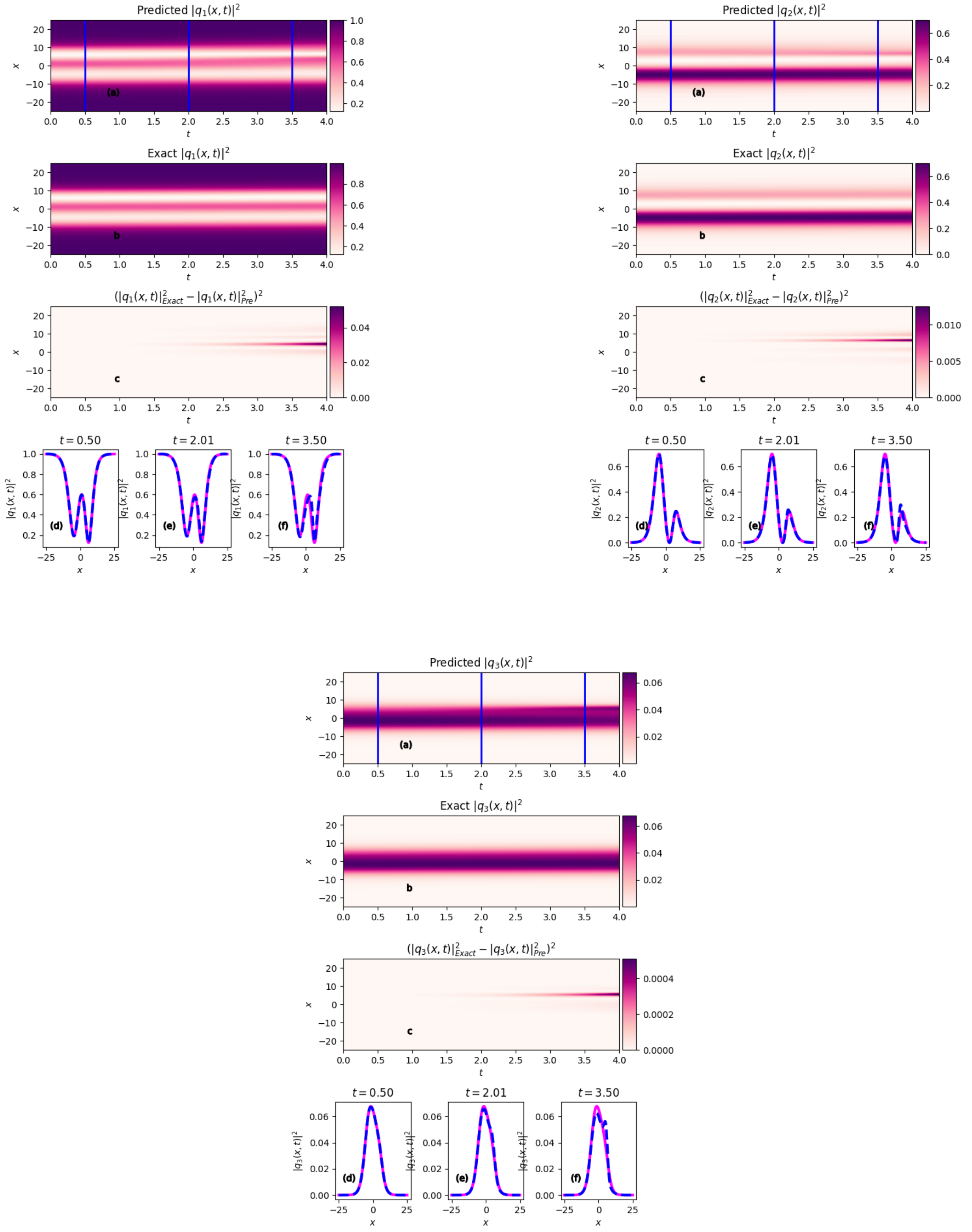

Figure 3: Performance of the PINN under imbalanced initialization with uniform random sampling. Shown are predicted spatiotemporal intensities $|q_1(x,t)|^2$, $|q_2(x,t)|^2$, and $|q_3(x,t)|^2$, together with their exact counterparts and error fields.

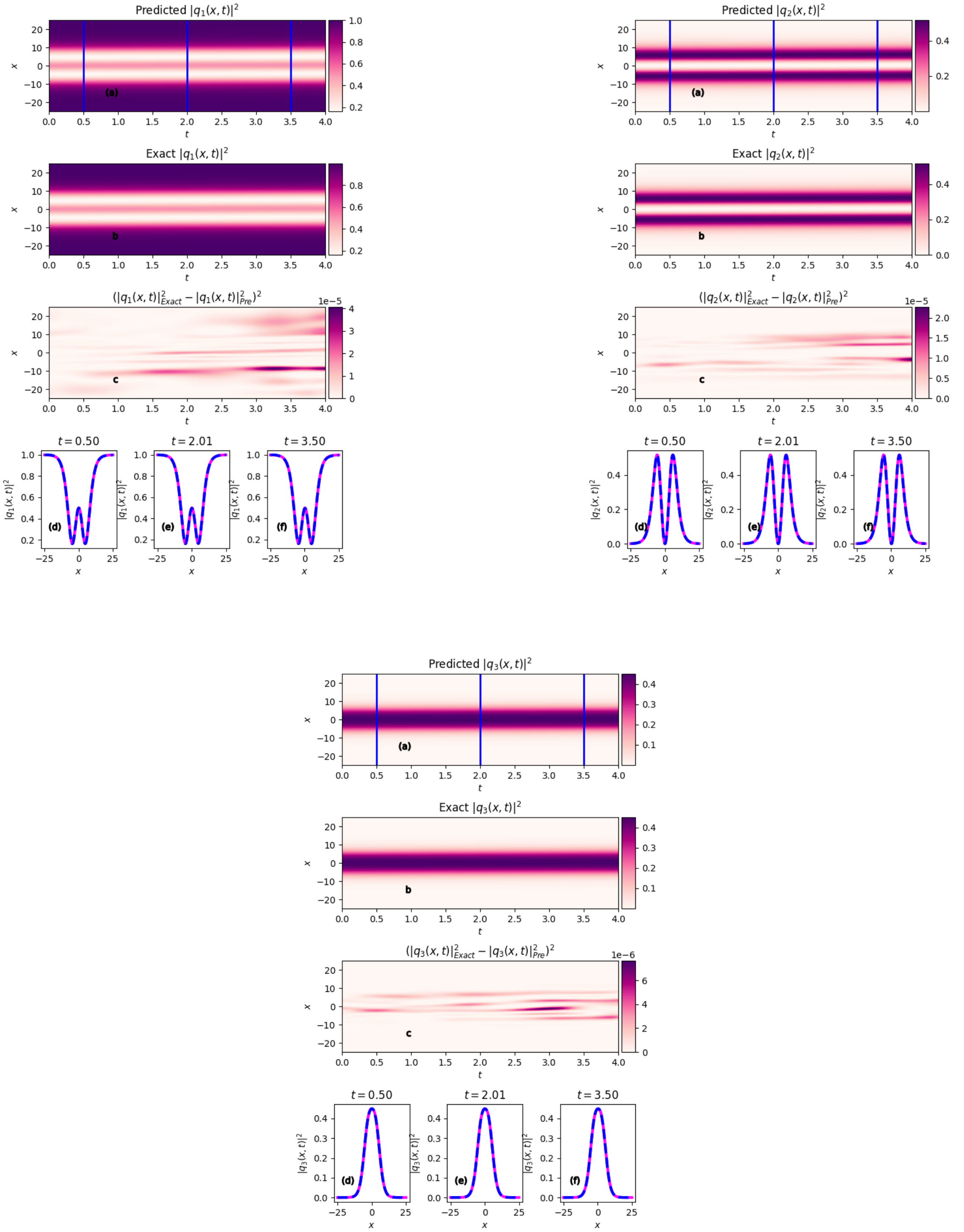

Figure 4: Performance of the PINN under balanced initialization with uniform random sampling. Shown are predicted spatiotemporal intensities $|q_1(x,t)|^2$, $|q_2(x,t)|^2$, and $|q_3(x,t)|^2$, together with their exact counterparts and error fields.

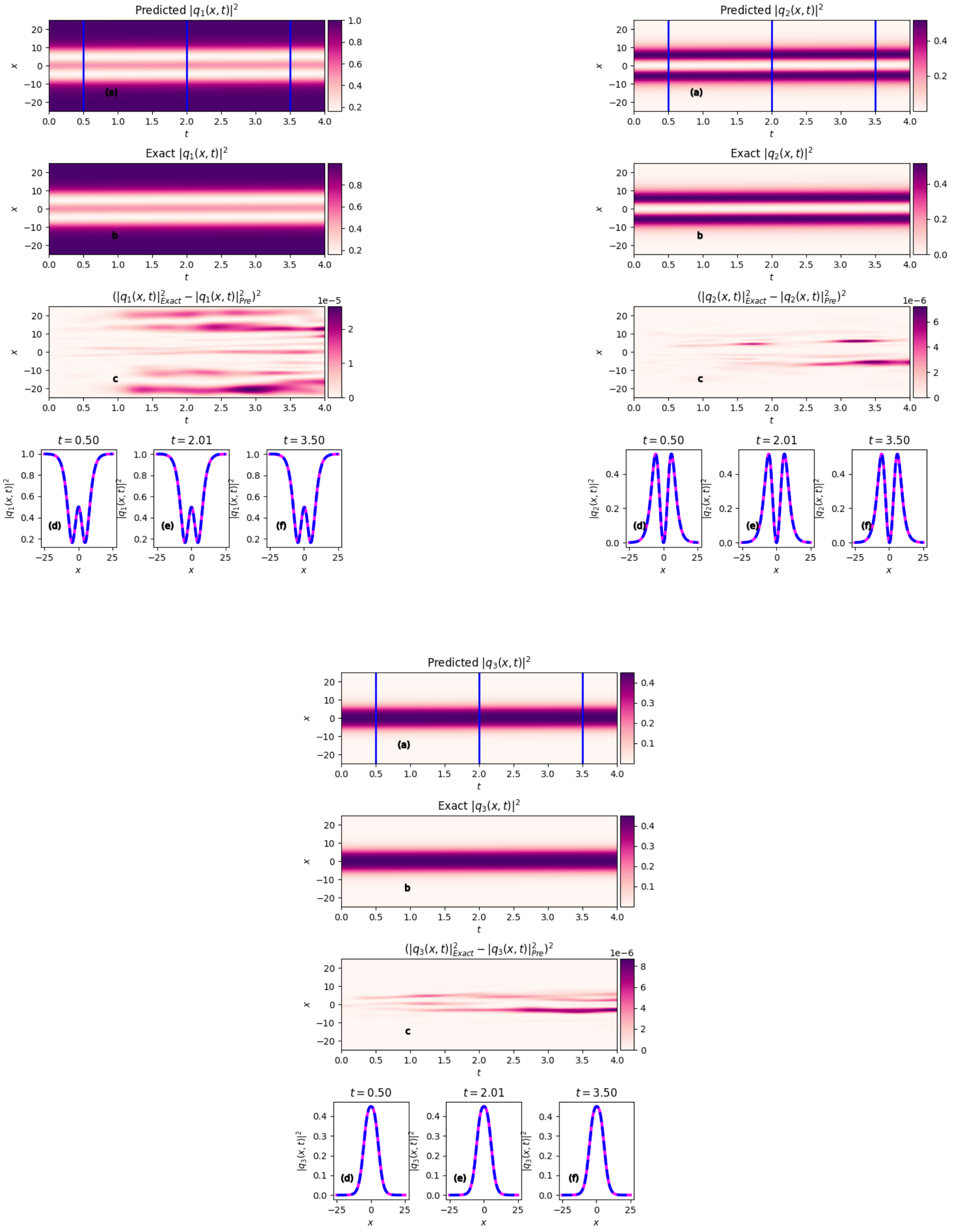

Figure 5: Performance of the PINN under balanced initialization with residual-driven active sampling. Shown are predicted spatiotemporal intensities $|q_1(x,t)|^2$, $|q_2(x,t)|^2$, and $|q_3(x,t)|^2$, together with their exact counterparts and error fields.

Appendix B. Four component System

We extend our investigation to a 4-component coupled system with imbalanced initialization. The observed results for both *static* and *residual-driven active sampling* configurations are summarized in Table 2, with corresponding metrics including total loss, number of training points, and mean absolute errors for the predicted densities $|q_j(x,t)|^2$ across all four components ($j = 1, \dots, 4$). The **static model** trained on 20,000 uniformly sampled collocation points converged in approximately 1.66×10^3 seconds, achieving a loss of 1.12×10^{-4}. In contrast, the **active learning model** demonstrates significantly improved efficiency and accuracy. With just 4,200 points and an elapsed time of 8.81×10^3 seconds, the loss drops to 6.62×10^{-5}, and the error for all components is reduced or maintained, with q_4 improving to 2.78×10^{-2}. Active setup when trained up to 18800 points (after which adding more points did not result in improvement) further reduces the loss by an order of magnitude (4.70×10^{-6})—demonstrating robust generalization across all components. These results support the hypothesis that this methods generalizes to more complex systems.

Experiment	Elapsed	#points	L^2 error	q_1	q_2	q_3	q_4
static	1.66×10^3	20000	1.12×10^{-4}	1.89×10^{-2}	5.69×10^{-2}	7.84×10^{-2}	2.68×10^{-2}
active	8.81×10^3	4200	6.62×10^{-5}	1.86×10^{-2}	5.12×10^{-2}	7.73×10^{-2}	2.77×10^{-2}
	3.05×10^4	18800	4.70×10^{-6}	1.65×10^{-2}	4.61×10^{-2}	6.76×10^{-2}	1.27×10^{-2}

Table 2: Quantitative comparison of PINN performance on the four-component Manakov system. Results are shown for balanced vs. imbalanced coefficient initialization, each with static or residual-driven active sampling. Reported values include elapsed training time (in seconds elapsed), number of collocation points, overall relative L^2 error, and component-wise errors for (q_1, q_2, q_3, q_4).

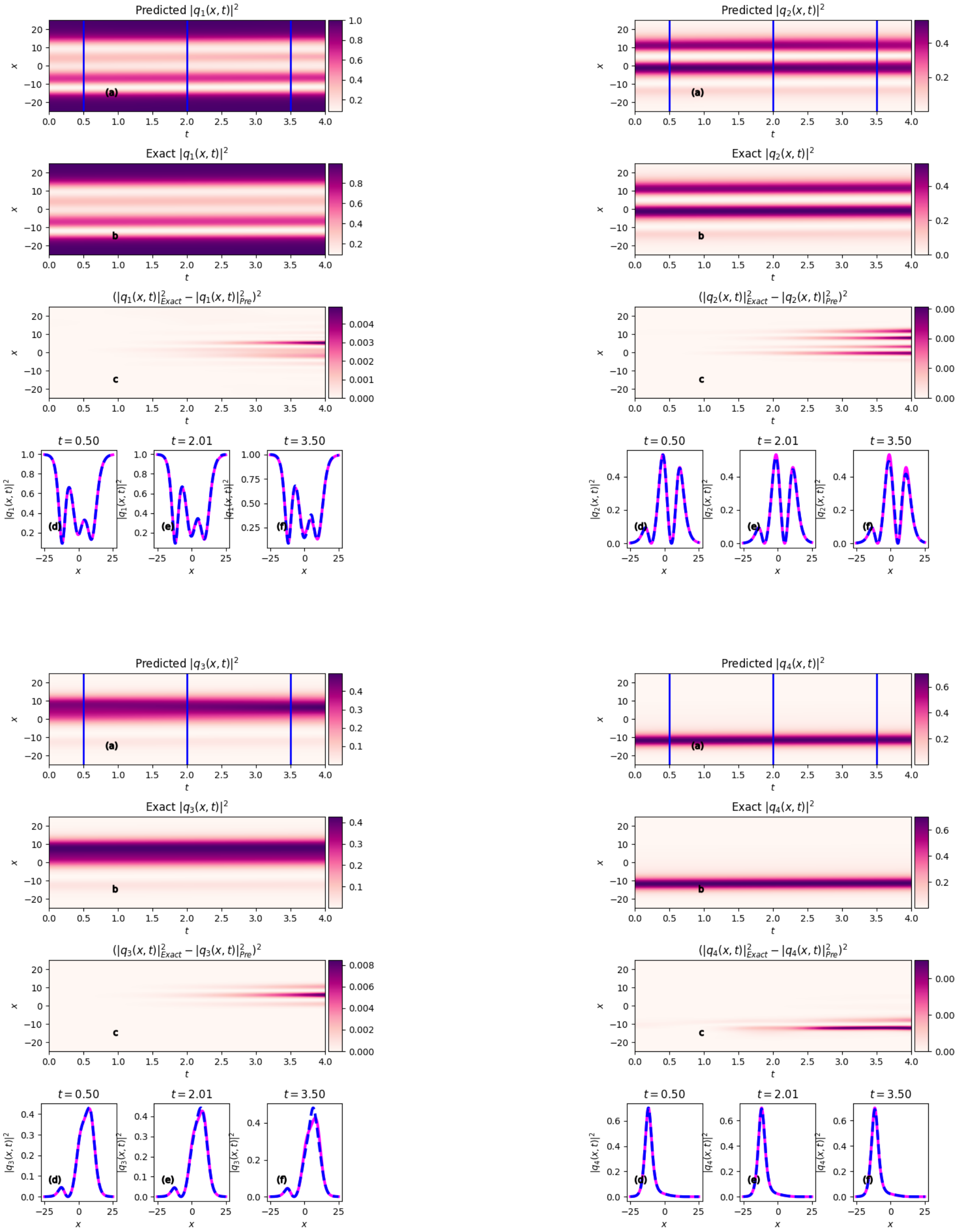

Figure 6: Performance of the PINN under balanced initialization with residual-driven active sampling. Shown are predicted spatiotemporal intensities $|q_1(x,t)|^2$, $|q_2(x,t)|^2$, and $|q_3(x,t)|^2$, together with their exact counterparts and error fields.

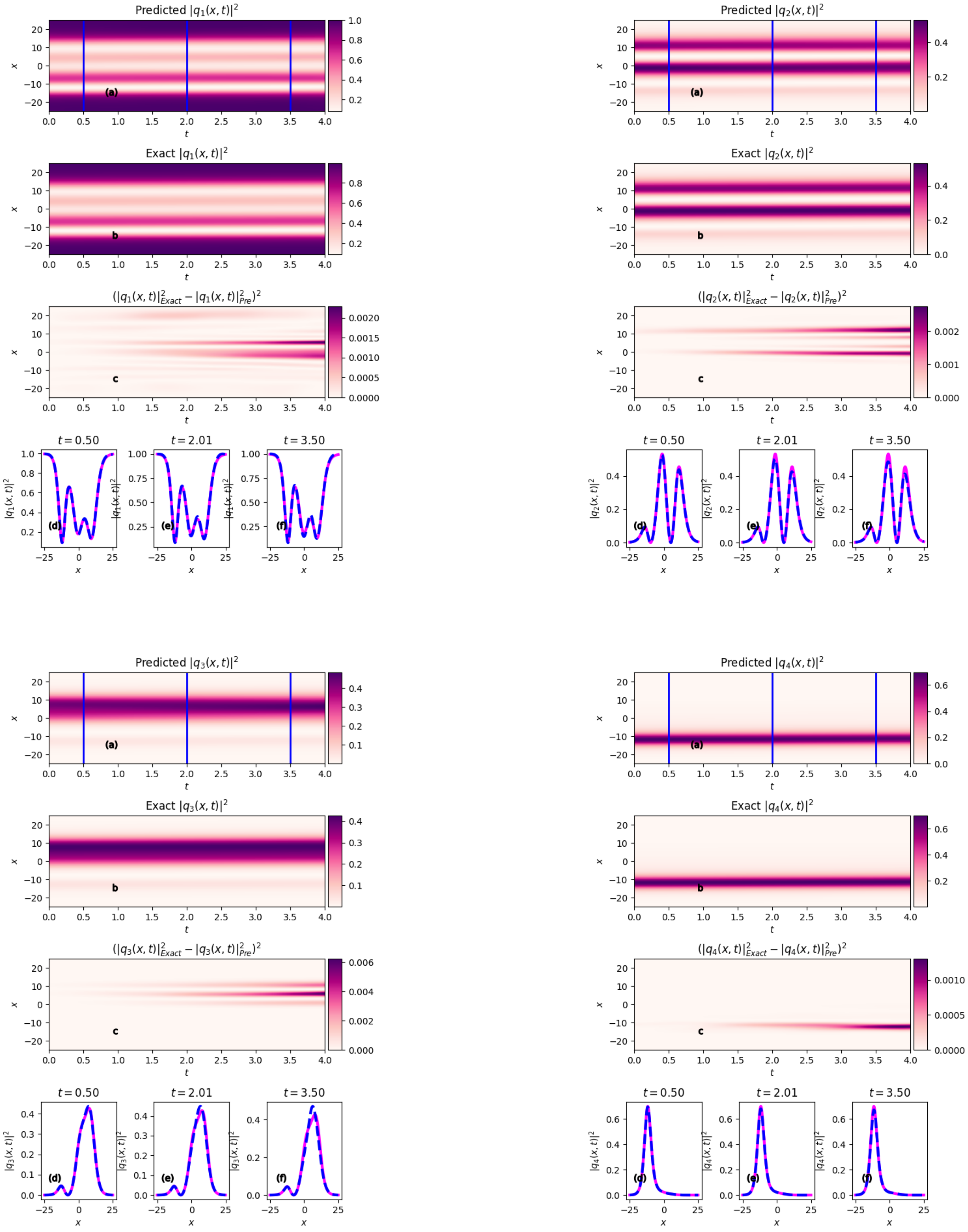

Figure 7: Performance of the PINN under balanced initialization with uniform random sampling. Shown are predicted spatiotemporal intensities $|q_1(x,t)|^2$, $|q_2(x,t)|^2$, and $|q_3(x,t)|^2$, together with their exact counterparts and error fields.