HiLD 2024: 2nd Workshop on High-dimensional Learning Dynamics

On the metastability of learning algorithms in physics-informed neural networks: a case study on Schrödinger operators

Alessandro Maria Selvitella ASELVITE@PFW.EDU
Department of Mathematical Sciecnes & Laboratory of Data Science, Purdue University Fort Wayne, Fort Wayne (IN), USA; eScience Institute, University of Washington, Seattle (WA), USA; NSF-Simons Center for Quantitative Biology, Northwestern University, Evanston (IL), USA.

Abstract

In this manuscript, we discuss an interesting phenomenon that happens in the training of physics-informed neural networks: PINNs seem to go through metastable states during the optimization process. This behaviour is present in several dynamical systems of interest to physics and was first noticed in the Fermi-Pasta-Ulam-Tsingou model, in which the system spends a lot of time in an intermediate state, before, eventually, reaching thermalization. We concentrate on some examples of Schrödinger equations in spatial dimension $n = 1$, including the nonlinear Schrödinger equation with quintic polynomial nonlinearity, the linear Schrödinger equation with trapping potential, and and the linear Schrödinger equation with asymptotically constant potential.
Keywords: Metastability, physics-informed neural networks, Schrödinger operators.

1. Introduction

In the past few years, scientific and physics informed machine learning have emerged as powerful tools in the search for solutions to many scientific problems in areas such as physics, engineering, and biology. Among the most important new methodologies, physics-informed neural networks (PINNs) and derived algorithms have been powerful strategies for the understanding of the dynamics of solutions to partial differential equations (PDEs) and for the discovery of PDEs from data [7, 11]. PINNs encode physical laws as regularizers and are good approximators in the small-data regime [3, 7, 11]. Key questions in PDEs concern the long-time dynamics of solutions and the problem of singularity formation. Global solutions exist for all time in the functional space of interest, while blow-solutions stop to belong to the functional space of the initial datum in finite time [12, 13]. This dicothomy is a fundamental question in PDE research and, actually, at the heart of the Navier Stokes Millenium problem [5]. It is still unknown if smooth solutions exist for all times for the 3D Navier Stokes system. Several PDEs admit equilibrium solutions, namely solutions that are constant in time [2]. Often, these solutions appear as attractors in the long-term dynamics of a physical system or as stable states [13, 14]. It is often harder (and less studied) to understand the properties of dynamical systems in intermediate time ranges, but there are fundamental examples of systems in which some important dynamical features emerge.

In a celebrated computational experiment involving the simulation of a vibrating string with cubic interaction, Fermi, Pasta, Ulam, and Tsingou (FPUT) noticed that a system of nonlinearly interacting particles on a line exhibited a complicated quasi-periodic behavior, instead of the expected ergodic one [6]. Instead of quickly converging towards equipartition, the energy, initially distributed to the lowest frequency modes remains in the lower frequency modes for long times [6]. The energy does eventually equipartition, but after long periods spent around the so-called *metastable states*. This phenomenon fascinated people for decades and it is still of interest [4]. Metastability is intrinsically nonlinear. In fact, in the case of quadratic potentials, the FPUT system can be transformed into a linear one with independent harmonic oscillators. Metastability seems appearing also in optimization trajectories of PINNs and, in this manuscript, we bring some information regarding this possibility through a series of experiments. We concentrate on some examples of Schrödinger equations in spatial dimension $n = 1$, including the nonlinear Schrödinger equation (NLS) with quintic polynomial nonlinearity, the linear Schrödinger equation with trapping potential, and and the linear Schrödinger equation (LS) with asymptotically constant potential. The LS is the fundamental equation in quantum mechanics, while the NLS appears, among the others, in nonlinear optics, small-amplitude gravity waves, plasma physics [9], and Bose–Einstein condensation [10]. Given their importance in physics, it is interesting to understand PINNs' behaviour on LS and NLS.

2. Methods

Physics Informed Neural Networks. We consider PINNs' loss $Loss = Loss_1 + Loss_2$, with

$$Loss_1 = \frac{1}{N_1}\sum_{i=1}^{N_1} |u(t_1^i, x_1^i) - u^i|^2 \quad \text{and} \quad Loss_2 = \frac{1}{N_2}\sum_{i=1}^{N_2} |f(t_2^i, x_2^i)|^2.$$

Here $f := u_t + \mathcal{N}[u]$ is a differential operator, u is a neural network, $\left\{t_1^i, x_1^i, u^i\right\}_{i=1}^{N_1}$ denote the initial/boundary training data on $u(t, x)$, and $\left\{t_2^i, x_2^i\right\}_{i=1}^{N_2}$ specify the collocation points for $f(t, x)$.

Optimization. A 4-th order Runge-Kutta method with 1000 steps has been implemented to compute the exact solutions to the PDEs of interest. The hypothesis class included a feed-forward neural network with 1 hidden layer with 32 nodes and $\tanh$ as activation function. $Loss$ was optimized with the *Adam* algorithm with learning rate $5 * 10^{-4}$. We chose (both uniformly) $N_1 = 2, N_2 = 40$ for the NLS with $p = 5$, $N_1 = 20$ and $N_2 = 100$ for the LS with trapping potential, and $N_1 = 20$ and $N_2 = 40$ for the LS with potential constant at infinity.

The Linear and Nonlinear Schrödinger Equations. The LS in 1d ($n = 1$) is given by

$$i\psi_t = \psi_{xx} + V(x)\psi,$$

with $V(x)$, a real-valued potential. The NLS governs the dynamics of mutual interacting particles:

$$i\hbar\frac{\partial \psi}{\partial t} = -\hbar^2\Delta\psi + \lambda|\psi|^{p-1}\psi.$$

Here: i, the imaginary unit, $\hbar$ the Planck constant, $p \geq 3$, and $\psi : \mathbb{R}^n \times \mathbb{R} \to \mathbb{C}$ representing the wave function. We will consider only the focusing case ($\lambda = -1$). Particles at rest are represented by *standing waves*: $\psi(x,t) := \exp(i\omega/\hbar^{-1}t)u(x)$ with u solving the elliptic PDE:

$$-\hbar^2\Delta u + \omega u = u^p, u \in H^1(\mathbb{R}^n),$$

which, in scaled coordinates, becomes:

$$-\Delta u + u = u^p, u \in H^1(\mathbb{R}^n).$$

$H^1(\mathbb{R}^n)$ represents the space of quadratic integrable functions in $\mathbb{R}^n$ with gradient quadratically integrable in $\mathbb{R}^n$, as well. In polar coordinates, this problem becomes:

$$u'' + \frac{n-1}{r}u' - u + u^p = 0, \quad u'(0) = 0$$

For ground states, we have $\lim_{r\to+\infty} u(r) = 0$. For the purpose of orbital stability, it is important to study the spectrum of the linearized operator around the ground state Q [12]:

$$-\Delta v + v - pQ^{p-1}v = \nu v,$$

with the eigenfunction $v \in H^1(\mathbb{R}^n)$ and ν representing the corresponding eigenvalue. By decomposing the problem into spherical harmonics $v(x) = \sum_{k=0}^{+\infty} \psi_k(r)Y_k(\theta)$ with $r > 0$ representing the radial component and $Y_k(\theta)$ for $\theta \in \mathbb{S}^{N-1}$ being the spherical harmonics, we obtain a sequence of decoupled ODEs that take the following form (see [1]):

$$A_k(\psi_k) := -\psi_k'' - \frac{N-1}{r}\psi_k' + \psi_k + \frac{\lambda_k}{r^2}\psi_k - pQ^{p-1}\psi_k = \nu\psi_k,$$

with $k = 0, 1, 2, \ldots$. Critical is the study of the kernel $\nu = 0$, which is known to be spanned by the following functions $\{Q_{x_i} : 1 \leq i \leq n\}$ [1]. This is proven by showing that the only solutions of the eigenvalue problem come from the mode $k = 1$. Solutions of the modes $k > 1$ are excluded using Perron-Frobenius Theorem, while solutions of the case $k = 0$ are proven to be unbounded using oscillation theory (See [1]). It is well known that the ground state of the NLS is smooth, positive, exponentially decaying, radial, and unique up to translations [1, 2, 8]. In dimension $n = 1$, such a solution has an explicit form: $Q(r) = \sqrt{2}\,\mathrm{sech}(r)$. Solutions that at $r = 0$ start below $Q(0)$, oscillate around the stationary point $u = 1$, those starting above $Q(0)$, become negative in finite time [13].

3. Experiments

In all figures/subfigures involving loss functions in this section, the number of epochs (x-axis) is scaled down by a factor of 1500 (e.g. $x = 30$ refers to the epoch 45000).

NLS with $p = 5$. In Figure 1, we consider the NLS with $p = 5$ and only mode $k = 0$ of the linearized problem around the grounds state Q. As expected, on unbounded domains the convergence is lost, because of representation limitations of the chosen hypothesis class on unbounded domains. However, another phenomenon appears. Although initially [Step=15000] the algorithm seems converging towards the correct solutions on bounded subsets, already at intermediate steps [Step=45000 and Step=48000], the algorithm starts to wiggle and seems deviating from the correct solution (metastability).

LS with Trapping Potential. We consider the LS with trapping potential $V(x) = (1 + x^2)$, with initial conditions $(u(0), u'(0)) = (1, 0)$. Also in this case, we observe metastability. In Figure 2, we can see that the fit of the solution seems going in the right direction until Step=36000, but already at Step=42000, there is evidence of metastability between the first two collocation points and the optimization trajectory continues to degenerate (see, for example, the fit at Step=198000).

LS with Potential Constant at Infinity. We consider the LS with a potential constant at infinity $V(x) = 1/(1 + |x|^{-0.5})$, with initial conditions $(u(0), u'(0)) = (1, 0)$. In Figure 3, we see that the presence of a Loss 2 (b) does not seem to substantially improve the fit with respect to the case Loss = Loss 1 (a). Furthermore, the algorithms seems converging to the correct solution at Step=6000, but it diverges from the metastable state to something highly oscillatory close to the first collocation points.

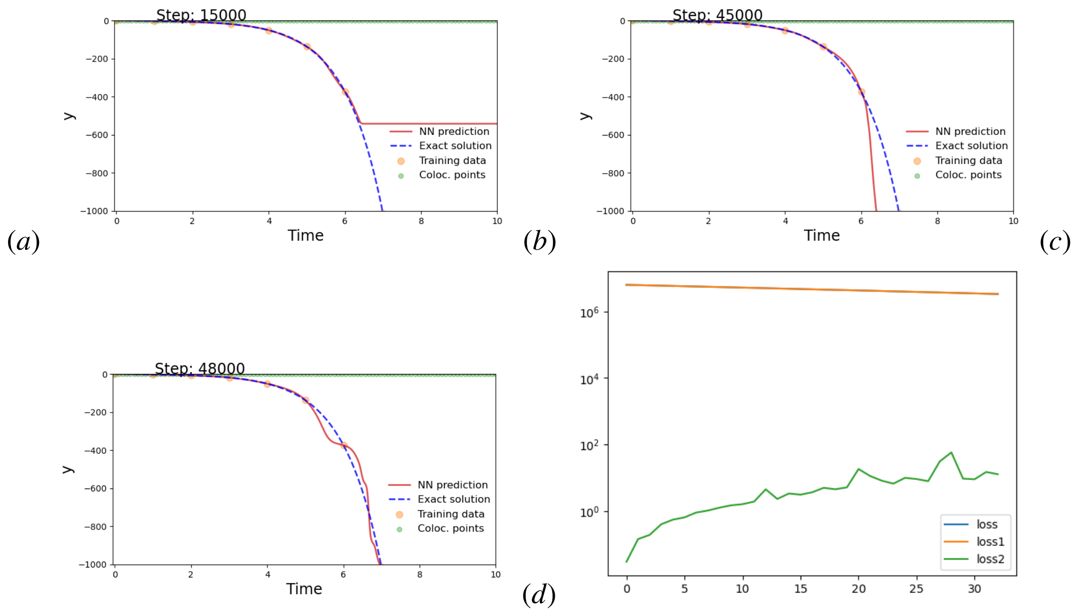

Figure 1: **Nonlinear Schrödinger Equation with $p = 5$: Linearization Mode 1.**

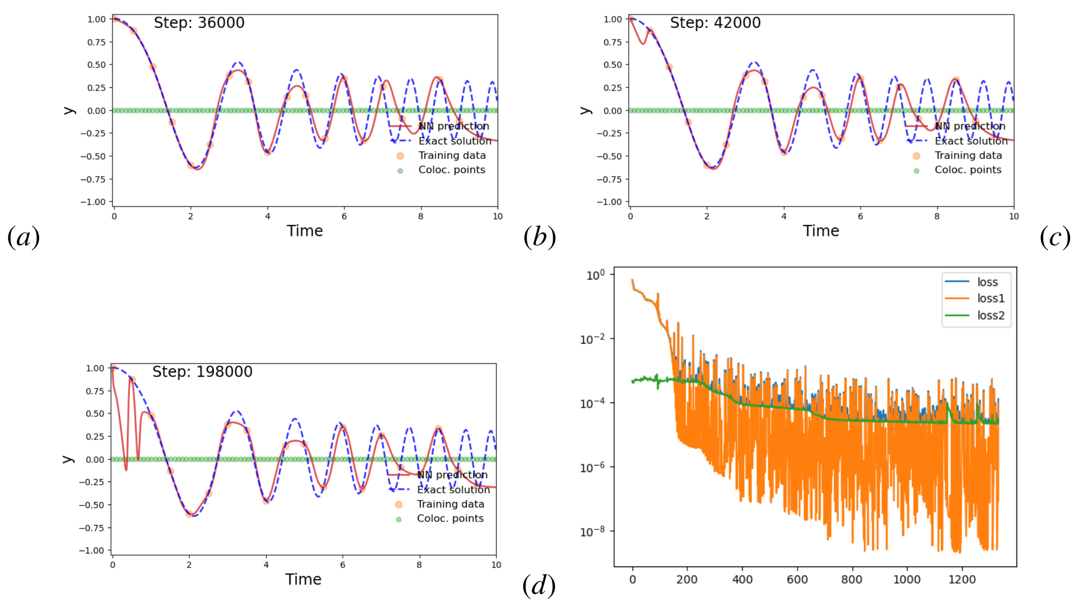

Figure 2: **Linear Schrodinger Equation: Trapping Potential**

4. Discussion and Conclusions

Our experiments on Schrödinger equations show that PINNs seem to go through metastable states during the optimization process, a behaviour that is present in several dynamical systems of interest to physics and that was first noticed in the celebrated FPUT model, in which the system spends a lot of time in an intermediate state, before, eventually, reaching equipartition of energy. We noticed that,

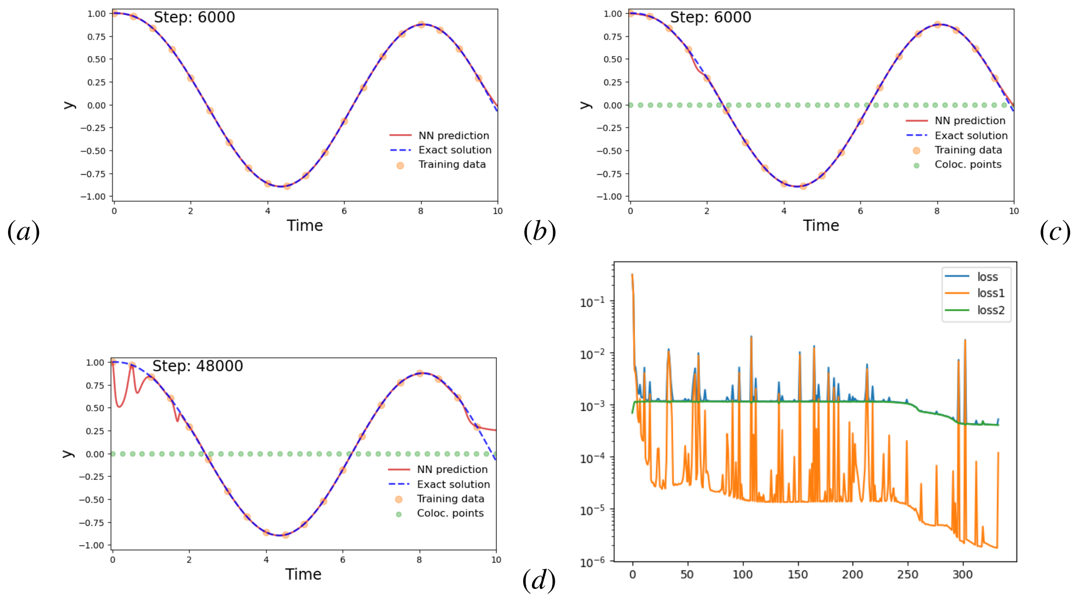

Figure 3: **Linear Schrodinger Equation: Potential Constant at Infinity.**

after a long period in the metastable state, PINNs seem collapsing to the zero solution of the PDE almost everywhere (everywhere, other than around collocation points). This is reminiscent of the known problems with PINNs tending to prefer regular solutions concentrated on lower modes. Note that, in our loss functions, nothing is present to force the solution to be positive. Recall that, in the continuous limit, the FPUT system converges to the KdV equation, a completely integrable and highly symmetric system, which possesses soliton solutions, namely solutions that travel at constant speed without loosing energy [15]. We do not know what equation is the equivalent of the KdV equation for PINNs. It would be interesting to see if such a limit-equation has any relationship with Neural ODEs. It would be also interesting to understand if metastability is related to over-parametrization (the network we used in the experiments is relatively small): Taking an over-parametrized network might alleviate metastability. It is well-known that it is challenging to train PINNs [11] and researchers have worked extensively to improve optimization methods. We took advantage of well-known results on the qualitative behaviour of solutions to LS and NLS, including radiality of the ground state, to reduce the problem to low-dimensional. The goal of this analysis was not to find optimal approximators, but to illustrate the emergence of metastability in PINNs' learning algorithms. We did not optimize on architectures, hyperparameters, and grids. Our choices were made for simplicity. We expect the main points of this manuscript to be valid also in more general cases.

Acknowledgements

AMS is supported by the NSF award # 2152789 on *RUI: Collaborative Research: DMS/NIGMS 1: The mathematical laws of morphology and biomechanics through ontogeny*. AMS thanks the referees and the editors for the time they spent reviewing this manuscript and for their insightful comments.

Impact Statement

This paper presents work whose goal is to advance the field of Machine Learning. There are many potential societal consequences of our work, none which we feel must be specifically highlighted here.

References

[1] Antonio Ambrosetti and Andrea Malchiodi. *Perturbation Methods and Semilinear Elliptic Problems on R^n*. Birkhäuser, Basel, 2006.

[2] Antonio Ambrosetti and Andrea Malchiodi. *Nonlinear Analysis and Semilinear Elliptic Problems*. Cambridge University Press, 2010.

[3] Steven Brunton and Nathan Kutz. *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control. 2ndEdition*. Cambridge University Press, 2024.

[4] A. Carati, L. Galgani, A. Ponno, and A. Giorgilli. The fermi-pasta-ulam problem and the question of the rates of thermalization. *IL NUOVO CIMENTO B*, 9-11:1017–1026, 2002.

[5] Charles Fefferman. The millennium prize problems. In *Clay Mathematics Institute*. Clay Mathematics Institute, 2006.

[6] E. Fermi, J. Pasta, and S. Ulam. Studies of nonlinear problems. document la-1940. Technical report, Los Alamos National Laboratory, New Brunswick, MA, 1955.

[7] George Em Karniadiakis, Ioannis G. Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang. Physics-informed machine learning. *Nature Reviews Physics*, 3:422–440, 2021.

[8] M.K. Kwong. Uniqueness of positive solutions of $\delta u - u + u^p = 0$ in $\mathbb{R}^n$. *Arch. Ration. Mech. Anal.*, 105(3):243–266, 1983.

[9] Boris Malomed. Encyclopedia of nonlinear science (639–643). In *Nonlinear Schrödinger Equations*. New York: Routledge, 2005.

[10] S. Pitaevskii, L.; Stringari. *Bose-Einstein Condensation*. Oxford, U.K.: Clarendon, 2003.

[11] M. Raissi, Paris Perdikaris, and George Em Karniadiakis. 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, 2019.

[12] Catherine Sulem and Pierre-Louis Sulem. *The Nonlinear Schrödinger Equation. Self-focusing and Wave Collapse*, volume 139. Springer-Verlag, New York, 1999.

[13] Terence Tao. *Nonlinear dispersive equations: local and global analysis*. CBMS regional conference series in mathematics,, 2006.

[14] Terence Tao. A (concentration-)compact attractor for high-dimensional non-linear schrödinger equations. `arXiv:math/0611402`, 2014.

[15] N. J. Zabusky and M. D. Kruskal. Interactions of solitons in a collisionless plasma and the recurrence of initial states. *Physical Review Letters*, 15(6):240–243, 1965.