
Exciton-Polariton Condensates: A Fourier Neural Operator Approach

Surya T. Sathujoda *
University of Cambridge
sts40@cam.ac.uk

Yuan Wang *†
University of Southampton
yw2e17@soton.ac.uk

Kanishk Gandhi
University of Southampton
kg7g19@soton.ac.uk

Abstract

Advancements in semiconductor fabrication over the past decade have catalyzed extensive research into all-optical devices driven by exciton-polariton condensates. Preliminary validations of such devices, including transistors, have shown encouraging results even under ambient conditions. A significant challenge still remains for large scale application however: the lack of a robust solver that can be used to simulate complex nonlinear systems which require an extended period of time to stabilize. Addressing this need, we propose the application of a machine-learning-based Fourier Neural Operator approach to find the solution to the Gross-Pitaevskii equations coupled with extra exciton rate equations. This work marks the first direct application of Neural Operators to an exciton-polariton condensate system. Our findings show that the proposed method can predict final-state solutions to a high degree of accuracy almost 1000 times faster than CUDA-based GPU solvers. Moreover, this paves the way for potential all-optical chip design workflows by integrating experimental data.

1 Introduction

The rapid advancement of exciton-polariton condensates [1] has led to the emergence of a wide range of all-optical devices, such as switches [2–8], analogue simulators [9, 10], neuromorphic computing [11–15], transistors [16–18], etc. The microcavity exciton-polariton (hereafter polariton) [19] system consists of two main strongly coupled components: quantum well (QW) excitons and photons trapped in microcavity. The former are electron-hole pairs bound by Coulomb interactions, are frequently observed in semiconductor QWs, and the latter are generated from the distributed Bragg reflectors (DBRs) aiming to create a stopband in the refractive spectrum of the microcavity structure [19], as illustrated in Figure 1. In this system, the exciton lifetime significantly surpasses that of photons in microcavity. Thus, these high-quality DBRs, engineered to extend the photon lifetime, ensure the preservation of the strong-coupling condition, where the light-matter coupling strength surpasses the decay rate of any component within the polaritonic system [20].

Polaritons, often described as quasiparticles with characteristics that are half-light and half-matter, have an impressively low effective mass due to their photonic components. This mass is approximately five orders of magnitude less than that of a bare electron. Thus, the necessary temperature for condensation is approximately 10 K for inorganic semiconductor materials [1], which contrasts

*Equal contribution.

†Corresponding author.

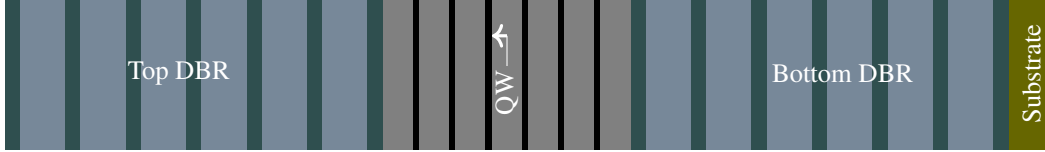


Figure 1: Sketch of the distributed Bragg reflectors (DBRs) microcavity integrated with quantum wells (QWs). The DBR is comprised of 6 bilayers of alternating low (depicted in light blue) and high (shown in dark green) refractive index materials. Also, 6 QWs (represented in black) with their respective barriers (in gray) are depicted. The substrate, located at the base of the structure, is presented in dark yellow.

significantly with atomic condensates such as Rubidium-87, which require temperatures around 170 nK [21]. Notably, using organic materials, polariton condensation can be realized even at room temperature [22–24]. In polariton condensates, due to their excitonic component, the predominant repulsive interaction between polaritons results in a blue-shifting effect on the potential and gives rise to rich nonlinear effects [25, 26]. In essence, it is the photonic (light-like) component of the polaritons that confers their notably light effective mass suitable for condensates, while the excitonic (matter-like) component is responsible for the observed nonlinear effects.

The emergence and adoption of all-optical devices necessitates the development of precise and adaptable simulation tools. Just as Electronic Design Automation (EDA) played a pivotal role in the evolution of chip design, there is a pressing need for emulators that can capture the rich of nonlinear characteristics inherent in optical devices. A case in point is the groundbreaking development of an optically-activated transistor switch. Rooted in advancements in polariton condensates, this transistor was initially conceptualized for operations at cryogenic temperatures [16]. However, strides in research have broadened its applicability, making it feasible even under ambient conditions [17, 18]. These rapid advancements emphasize the need for comprehensive simulations in the design of large-scale logic gates. Hence, crafting a robust theoretical framework becomes imperative, one that can effectively allow manipulation of the potential profile via specific input configurations.

To this aim, a robust solution to the Gross-Pitaevskii equation (GPE) which is used to theoretically describe the condensates is required [27]. While GPU-based GPE solvers tailored for both uniform [28–31] and non-uniform meshes [32] currently exist, we aim to adopt an even more efficient machine-learning (ML) based solver that is intended to notably accelerate the computational process, especially in the context of designing extensive transistor networks.

Thus far, a wide variety of ML architectures have been proposed to approximate solutions to general classes of PDEs. These range from convolution-based methods, such as variants of the U-Net architecture [33], to operator-learning methods such as Deep Operator Networks (DeepONets [34]), Graph Neural Operators [35], Multipole Graph Neural Operators [36], Fourier Neural Operators [37] and Physics-informed Neural Operators [38]. Though convolution-based methods have shown promise regarding accuracy of future state predictions, they fail to scale in compute efficiency to larger systems, even with recent advancements [39]. Operator-learning methods overcome this bottleneck by learning mappings between infinite-dimensional spaces, allowing them to predict solutions at different discretizations at a similar speed.

In this work, we study the application of the Fourier Neural Operator (FNO) architecture to approximate solutions to the GPE. We are interested in this specific variant of Neural Operator as its mathematical formulation relates very closely to the Split-step Fourier Method (SSFM), which is the numerical method used to solve the GPE in this instance. Moreover, FNOs have shown widespread success in application to many other areas of physics and engineering [40–43].

The structure of this paper is outlined as follows: In Section 2, we provide an introduction to exciton-polariton condensates and explain the physical quantities employed within the GPE. Section 3 introduces the Neural Operator architecture, detailing its compatibility with the GPE, especially when integrated with the rate equation. This section further elaborates on data generation and preparation methods. Our findings are presented in Section 4, future work in Section 5 and the general conclusions are given in Section 6.

2 Background

2.1 Exciton-polariton condensates

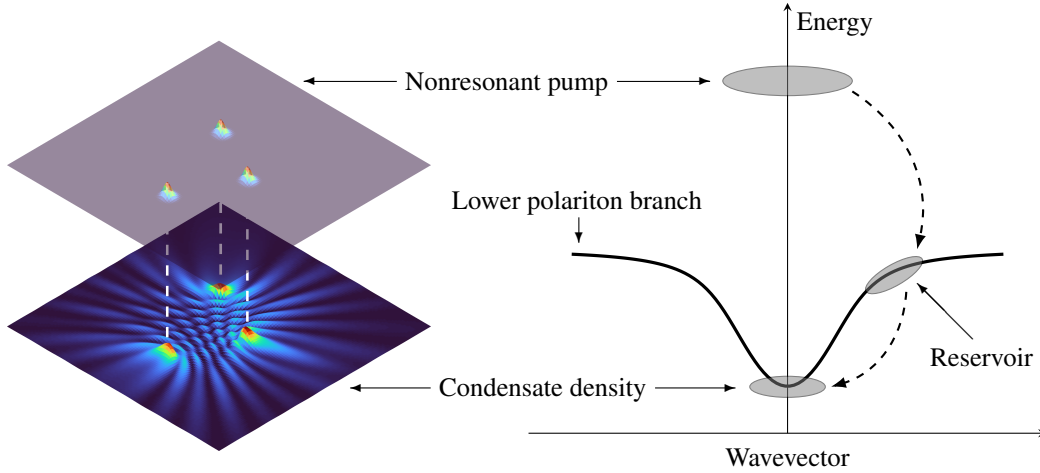


Figure 2: Comparison of pump profiles and wavefunction density with scattering process illustration. Left column: The upper panel shows the nonresonant pump profile featuring three Gaussian spots, while the lower panel shows the wavefunction density of the condensates at the final time. Three white dashed lines indicate the central positions of the pump regions and align with their corresponding locations on the condensate density map. Right column: Depiction of the scattering process, tracing the transition from the hot electron-hole plasma phase, through the reservoir cooling phase, to the scattering in the condensates. Only the lower polariton branch of the polariton energy mode is shown here.

Polariton condensates, being a non-equilibrium process, start with nonresonant excitation (see the right column of Figure 2). The hot electron-hole plasma initially undergoes rapid cooling, primarily driven by emissions mediated by the longitudinal-optical phonons, forming excitations at the high momentum of the lower polariton (LP) branch. Subsequent interactions between excitons and acoustic phonons, as well as exciton-exciton scatterings, lead polaritons to a transitional 'bottleneck region' within the LP branch, as illustrated in Figure 2. It is worth noting that the terminology 'polariton' instead of 'exciton' is used here because at very high position of the LP branch the photonic components of the polaritons are barely present and then the quasiparticles arrive all the way to the bottleneck region where more photons are coupled. By the end of the process, parametric scattering, together with momentum conservation, takes place, which results in a segment of the polaritons transitioning to a higher-momentum state, while another segment descends into the lower-momentum branch of the LP forming condensates. Ballistic polariton flow (see the left column of the Figure 2) is activated once the system is above the condensate threshold [44]. These condensates represent large-scale coherent states, distinct from the initial nonresonant inputs [1].

Technological advances, such as the use of spatial light modulators and precise semiconductor fabrications, allow the manipulation of the pump profile and rich nonlinear condensate outputs, leading to engineering of quantum fluids of light [45]. Extensive studies have been conducted on the ability of nonresonant optical techniques to manipulate the motion of condensate polaritons, such as customized momentum distributions [46], condensate amplifiers [44, 47], waveguides [48–51], directional superfluids near equilibrium [52]. The ability to manipulate the condensate flow through reservoir engineering shows its potential for quantum computing [53].

2.2 Gross-Pitaevskii equation

The dynamics of polariton condensates can be described by the GPE coupled with the rate equation of the exciton reservoir denoted as \mathcal{N} [27]:

$$i\hbar \frac{\partial}{\partial t} \Psi = \left\{ -\frac{\hbar^2}{2m} \nabla^2 + \alpha |\Psi|^2 + G \left[\mathcal{N} + \frac{\eta}{\Gamma} P(\mathbf{r}) \right] + i\frac{\hbar}{2} [R\mathcal{N} - \gamma] \right\} \Psi, \quad (1)$$

$$\frac{\partial}{\partial t} \mathcal{N} = -\left[\Gamma + R|\Psi|^2 \right] \mathcal{N} + P(\mathbf{r}), \quad (2)$$

where m is the effective mass of the polariton, α and G stands for, respectively, polariton-polariton and polariton-reservoir interaction, R denotes the scattering rate from the reservoir to the condensates, η refers to the ratio of the dark excitons, and γ (Γ) is the decay rate of the polariton (reservoir). The detuning between the exciton and the photon mode can greatly alter the interaction terms with the relationship $\alpha = g|\chi|^4$ and $G = 2g|\chi|^2$, and $g = g_0/N$ where g_0 is the exciton-exciton interaction, N_{QW} the number of QWs, and $|\chi|^2$, representing the presentage of exciton of which the polariton consists, is the Hopfield coefficient [54] of the excitonic branch. In this work, the continuous-wave (CW) pump, denoted by $P(\mathbf{r})$, is used to replenish the reservoir due to the losses from the system. The nonlinear term $|\psi|^2$ appearing in both the pump-to-reservoir transition [see Equation (2)] and the superfluid in the condensates [see Equation (1)], produce the rich nonlinear characteristic induced from the pump to the condensate.

In the case of CW excitation under a weak pumping regime, the approximate value of $|\psi|^2$ tends towards zero. In this situation, the rate of reservoir with respect to time maintains a steady state, or in mathematical terms, $\partial \mathcal{N} / \partial t = 0$. The determination of threshold power, denoted at p^{th} , is possible through an analysis of the right-hand side (r.h.s.) of Equation (1) where $R\mathcal{N} = \gamma$ serves as a representative of the equilibrium state between gain and loss. Therefore, the threshold power $p^{\text{th}} = \gamma\Gamma/R$ is obtained. This suggests that when the population of polaritons exceeds the condensation threshold p^{th} , a detectable density value manifests itself. The real potential of Equation (1) denoted V in the stationary state of the system, therefore, is

$$V(\mathbf{r}) = \alpha |\psi|^2 + G \left(\frac{1}{\Gamma + R|\psi|^2} + \frac{\eta}{\Gamma} \right) P(\mathbf{r}). \quad (3)$$

The real potential is composed of two main components: one originating from the pumping region [first term on the r.h.s of Equation (3)] and the other stemming from the interactions among the polaritons outside this region [second term on the r.h.s of Equation (3)]. When the pumping power is below the threshold, the direct contribution of the potential goes directly into the pumping profile. This relationship is represented as $V(\mathbf{r}) = (1 + \eta)(G/\Gamma)P(\mathbf{r})$. The spatial profile chosen for the demonstration of N_G Gaussian spots. That is

$$P(\mathbf{r}) = \sum_i^{N_G} p_i G_i(\mathbf{r}), \quad (4)$$

where p_i stands for strength of each spot and the normalized Gaussian function $G_i(\mathbf{r})$, with full width at half maximum (FWHM) denoted σ , is defined as

$$G_i(\mathbf{r}) = \frac{1}{2\pi\sigma} \exp\left(\frac{-|\mathbf{r} - \mathbf{r}_i|}{2\sigma^2}\right). \quad (5)$$

Note that \mathbf{r}_i represents different location of spots.

3 Methodology

3.1 Fourier Neural Operators

The numerical solution to Equations (1) and (2) is derived using the SSFM, detailed in Appendix A. A natural ML analog to this classical method is the FNO architecture [37]. More generally, Neural Operators [55] are a class of models which learn mappings between two infinite-dimensional spaces from a finite set of input-output pairs. Many variants of the Neural Operator architecture have been applied to approximate solutions to Partial Differential Equations, such as in [40–43]. The Neural

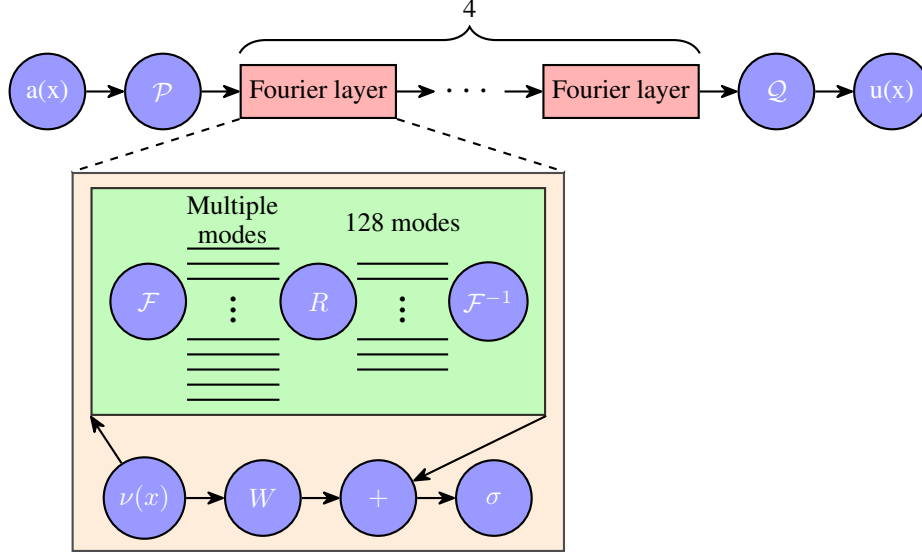


Figure 3: Architecture of the Fourier Neural Operator: The process begins with the input $a(x)$ which undergoes a lifting operation, denoted as \mathcal{P} . This is followed by 4 consecutive Fourier layers. Subsequently, a projector \mathcal{Q} transforms the data to the desired target dimension, resulting in the output $u(x)$. The inset provides a detailed view of the structure of a Fourier layer. Data initially flows to the layer as $\nu(x)$ and is bifurcated into two branches: one undergoes a linear transformation W , and the other first experiences a Fourier transformation, from which the 128 lowest Fourier modes are kept and the other higher modes are filtered out by undergoing a transformation R , and ends with an inverse Fourier transformation with these left modes. The two data streams then converge, followed by the application of an activation function σ .

Operator architecture consists of a lifting operation \mathcal{P} , followed by iterative updates using a Kernel Integral Operator \mathcal{K} , and a final projection operator \mathcal{Q} , as defined in Equation 6.

$$\mathcal{G}_\theta := \mathcal{Q} \circ \sigma_T(W_{T-1} + \mathcal{K}_{T-1} + b_{T-1}) \circ \dots \circ \sigma_1(W_0 + \mathcal{K}_0 + b_0) \circ \mathcal{P} \quad (6)$$

Here σ corresponds to a non-linearity and W and b correspond to the weights and biases of the Kernel Integral Layer, respectively. \mathcal{P} and \mathcal{Q} are point-wise fully local projection and lifting operators. The choice of the Kernel Integral Operator \mathcal{K} delineates the class of the Neural Operator. Specifically, the FNO (see Figure 3) uses the Kernel Integral Operator defined by Equation 7 below.

$$(\mathcal{K}_t(v_t))(x) = \mathcal{F}^{-1}(R_\phi \cdot \mathcal{F}(v_{t-1}))(x) \quad \forall x \in \mathbb{R}^n \quad (7)$$

Here \mathcal{F} and \mathcal{F}^{-1} correspond to the Fourier and Inverse Fourier Transforms and R_ϕ corresponds to the Fourier Transform of a periodic function arising from the definition of a Kernel Integral Operator given in [37]. This object is parameterised by a linear transformation of the top k modes pertaining to the given layer, which acts as a hyperparameter in the model.

The natural choice of the FNO architecture for approximating the solution to Equations 1 and 2 is due to the inductive bias that arises from the SSFM - FNO correspondence stated below.

Theorem 1. (SSFM-FNO Correspondence) Suppose that $\sigma \in (TW)$ is a Tauber-Wiener function, X is a Banach Space, $K \subset X$ is a compact set, V is a compact set in $C(K)$, Ψ_t is a nonlinear continuous operator representing the solution of the first-order Split-step Fourier Method at time t , then for any $\epsilon > 0$, there are a positive integer n , m points $x_1, \dots, x_m \in K$, and real constants c_i, θ_i, ξ_{ij} , $i = 1, \dots, n$, $j = 1, \dots, m$, such that for

$$R_\phi := \sum_{i=1}^n c_i \sigma \left(\sum_{j=1}^m \xi_{ij} u(x_j) + \theta_i \right), \quad (8)$$

$$\left| \Psi_{t+\Delta t}(\Psi_t)(x) - \mathcal{F}^{-1}(R_\phi \cdot \mathcal{F}(v_t))(x) \right| < \epsilon \quad (9)$$

holds for all $u \in V$.

Proof. See Appendix B.

3.2 Data Generation

The training and testing datasets are constructed based on varying pump profiles, $P(\mathbf{r})$, as elucidated in Equation (4). This profile is characterized by four Gaussian spots, represented by $N_G = 4$. Out of these, three spots have their power adjusted to exceed the threshold, specifically $p_i = 1.6 p^{\text{th}}$, while the power of the remaining spot is set below the threshold at $p_i = 0.5 p^{\text{th}}$. Each spot's spatial profile is represented as $G_i(\mathbf{r})$. Moreover, the boundary conditions that the condensate density vanishes near the computation frame in both real and Fourier space are also applied.

These profiles are stochastically determined within a square region that measures $64 \mu\text{m} \times 64 \mu\text{m}$ out of the entire configuration with $128 \mu\text{m} \times 128 \mu\text{m}$. The region where Gaussian spots stay is smaller than the full grid, to make sure that they are still far from the region where the boundary condition is applied. Care has also been taken to ensure that the Gaussian spots are non-overlapping. Additionally, every pump profile is unique, and among the spots with power exceeding the threshold, each one is distinct from the others, thereby eliminating any potential redundancy. Given $0.5 \mu\text{m}$ resolution per pixel per dimension, the total datasets for pump configuration is with size $256 \times 256 \times 1568$ where 256 represents each square map size per dimension and 1568 is the number of different pump configurations (of which 314 configurations are used for testing and 1254 for training respectively). The datasets for the density map is with size $256 \times 256 \times 2 \times 1568$ where 2 refers to the density at the initial and final time.

It is worth mentioning that the systems of all the datasets are chosen with system only at stationary state with single energy mode, which means that the results with multiple energy modes are excluded. In multimode cases, the wavefunction density changes at different times, which can be found in experiments [56, 57].

For better emulating the experiment $\sigma \approx 0.85 \mu\text{m}$, standing for FWHM of each Gaussian spot equaling to $2 \mu\text{m}$, is chosen. The simulation is based on InGaAs QWs [58] with slightly negatively detuned cavities. The parameters are the following: $m = 0.28 \text{ meV ps}^2 \mu\text{m}^{-2}$, $|\chi|^2 = 0.4$, $N_{\text{QW}} = 6$, $g_0 = 0.01 \text{ meV } \mu\text{m}^2$, $\hbar R = 10g$, $\eta = 2$, and $\gamma^{-1} = \Gamma^{-1} = 5.5 \text{ ps}$.

4 Results

In Figure 4, we present model predictions for 4 representative test cases. Note that we have carefully chosen four distinct pump profiles to visualize model performance over varying inputs. As we see from the second row of the figure, the model is highly accurate in predicting the steady-state solution for the GPE. It is not only able to predict accurate values at pump locations and their locality but it is also able to predict direction of flow for the wave fronts as well as the scattering from the barrier produced from the spot below the threshold. Where the model exceeds expectations most is in the detail to which it also captures the interference pattern for different pump configurations. We see an almost identical pattern, including the parity of fringes among spots, between the predictions and simulation ground truths. The parity of these fringes are responsive to the distance between spots (see experiment in [57]), which also indicates that our model is capable of capturing these details.

Looking at error plots between predictions and ground truth values, we see that the highest errors occur very close to the pump locations. This is due to the high degree of nonlinearity which is exhibited near this region and the information loss that occurs from the Fast Fourier Transform cutoff modes in the FNO architecture. We see other smaller errors around the edges of wave fronts which could potentially be tackled using physics-informed gradient losses, which we intend to test in future work.

We present a full overview of cell averaged absolute difference errors in Figure D1 in Appendices D for all 314 test cases. The 4 test predictions shown in Figure 4 correspond to cases 75, 97, 181 and 291 respectively. The cell averaged errors on the test set range between 1.774×10^{-2} , for case 198, and 8.080×10^{-2} , for case 31. Empirically, the lower errors correspond to pump configurations where distance between the pumps is smaller, leading to better interference predictions. The higher errors correspond to pump configurations where at least one pump is far away from the other two,

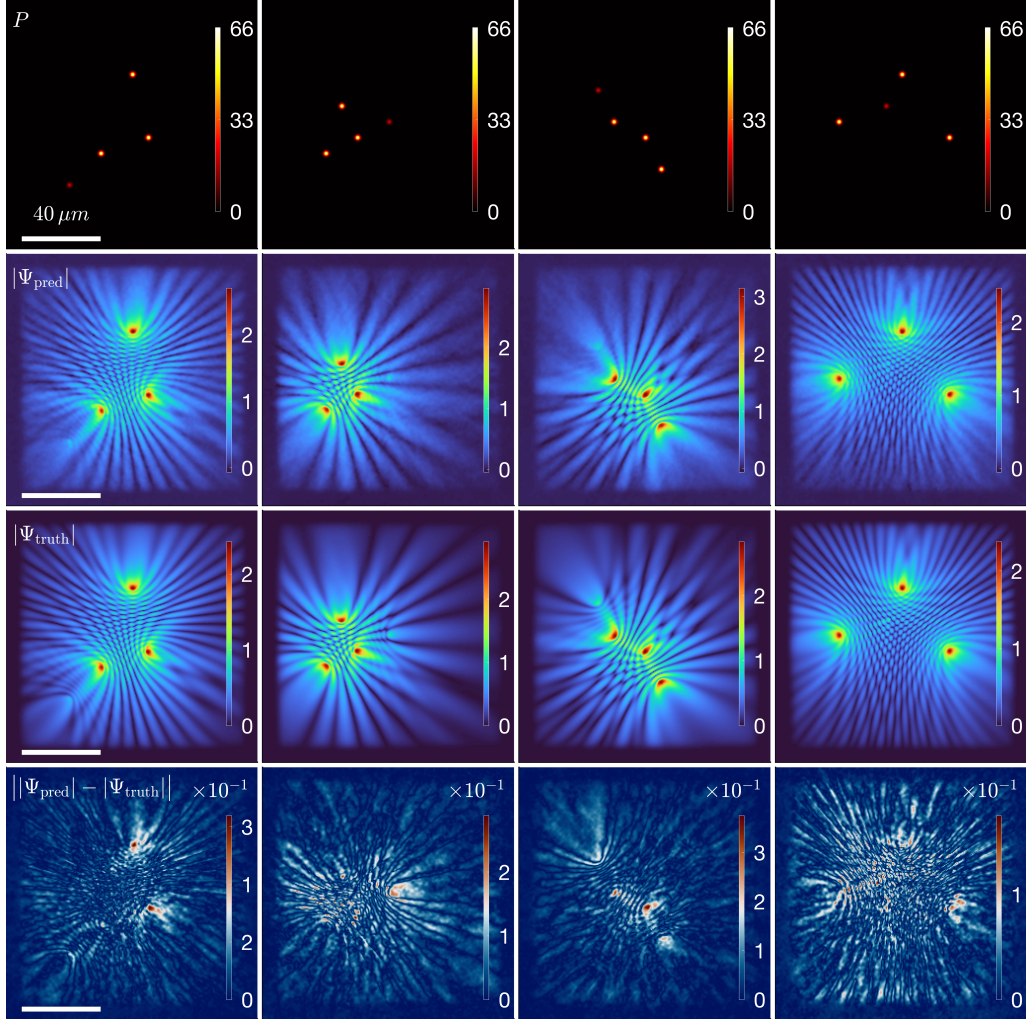


Figure 4: Comparison of the prediction from the Fourier Neural Operator approach. The columns, left to right, show different pump configurations. The rows, from top to bottom, are pump profiles P , predictions $|\Psi_{\text{pred}}|$, ground truths $|\Psi_{\text{truth}}|$, and errors $||\Psi_{\text{pred}}| - |\Psi_{\text{truth}}||$.

leading to worse interference pattern predictions. This is evident in the results from Figure D2 in Appendices D where pumps are very far apart as opposed to Figure 4 where pumps are closer together.

We found that, when we trained the FNO with a cutoff smaller than 128 Fourier modes for both spatial directions, which is conventional, that the model was not able to capture nonlinearities in the system to a great extent and that it would take longer to converge. We also observed that predictions were accurate in a large square sub-region of the grid but performance dropped outside this region. This was likely due to the fact that with a 256×256 grid, in order to capture the finer details, we would need to also include higher frequency modes of the Fourier expansion. To this avail, we increased the cutoff to 128 Fourier modes in both directions and observed a significant improvement in performance. For larger grid sizes, there is a higher probability that the distance between the randomly generated pump locations will be greater. This would require us to increase further the number of Fourier modes to retain similar levels of accuracy or employ dimensionality reduction techniques.

The hyperparameters are detailed in Appendix C. It is worth noting that the usage of the domain padding can reduce the artifacts around the boundary conditions. Moreover, the loss function named 'H1Loss' (see Appendix C) considers not only its true values but also in their derivatives. With both the FNO model and numerical solver run in parallel computing using a GPU, the time to predict

solutions for 314 cases using the CUDA-based numerical method took 2808.03 s, whereas the FNO model took 2.66 s.

5 Future Work

In this work, we use numerical datasets for training and prediction. However, in the future we hope to directly use experimentally collected data to train the model on ground truth physics. Pump profiles and wavefunction density maps shown in Figure 4 can be obtained directly from spatial light modulator and photoluminescence spectroscopy, respectively. Due to the improvement in semiconductor fabrication, clear interference patterns, which are very close to those simulated shown here, can be found from inorganic semiconductor materials [57, 59]. Furthermore, with the help of a streak camera, photoluminescence can be captured at the picosecond level, which makes it possible to make predictions of a time-resolved condensate formation on the basis of purely experimental data.

Various general ML methods have been proposed to incorporate underlying physics-based losses and information into the model to aid the learning task, such as in [60–63]. In this work, we have taken a purely data-driven approach to training, however, we believe that incorporating additional physics-informed loss terms will strictly increase the rate of convergence and accuracy. This is especially appealing given that we have a strong theoretical understanding of the underlying system.

In the future, we aim to also propose a novel Neural Operator architecture entirely, tailored to closer align to the computational procedure of the SSFM. As shown in Appendices A and B, the FNO can be seen as a first-order approximation to the SSFM. We aim to instead take structure from the BCH formula at second-order and embed this in the model. This will better guide convergence dynamics in the weight landscape, which in theory will produce more accurate solutions at a faster rate.

6 Conclusions

In the present study, we explored the potential of the FNO in the context of polariton condensates. Our findings, as detailed in Section 4, demonstrate a notable alignment with the simulation data with an approximate $1000\times$ speed up in solution generation compared to CUDA-based GPU solvers. This research paves the way for the conceptualization and development of advanced large-scale all-optical devices. Furthermore, this method draws parallels with the principles of EDA traditionally used in chip design. It introduces an innovative avenue to meet the growing demand for fast and reliable solutions in the realm of all-optical chip design.

7 Acknowledgements

We are greatly indebted to Prof. Pavlos Lagoudakis for giving feedback on this manuscript.

References

- [1] J. Kasprzak, M. Richard, S. Kundermann, A. Baas, P. Jeambrun, J. M. J. Keeling, F. M. Marchetti, M. H. Szymańska, R. André, J. L. Staehli, V. Savona, P. B. Littlewood, B. Deveaud, and Le Si Dang. Bose–einstein condensation of exciton polaritons. *Nature*, 443(7110):409–414, sep 2006.
- [2] A. Amo, T. C. H. Liew, C. Adrados, R. Houdré, E. Giacobino, A. V. Kavokin, and A. Bramati. Exciton–polariton spin switches. *Nature Photonics*, 4(6):361–366, apr 2010.
- [3] M. De Giorgi, D. Ballarini, E. Cancellieri, F. M. Marchetti, M. H. Szymanska, C. Tejedor, R. Cingolani, E. Giacobino, A. Bramati, G. Gigli, and D. Sanvitto. Control and ultrafast dynamics of a two-fluid polariton switch. *Physical Review Letters*, 109(26):266407, dec 2012.
- [4] T. Gao, P. S. Eldridge, T. C. H. Liew, S. I. Tsintzos, G. Stavrinidis, G. Deligeorgis, Z. Hatzopoulos, and P. G. Savvidis. Polariton condensate transistor switch. *Physical Review B*, 85(23):235102, jun 2012.

- [5] Alexander Dreismann, Hamid Ohadi, Yago del Valle-Inclan Redondo, Ryan Balili, Yuri G. Rubo, Simeon I. Tsintzos, George Deligeorgis, Zacharias Hatzopoulos, Pavlos G. Savvidis, and Jeremy J. Baumberg. A sub-femtojoule electrical spin-switch based on optically trapped polariton condensates. *Nature Materials*, 15(10):1074–1078, aug 2016.
- [6] Xuekai Ma, Bernd Berger, Marc Aßmann, Rodislav Driben, Torsten Meier, Christian Schneider, Sven Höfling, and Stefan Schumacher. Realization of all-optical vortex switching in exciton-polariton condensates. *Nature Communications*, 11(1), feb 2020.
- [7] Jiangang Feng, Jun Wang, Antonio Fieramosca, Ruiqi Bao, Jiabin Zhao, Rui Su, Yutian Peng, Timothy C. H. Liew, Daniele Sanvitto, and Qihua Xiong. All-optical switching based on interacting exciton polaritons in self-assembled perovskite microwires. *Science Advances*, 7(46), nov 2021.
- [8] Fei Chen, Hui Li, Hang Zhou, Song Luo, Zheng Sun, Ziyu Ye, Fenghao Sun, Jiawei Wang, Yuanlin Zheng, Xianfeng Chen, Huailiang Xu, Hongxing Xu, Tim Byrnes, Zhanghai Chen, and Jian Wu. Optically controlled femtosecond polariton switch at room temperature. *Physical Review Letters*, 129(5):057402, jul 2022.
- [9] Natalia G. Berloff, Matteo Silva, Kirill Kalinin, Alexis Askitopoulos, Julian D. Töpfer, Pasquale Cilibrizzi, Wolfgang Langbein, and Pavlos G. Lagoudakis. Realizing the classical XY hamiltonian in polariton simulators. *Nature Materials*, 16(11):1120–1126, sep 2017.
- [10] Pavlos G Lagoudakis and Natalia G Berloff. A polariton graph simulator. *New Journal of Physics*, 19(12):125008, dec 2017.
- [11] Andrzej Opala, Sanjib Ghosh, Timothy C.H. Liew, and Michał Matuszewski. Neuromorphic computing in ginzburg-landau polariton-lattice systems. *Physical Review Applied*, 11(6):064029, jun 2019.
- [12] Dario Ballarini, Antonio Gianfrate, Riccardo Panico, Andrzej Opala, Sanjib Ghosh, Lorenzo Dominici, Vincenzo Ardizzone, Milena De Giorgi, Giovanni Lerario, Giuseppe Gigli, Timothy C. H. Liew, Michał Matuszewski, and Daniele Sanvitto. Polaritonic neuromorphic computing outperforms linear classifiers. *Nano Letters*, 20(5):3506–3512, apr 2020.
- [13] Rafał Mirek, Andrzej Opala, Paolo Comaron, Magdalena Furman, Mateusz Król, Krzysztof Tyszka, Bartłomiej Seredyński, Dario Ballarini, Daniele Sanvitto, Timothy C. H. Liew, Wojciech Pacuski, Jan Suffczyński, Jacek Szczytko, Michał Matuszewski, and Barbara Piętka. Neuromorphic binarized polariton networks. *Nano Letters*, 21(9):3715–3720, feb 2021.
- [14] Sanjib Ghosh, Kohei Nakajima, Tanjung Krisnanda, Keisuke Fujii, and Timothy C. H. Liew. Quantum neuromorphic computing with reservoir computing networks. *Advanced Quantum Technologies*, 4(9), jul 2021.
- [15] A. Opala, R. Panico, V. Ardizzone, B. Piętka, J. Szczytko, D. Sanvitto, M. Matuszewski, and D. Ballarini. Training a neural network with exciton-polariton optical nonlinearity. *Physical Review Applied*, 18(2):024028, aug 2022.
- [16] D. Ballarini, M. De Giorgi, E. Cancellieri, R. Houdré, E. Giacobino, R. Cingolani, A. Bramati, G. Gigli, and D. Sanvitto. All-optical polariton transistor. *Nature Communications*, 4(1), apr 2013.
- [17] Anton V. Zasedatelev, Anton V. Baranikov, Darius Urbonas, Fabio Scafirimuto, Ullrich Scherf, Thilo Stöferle, Rainer F. Mahrt, and Pavlos G. Lagoudakis. A room-temperature organic polariton transistor. *Nature Photonics*, 13(6):378–383, mar 2019.
- [18] Anton V. Zasedatelev, Anton V. Baranikov, Denis Sannikov, Darius Urbonas, Fabio Scafirimuto, Vladislav Yu. Shishkov, Evgeny S. Andrianov, Yurii E. Lozovik, Ullrich Scherf, Thilo Stöferle, Rainer F. Mahrt, and Pavlos G. Lagoudakis. Single-photon nonlinearity at room temperature. *Nature*, 597(7877):493–497, sep 2021.
- [19] C. Weisbuch, M. Nishioka, A. Ishikawa, and Y. Arakawa. Observation of the coupled exciton-photon mode splitting in a semiconductor quantum microcavity. *Physical Review Letters*, 69(23):3314–3317, dec 1992.

- [20] Anton Frisk Kockum, Adam Miranowicz, Simone De Liberato, Salvatore Savasta, and Franco Nori. Ultrastrong coupling between light and matter. *Nature Reviews Physics*, 1(1):19–40, jan 2019.
- [21] M. H. Anderson, J. R. Ensher, M. R. Matthews, C. E. Wieman, and E. A. Cornell. Observation of bose-einstein condensation in a dilute atomic vapor. *Science*, 269(5221):198–201, jul 1995.
- [22] S. Christopoulos, G. Baldassarri Höger von Högersthal, A. J. D. Grundy, P. G. Lagoudakis, A. V. Kavokin, J. J. Baumberg, G. Christmann, R. Butté, E. Feltin, J.-F. Carlin, and N. Grandjean. Room-temperature polariton lasing in semiconductor microcavities. *Phys. Rev. Lett.*, 98:126405, Mar 2007.
- [23] Daniele Sanvitto and Stéphane Kéna-Cohen. The road towards polaritonic devices. *Nature materials*, 15(10):1061–1073, 2016.
- [24] Rui Su, Jun Wang, Jiaxin Zhao, Jun Xing, Weijie Zhao, Carole Diederichs, Timothy CH Liew, and Qihua Xiong. Room temperature long-range coherent exciton polariton condensate flow in lead halide perovskites. *Science advances*, 4(10):eaau0244, 2018.
- [25] E. Wertz, L. Ferrier, D. D. Solnyshkov, R. Johné, D. Sanvitto, A. Lemaître, I. Sagnes, R. Grousseau, A. V. Kavokin, P. Senellart, G. Malpuech, and J. Bloch. Spontaneous formation and optical manipulation of extended polariton condensates. *Nature Physics*, 6(11):860–864, aug 2010.
- [26] Johannes Schmutzler, Tomasz Kazimierczuk, Ömer Bayraktar, Marc Aßmann, Manfred Bayer, Sebastian Brodbeck, Martin Kamp, Christian Schneider, and Sven Höfling. Influence of interactions with noncondensed particles on the coherence of a one-dimensional polariton condensate. *Physical Review B*, 89(11):115119, mar 2014.
- [27] Michiel Wouters and Iacopo Carusotto. Excitations in a nonequilibrium bose-einstein condensate of exciton polaritons. *Physical Review Letters*, 99(14):140402, oct 2007.
- [28] Vladimir Lončar, Antun Balaž, Aleksandar Bogojević, Srdjan Škrbić, Paulsamy Muruganandam, and Sadhan K. Adhikari. CUDA programs for solving the time-dependent dipolar gross-pitaevskii equation in an anisotropic trap. *Computer Physics Communications*, 200:406–410, mar 2016.
- [29] James Schloss and Lee O’Riordan. GPUE: Graphics processing unit gross-pitaevskii equation solver. *Journal of Open Source Software*, 3(32):1037, dec 2018.
- [30] Joshua P. Wilson. Generalized finite-difference time-domain method with absorbing boundary conditions for solving the nonlinear schrödinger equation on a GPU. *Computer Physics Communications*, 235:279–292, feb 2019.
- [31] Benjamin D. Smith, Logan W. Cooke, and Lindsay J. LeBlanc. GPU-accelerated solutions of the nonlinear schrödinger equation for simulating 2d spinor BECs. *Computer Physics Communications*, 275:108314, jun 2022.
- [32] Markus Kivioja, Sanna Mönkölä, and Tuomo Rossi. GPU-accelerated time integration of gross-pitaevskii equation with discrete exterior calculus. *Computer Physics Communications*, 278:108427, sep 2022.
- [33] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pages 234–241, Cham, 2015. Springer International Publishing.
- [34] Lu Lu, Pengzhan Jin, Guofei Pang, Zhongqiang Zhang, and George Em Karniadakis. Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. *Nature Machine Intelligence*, 3(3):218–229, mar 2021.
- [35] Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Neural operator: Graph kernel network for partial differential equations. In *Advances in Neural Information Processing Systems*, 2020.

- [36] Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Multipole graph neural operator for parametric partial differential equations. In *ICLR 2020 Workshop on Integration of Deep Neural Models and Differential Equations*, 2020.
- [37] Zongyi Li, Nikola Borislavov Kovachki, Kamyar Azizzadenesheli, Burigede liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier neural operator for parametric partial differential equations. In *International Conference on Learning Representations*, 2021.
- [38] Zongyi Li, Hongkai Zheng, Nikola Kovachki, David Jin, Haoxuan Chen, Burigede Liu, Kamyar Azizzadenesheli, and Anima Anandkumar. Physics-informed neural operator for learning partial differential equations, 2021.
- [39] Surya Sathujoda and Soham M Sheth. Physics-informed localized learning for advection-diffusion-reaction systems. In *ICML Workshop on New Frontiers in Learning, Control, and Dynamical Systems*, 2023.
- [40] Vignesh Gopakumar, Stanislas Pamela, Lorenzo Zanisi, Zongyi Li, Anima Anandkumar, and MAST Team. Fourier neural operator for plasma modelling. In *Conference on Neural Information Processing Systems, Workshop on AI4Science*, 2021.
- [41] Gege Wen, Zongyi Li, Kamyar Azizzadenesheli, Anima Anandkumar, and Sally M. Benson. U-fno—an enhanced fourier neural operator-based deep-learning model for multiphase flow. *Advances in Water Resources*, 163:104180, 2022.
- [42] Tianze Zhang, Daniel Trad, and Kristopher Innanen. Learning to solve the elastic wave equation with Fourier neural operators. *Geophysics*, 88(3):T101–T119, 04 2023.
- [43] Zhijie Li, Wenhui Peng, Zelong Yuan, and Jianchun Wang. Fourier neural operator approach to large eddy simulation of three-dimensional turbulence. *Theoretical and Applied Mechanics Letters*, 12(6):100389, 2022.
- [44] E. Wertz, A. Amo, D. D. Solnyshkov, L. Ferrier, T. C. H. Liew, D. Sanvitto, P. Senellart, I. Sagnes, A. Lemaître, A. V. Kavokin, G. Malpuech, and J. Bloch. Propagation and amplification dynamics of 1d polariton condensates. *Physical Review Letters*, 109(21):216404, nov 2012.
- [45] Iacopo Carusotto and Cristiano Ciuti. Quantum fluids of light. *Reviews of Modern Physics*, 85(1):299–366, feb 2013.
- [46] M. Aßmann, F. Veit, M. Bayer, A. Löffler, S. Höfling, M. Kamp, and A. Forchel. All-optical control of quantized momenta on a polariton staircase. *Physical Review B*, 85(15):155320, apr 2012.
- [47] Dominik Niemietz, Johannes Schmutzler, Przemyslaw Lewandowski, Karol Winkler, Marc Aßmann, Stefan Schumacher, Sebastian Brodbeck, Martin Kamp, Christian Schneider, Sven Höfling, and Manfred Bayer. Experimental realization of a polariton beam amplifier. *Physical Review B*, 93(23):235301, jun 2016.
- [48] Johannes Schmutzler, Przemyslaw Lewandowski, Marc Aßmann, Dominik Niemietz, Stefan Schumacher, Martin Kamp, Christian Schneider, Sven Höfling, and Manfred Bayer. All-optical flow control of a polariton condensate using nonresonant excitation. *Physical Review B*, 91(19):195308, may 2015.
- [49] Peter Cristofolini, Z. Hatzopoulos, Pavlos G. Savvidis, and Jeremy J. Baumberg. Generation of quantized polaritons below the condensation threshold. *Physical Review Letters*, 121(6):067401, aug 2018.
- [50] Y. Wang, H. Sigurdsson, J. D. Töpfer, and P. G. Lagoudakis. Reservoir optics with exciton-polariton condensates. *Physical Review B*, 104(23):235306, dec 2021.
- [51] Denis Aristov, Stepan Baryshev, Julian D. Töpfer, Helgi Sigurðsson, and Pavlos G. Lagoudakis. Directional planar antennae in polariton condensates. *Applied Physics Letters*, 123(12), sep 2023.

- [52] Franziska Barkhausen, Matthias Pukrop, Stefan Schumacher, and Xuekai Ma. Structuring coflowing and counterflowing currents of polariton condensates in concentric ring-shaped and elliptical potentials. *Physical Review B*, 103(7):075305, feb 2021.
- [53] Alexey Kavokin, Timothy C. H. Liew, Christian Schneider, Pavlos G. Lagoudakis, Sebastian Klemmt, and Sven Hoefling. Polariton condensates for classical and quantum computing. *Nature Reviews Physics*, 4(7):435–451, apr 2022.
- [54] J. J. Hopfield. Theory of the contribution of excitons to the complex dielectric constant of crystals. *Physical Review*, 112(5):1555–1567, dec 1958.
- [55] Nikola Kovachki, Zongyi Li, Burigede Liu, Kamyar Azizzadenesheli, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Neural operator: Learning maps between function spaces with applications to pdes. *Journal of Machine Learning Research*, 24(89):1–97, 2023.
- [56] D. N. Krizhanovskii, K. G. Lagoudakis, M. Wouters, B. Pietka, R. A. Bradley, K. Guda, D. M. Whittaker, M. S. Skolnick, B. Deveaud-Plédran, M. Richard, R. André, and Le Si Dang. Coexisting nonequilibrium condensates with long-range spatial coherence in semiconductor microcavities. *Physical Review B*, 80(4):045317, jul 2009.
- [57] J. D. Töpfer, H. Sigurdsson, L. Pickup, and P. G. Lagoudakis. Time-delay polaritonics. *Communications Physics*, 3(1), jan 2020.
- [58] Pasquale Cilibrizzi, Alexis Askitopoulos, Matteo Silva, Faebian Bastiman, Edmund Clarke, Joanna M. Zajac, Wolfgang Langbein, and Pavlos G. Lagoudakis. Polariton condensation in a strain-compensated planar microcavity with InGaAs quantum wells. *Applied Physics Letters*, 105(19), nov 2014.
- [59] J. D. Töpfer, I. Chatzopoulos, H. Sigurdsson, T. Cookson, Y. G. Rubo, and P. G. Lagoudakis. Engineering spatial coherence in lattices of polariton condensates. *Optica*, 8(1):106, jan 2021.
- [60] M. Raissi, P. Perdikaris, and G.E. 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, feb 2019.
- [61] George Em Karniadakis, Ioannis G. Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang. Physics-informed machine learning. *Nature Reviews Physics*, 3(6):422–440, may 2021.
- [62] Lu Lu, Xuhui Meng, Zhiping Mao, and George Em Karniadakis. DeepXDE: A deep learning library for solving differential equations. *SIAM Review*, 63(1):208–228, jan 2021.
- [63] Salvatore Cuomo, Vincenzo Schiano Di Cola, Fabio Giampaolo, Gianluigi Rozza, Maziar Raissi, and Francesco Piccialli. Scientific machine learning through physics-informed neural networks: Where we are and what’s next. *Journal of Scientific Computing*, 92(3), jul 2022.
- [64] G. Cybenko. Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals, and Systems (MCSS)*, 2(4):303–314, December 1989.
- [65] T. Chen and H. Chen. Approximations of continuous functionals by neural networks with application to dynamic systems. *IEEE Transactions on Neural Networks*, 4(6):910–918, 1993.
- [66] Tianping Chen and Hong Chen. Approximation capability to functions of several variables, nonlinear functionals, and operators by radial basis function neural networks. *IEEE Transactions on Neural Networks*, 6(4):904–910, 1995.
- [67] Tianping Chen and Hong Chen. Universal approximation to nonlinear operators by neural networks with arbitrary activation functions and its applications to dynamic systems. *IEEE Transactions on Neural Networks*, pages 911 – 917, 08 1995.

Appendix A Split-step Fourier method

In this Appendix, the numerical solution to Equations (1) and (2) is derived using SSFM, which serves as the supplementary information for Section 3.1. Let us start with the GPE first and Equation 1 can be rearranged as

$$\frac{\partial}{\partial t} \Psi = \left\{ i \frac{\hbar}{2m} \nabla^2 - i \frac{\alpha}{\hbar} |\Psi|^2 - i \frac{G}{\hbar} \left[\mathcal{N} + \frac{\eta}{\Gamma} P \right] + \frac{1}{2} [R\mathcal{N} - \gamma] \right\} \Psi. \quad (\text{A1})$$

The direct solution from Equation (A1) at time interval $[t, t + \Delta t]$ is

$$\Psi_{t+\Delta t} = \exp[(\hat{f}_L + \hat{f}_N)\Delta t] \Psi_t, \quad (\text{A2})$$

where

$$\hat{f}_L = i \frac{\hbar}{2m} \nabla^2, \quad (\text{A3})$$

$$\hat{f}_N = -i \frac{\alpha}{\hbar} |\Psi|^2 - i \frac{G}{\hbar} \left[\mathcal{N} + \frac{\eta}{\Gamma} P \right] + \frac{1}{2} [R\mathcal{N} - \gamma]. \quad (\text{A4})$$

Note that since $[\hat{f}_L, \hat{f}_N] \neq 0$, the exponentiation identity cannot be applied directly, namely, $\exp[\hat{f}_L + \hat{f}_N] \neq \exp[\hat{f}_L] \exp[\hat{f}_N]$. By applying the Baker–Campbell–Hausdorff (BCH) formula at second order and the strang splitting, we have

$$\Psi_{t+\Delta t} = \exp\left[\frac{1}{2}\hat{f}_N\Delta t\right] \exp[\hat{f}_L\Delta t] \exp\left[\frac{1}{2}\hat{f}_N\Delta t\right] \Psi_t, \quad (\text{A5})$$

which give the accuracy of Δt^3 . The essence of SSFM is that it can convert the nonlinear operator \hat{f}_N into the linear one through the Fourier transform. We can always construct the relation that

$$\frac{\partial}{\partial t} \Psi = \hat{f}_L \Psi. \quad (\text{A6})$$

Then, applying the Fourier transform for above equation, we have

$$\frac{\partial}{\partial t} \mathcal{F}[\Psi] = \hat{f}_P \mathcal{F}[\Psi], \quad (\text{A7})$$

where

$$\hat{f}_P = -i \frac{\hbar}{2m} |\mathbf{k}|^2, \quad (\text{A8})$$

is the momentum operator. This implies that Equation A6 can be rewritten as

$$\Psi_{t+\Delta t} = \exp\left[\frac{1}{2}\hat{f}_N\Delta t\right] \mathcal{F}^{-1} \left[\exp[\hat{f}_P\Delta t] \mathcal{F} \left[\exp\left[\frac{1}{2}\hat{f}_N\Delta t\right] \Psi_t \right] \right]. \quad (\text{A9})$$

Ultimately, the final state wavefunction can be obtained by iteration of Equation (A9) overall the infinitesimal time steps. During each iteration, \mathcal{N} is updated within the operator \hat{f}_N . That is

$$\mathcal{N}_{t+\Delta t} = \exp[(-\Gamma + R|\Psi|^2)\Delta t] \mathcal{N}_t + P\Delta t. \quad (\text{A10})$$

The kernel defined by Equation (A5) is used to calculate the ground truth for the simulation with BCH at second order. However, if we limit ourselves at BCH at first order, then corresponding Equation (A9) can be simplified as

$$\Psi_{t+\Delta t} = \mathcal{F}^{-1} \left[\exp[\hat{f}_P\Delta t] \mathcal{F} \left[\exp[\hat{f}_N\Delta t] \Psi_t \right] \right]. \quad (\text{A11})$$

From here, Equation (A11) takes the functional form of the Kernel Integral Operator of the FNO introduced in Section 3.1. This is stated formally in the SSFM - FNO correspondence given in Theorem 1.

Appendix B Theorems

Theorem B1. (Universal Approximation Theorem for Functionals.) Suppose that $\sigma \in (TW)$ is a Tauber-Wiener function, X is a Banach Space, $K \subset X$ is a compact set, V is a compact set in $C(K)$, f is a continuous functional defined on V , then for any $\epsilon > 0$, there are a positive integer n , m points $x_1, \dots, x_m \in K$, and real constants $c_i, \theta_i, \xi_{ij}, i = 1, \dots, n, j = 1, \dots, m$, such that

$$\left| f(u) - \sum_{i=1}^n c_i \sigma \left(\sum_{j=1}^m \xi_{ij} u(x_j) + \theta_i \right) \right| < \epsilon \quad (\text{B1})$$

holds for all $u \in V$.

Proof. See [S34] and [S64–S66].

Theorem B2. (Universal Approximation Theorem for Operators.) Suppose that $\sigma \in (TW)$ is a Tauber-Wiener function, X is a Banach Space, $K_1 \subset X, K_2 \subset \mathbb{R}^d$ are two compact sets in X and \mathbb{R}^d , respectively, V is a compact set in $C(K_1)$, G is a nonlinear continuous operator, which maps V into $C(K_2)$, then for any $\epsilon > 0$, there are positive integers n, p, m , constants $c_i^k, \xi_{ij}^k, \theta_i^k, \zeta_k \in \mathbb{R}, w_k \in \mathbb{R}^d, x_j \in K_1, i = 1, \dots, n, k = 1, \dots, p, j = 1, \dots, m$, such that

$$\left| G(u)(y) - \sum_{k=1}^p \sum_{i=1}^n c_i^k \sigma \left(\sum_{j=1}^m \xi_{ij}^k u(x_j) + \theta_i^k \right) \sigma(w_k \cdot y + \zeta_k) \right| < \epsilon \quad (\text{B2})$$

holds for all $u \in V$ and $y \in K_2$.

Proof. See [S34] and [S64, S67, S66].

Theorem B3. SSFM-FNO Correspondence. Suppose that $\sigma \in (TW)$ is a Tauber-Wiener function, X is a Banach Space, $K \subset X$ is a compact set, V is a compact set in $C(K)$, Ψ_t is a nonlinear continuous operator representing the solution of the first-order Split-step Fourier Method at time t , then for any $\epsilon > 0$, there are a positive integer n, m points $x_1, \dots, x_m \in K$, and real constants $c_i, \theta_i, \xi_{ij}, i = 1, \dots, n, j = 1, \dots, m$, such that for

$$R_\phi := \sum_{i=1}^n c_i \sigma \left(\sum_{j=1}^m \xi_{ij} u(x_j) + \theta_i \right), \quad (\text{B3})$$

$$\left| \Psi_{t+\Delta t}(\Psi_t)(x) - \mathcal{F}^{-1}(R_\phi \cdot \mathcal{F}(v_t))(x) \right| < \epsilon \quad (\text{B4})$$

holds for all $u \in V$.

Proof. Limiting the solution of Equation 1 to first-order in the BCH formula, we get

$$\Psi_{t+\Delta t}(\Psi_t)(x) = \mathcal{F}^{-1} \left[\exp[\widehat{f}_P \Delta t] \mathcal{F} \left[\exp[\widehat{f}_N \Delta t] \Psi_t \right] \right]. \quad (\text{B5})$$

Also considering the Kernel Integral Operator for the FNO as defined in Equation 7, we have

$$(\mathcal{K}_{t+1}(v_t))(x) = \mathcal{F}^{-1}(R_\phi \cdot \mathcal{F}(v_t))(x) \quad \forall x \in \mathbb{R}^n \quad (\text{B6})$$

It is clear that these two equations are equivalent for some $R_\phi = \exp[\widehat{f}_P \Delta t]$ and $v_t = \exp[\widehat{f}_N \Delta t] \Psi_t$. Now, appealing to the Universal Approximation Theorems B1 and B2, and using the fact that a non-linearity σ is applied following every Kernel Integral layer, it is clear that there exists some

learnable R_ϕ which can approximate $\exp[\widehat{f}_P \Delta t]$, and indirectly learnable v which can approximate $\exp[\widehat{f}_N \Delta t] \Psi_t$, to an arbitrary margin ϵ . It directly follows that the Fourier Kernel Integral Operator can approximate the first-order SSFM time step update to an arbitrary degree of accuracy which then implies that the FNO as a whole can approximate the solution to the GPE to an arbitrary degree of accuracy, with the number of Kernel Integral layers in the FNO dictating the time discretization Δt of the approximation.

Appendix C Model Hyperparameters

Table C1: Model Hyperparameters

Hyperparameter	Value
Learning Rate	0.001
Batch Size	32
Epochs	560
Optimizer	Adam
Loss Function	H1Loss
Fourier Layers	4
Height Modes	128
Width Modes	128
Hidden Dimension	64
Domain Padding	0.125
Domain Padding Mode	symmetric
Activation Function	GeLU

The training loss as a function of the epochs is shown in Figure C1. The model was trained on an NVIDIA RTX 4090 GPU with an Intel i9 13900KF and each epoch took on average 15.1s to train.

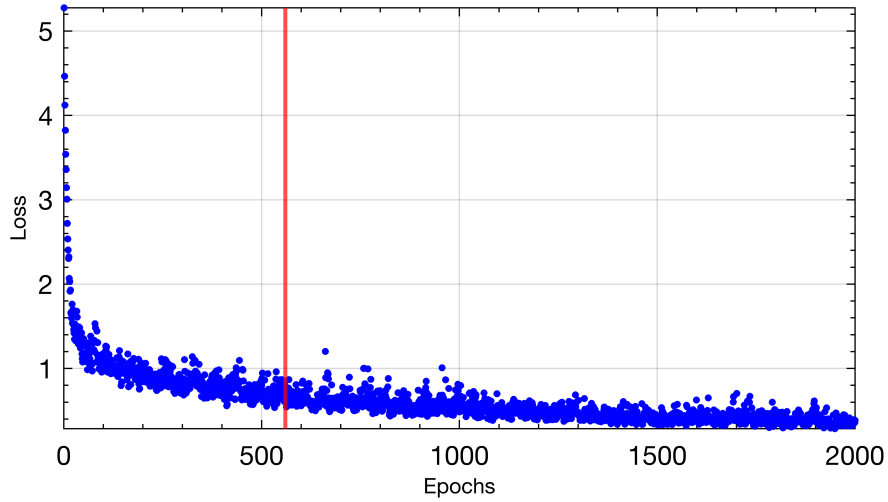


Figure C1: Training Loss per Epoch for Fourier Neural Operator. The red line, representing epoch number 560, is selected for the purpose of prediction analysis.

Appendix D Extended Results

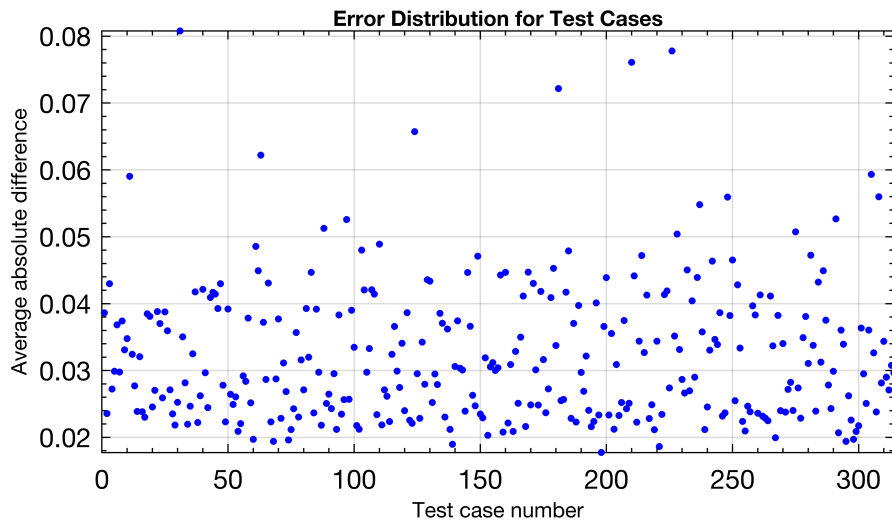


Figure D1: Average absolute difference for each test case normalised by grid size.

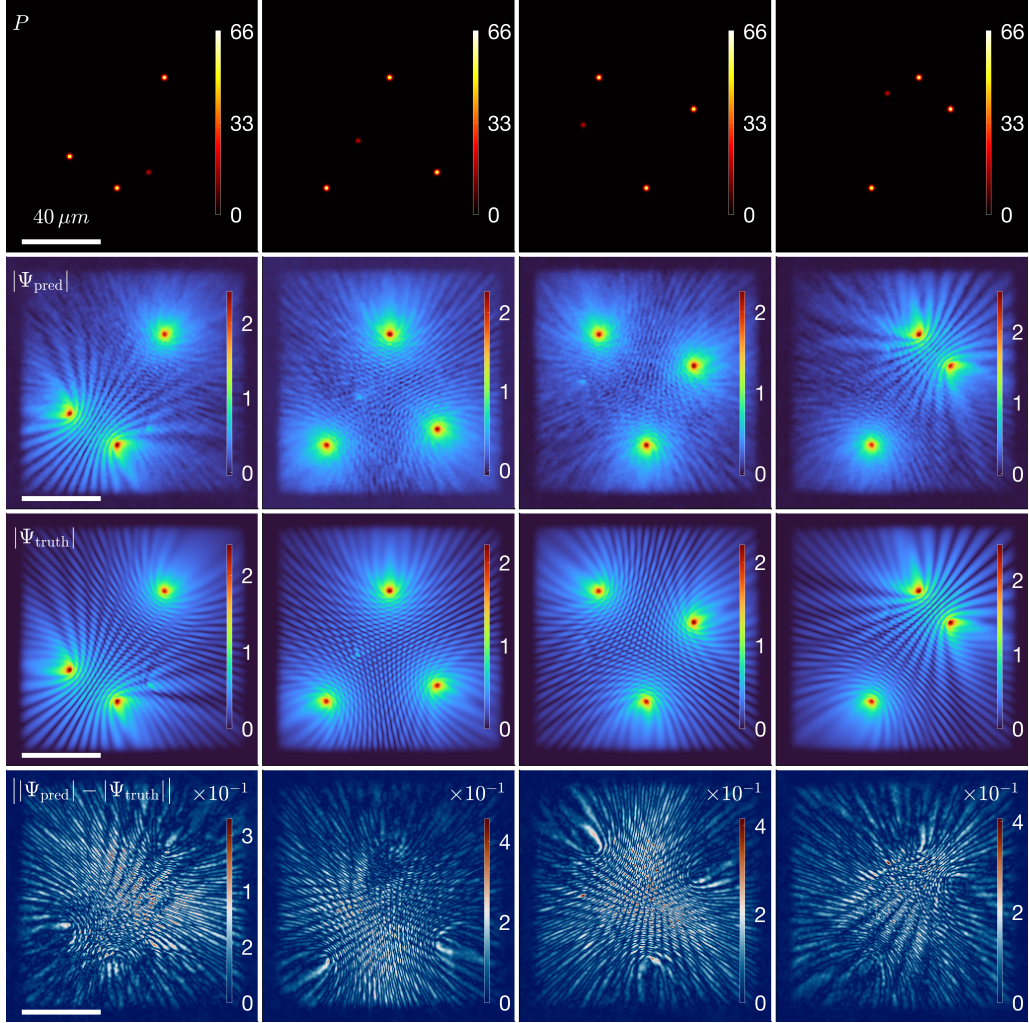


Figure D2: Comparison of the prediction from the Fourier Neural Operator approach. The columns, left to right, show different pump configurations (test case number 69, 103, 148, and 258 from Figure D1). The rows, from top to bottom, are pump profiles P , predictions $|\Psi_{\text{pred}}|$, ground truths $|\Psi_{\text{truth}}|$, and errors $||\Psi_{\text{pred}}| - |\Psi_{\text{truth}}||$.