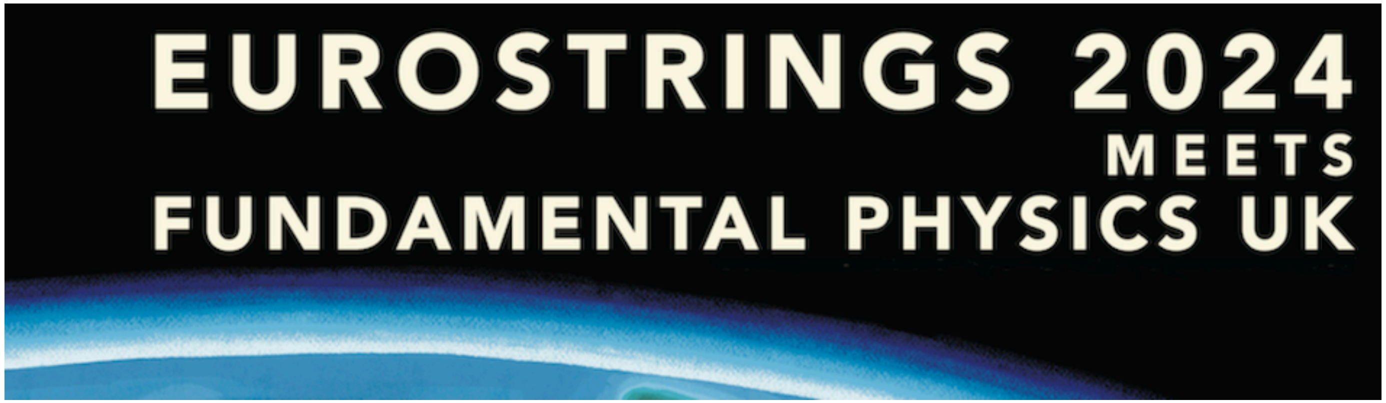
Non-perturbative Quantum Physics and Deep Learning

Costis Papageorgakis and Vasilis Niarchos

Queen Mary University of London and University of Crete



Eurostrings 2024, 2-6 September, 2024

Machine Learning (ML) & AI in Science: *a teaser*

https://github.com/SakanaAI/AI-Scientist

Lu-Lu-Lange-Foerster-Clune-Ha

arXiv:2408.06292v2 [cs.AI] 15 Aug 2024

Abstract

``One of the grand challenges of **artificial general intelligence** is developing agents capable of conducting scientific research and discovering new knowledge. While frontier models have already been used as aides to human scientists, e.g. for brainstorming ideas, writing code, or prediction tasks, they still conduct only a small part of the scientific process. This paper presents the first comprehensive framework for fully ***automatic scientific discovery***, enabling frontier **large language models (LLMs)** to perform research independently and communicate their findings. We introduce The AI Scientist, which **generates novel research ideas**, **writes code**, **executes experiments**, visualizes results, describes its findings by **writing a full scientific paper**, and then **runs a simulated review process for evaluation**…"

- Other scientific fields have already incorporated aggressively many of the recent developments in ML/AI

- Our field is not naturally data-driven, it requires rigor & interpretability, but ML methods are usually stochastic & error-prone.
  There are, however, examples of rigor & interpretability in ML/AI
  (see [Gukov-Halverson-Ruehle 2402.13321] for a recent perspective article)

- And ML/AI applications to HEP-th are slowly growing:
  - **CY metrics**: Anderson-et al (2012.04656), Douglas-et al (2012.04797), Jejjala-et al (2012.15821)
  - **Conformal bootstrap**: VN-CP-et al (2108.08859, 2108.09330, 2209.02801, 2306.15730)
  - **Amplitude bootstrap**: Cai-…-Dixon (2405.06107 [cs.LG])
  - And much more… see eg upcoming talks and Ruehle review Phys. Rept. 839 (2020) 1–117

- HEP-th is faced with many challenging conceptual & computational problems

- Applying unorthodox approaches (like ML/AI) to such problems will require:

  a) Clever reformulations of traditional problems in HEP-th

  b) Clever ways to encode complex (typically, infinite-dimensional) problems - a CS challenge

  c) Suitable incorporation of multiple domain-specific analytical & numerical results

- **Exciting, novel research avenues!**

QFT as a specific domain of exploration

The **non-perturbative** dynamics of **Quantum Field Theories** (QFTs) is a fundamental, still largely open problem, which is central in our field

Using ML/AI we can try to:

1) Solve `old problems' in new ways (eg ML for efficient Lattice FT)

2) Perform efficient searches of solutions satisfying known constraints (eg some PDE, or some bootstrap constraints)

3) Identify structures in tractable corners and generalize to currently inaccessible regimes [generative AI techniques can be useful in this context]

In this talk

I want to illustrate 2-3 in the context of a specific problem in Quantum Mechanics

a) I will set up the problem

b) I will present a more-or-less traditional implementation of Neural-Network methods to this problem along the lines of 2

c) I will discuss a less traditional approach using **Neural Operators** along the lines of 3 [VN-CP, 2404.14551]

Scattering amplitude phases

Consider elastic 2→2 scattering in Quantum Mechanics

The differential cross-section can be deduced from the scattering amplitude

$$\frac{d\sigma}{d\Omega} = |f(\theta)|^2$$

The scattering amplitude $f(z) = b(z)e^{i\phi(z)}$ ($z := \cos\theta$), is a complex function.

Question: if we know the modulus $b(z)$ (at fixed energy) can we reconstruct the phase $\phi(z)$?

Given an amplitude with a **finite** partial wave expansion

$$f(z) = \frac{1}{k} \sum_{\ell=0}^{L} (2\ell + 1) \sin \delta_\ell \, e^{i\delta_\ell} \, P_\ell(z)$$

in terms of a finite number of phase-shifts δ_ℓ, one can easily express both $b(z)$ and $\phi(z)$ in terms of δ_ℓ and attempt to reconstruct them.

The problem is much harder for generic amplitudes that have an **infinite** partial wave expansion.

This problem was studied in the 60s and 70s by various people (Martin, Atkinson,…) and the standard approach goes as follows.

A simple argument shows that unitarity implies the following integral relation between the modulus $B(z) := k\, b(z)$ and the phase $\phi(z)$

$$\sin \phi(z) = \int_{-1}^{1} dz_1 \int_{0}^{2\pi} d\phi_1 \frac{B(z_1)B(z_2)}{4\pi B(z)} \cos\left[\phi(z_1) - \phi(z_2)\right]$$

where $z_2(z, z_1, \phi_1) := z z_1 + \sqrt{1-z^2}\sqrt{1-z_1^2}\cos\phi_1$

The analysis of this equation yielded some interesting results:

a) Setting $z = 1$ gives the dual bound

$$\int_{-1}^{1} dz_1 \frac{B(z_1)^2}{2B(1)} \leq 1$$

b) Additional bounds on existence and uniqueness can be obtained by defining the function

$$K(z) := \int_{-1}^{1} dz_1 \int_{0}^{2\pi} d\phi_1 \frac{B(z_1)B(z_2)}{4\pi B(z)}$$

and the `Martin parameter'

$$\sin\mu := \max_{-1\leq z\leq 1} K(z)$$

One can prove that given a modulus $B(z)$ a phase $\phi(z)$ always exists when $\sin\mu \leq 1$.

For $\sin\mu > 1$ solutions may or may not exist, and when they exist they can be unique or doubly-ambiguous. It is unclear what happens in general.

A more modern ML approach towards this problem would attack the unitarity equation

$$\sin\phi(z) = \int_{-1}^{1} dz_1 \int_{0}^{2\pi} d\phi_1 \frac{B(z_1)B(z_2)}{4\pi B(z)} \cos\left[\phi(z_1) - \phi(z_2)\right]$$

in the following manner. Given a specific modulus $B(z)$

i) Model $\phi(z)$ by a Neural Network (NN): $\phi_{\mathbf{\Theta}}(z)$

ii) Formulate a loss function, e.g.

$$\mathscr{L}(\mathbf{\Theta}) := N_c^{-1} \sum_i \left(\sin\phi_{\mathbf{\Theta}}(z_i) - \frac{1}{4\pi B(z_i)} \int_{-1}^{1} dz_1 \int_{0}^{2\pi} d\phi_1 B(z_1)B(z_{2i})\cos(\phi_{\mathbf{\Theta}}(z_1) - \phi_{\mathbf{\Theta}}(z_{2i})) \right)^2$$

iii) minimize $\mathscr{L}(\mathbf{\Theta})$ wrt the NN parameters $\mathbf{\Theta}$

Finding the optimum NN parameters $\boldsymbol{\Theta}$ is called training and involves solving a generally hard non-convex optimization problem.

[Chen…, Lagaris,… 90's]

[Raissi et al 2017]

This approach is called a **PINN** (Physics-Informed Neural Network) and is a standard NN method for solving PDEs etc.

It was recently used by [Dersy-Schwartz-Zhiboedov, 2308.09451] to solve the above unitarity equation producing notable results that include a new ambiguous solution with the lowest known Martin parameter $\sin\mu \simeq 1.67$ (improving the relevant bound for the first time in 50 yrs!)

Note: this approach requires retraining from scratch for each new input $B(z)$, which makes scanning over functions cumbersome.

A different approach

[VN-CP, 2404.14551]

Assume that the unitarity equation is unknown.

The only tractable information that we will allow ourselves to have access to comes from finite partial wave expansions.

Question: Is it still possible to learn the relation between $B(z)$ and $\phi(z)$, or when a $\phi(z)$ does not exist?

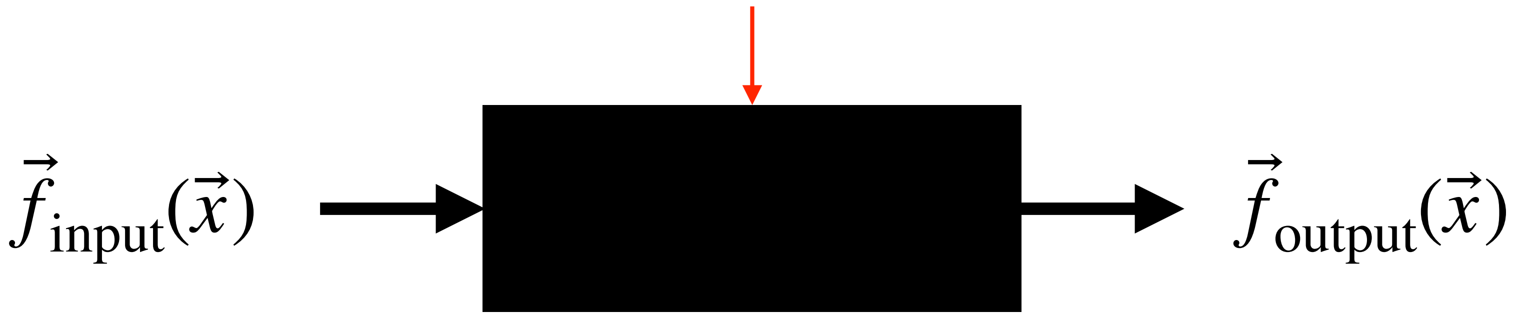
Note: Learning maps between infinite-dimensional spaces of functions is also interesting from the continuum-QFT perspective, because QFTs are naturally structures operating on such spaces.

More involved NN architectures (called Neural Operators (NOs)) can be used for this purpose

[Chen-Chen 90's]

[Li et al 2021 **Fourier NOs**]

integral transforms + linear operations + non-linear activations



[Supervised learning]

- In this case one trains by exposing the NO to many pairs $(\vec{f}_{\text{input}}, \vec{f}_{\text{output}})$
- Train only once
- `Zero-shot super-resolution'
- **PINN + NO = PINOs** [Li et al 2021]

Unique phases: Training on 300K valid input-output samples

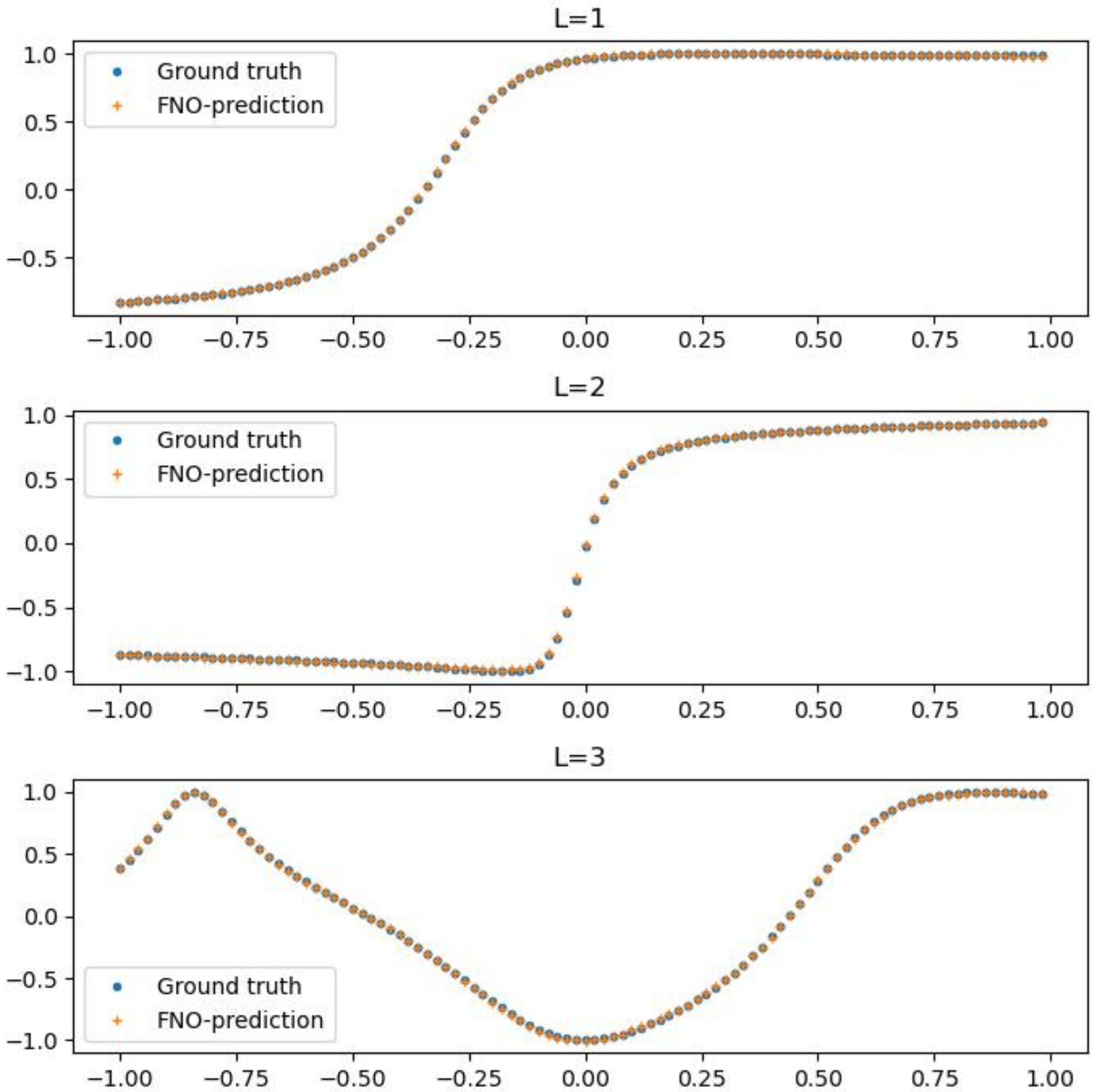
- We trained an FNO on 300K $(B(z), \phi(z))$ random samples consisting of 100K L=1 amplitudes, 100K L=2 amplitudes, 100K L=3 amplitudes

$$F(z) = B(z)e^{i\phi(z)} = \sum_{\ell=0}^{L} (2\ell + 1)\sin \delta_\ell e^{i\delta_\ell} P_\ell(z)$$

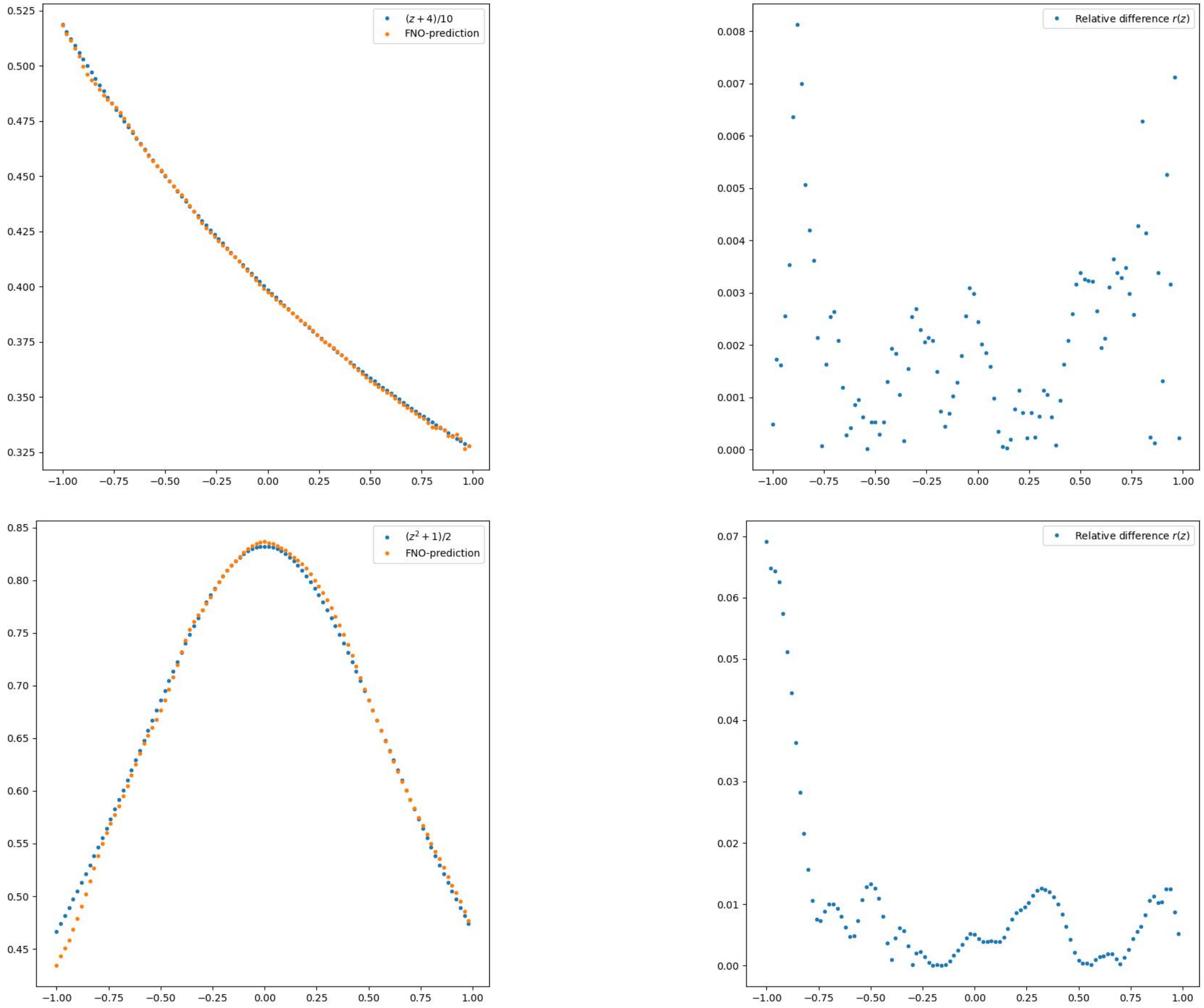
- $z \in [-1,1]$ was discretized on a uniform grid with 100 points.

  We can train with this resolution and evaluate on much higher resolution!

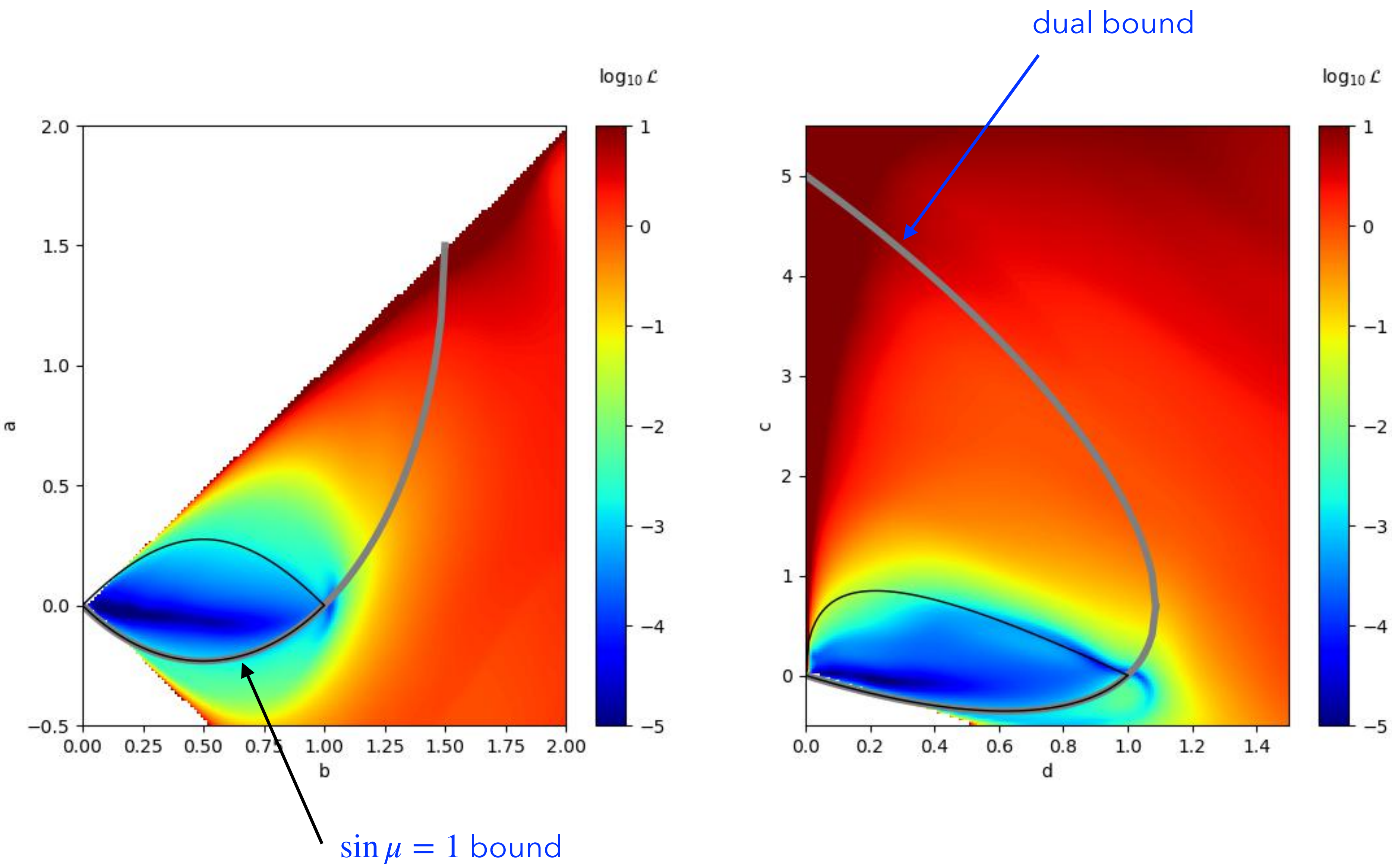
Performance within the training dataset: plots of $\sin(\phi(z))$



Performance on **unseen** linear and quadratic moduli $B(z)$ with infinite partial wave expansions (ground-truth computed by solving independently the unitarity equation)



Heatmaps of the $\log_{10}$ of the unitarity loss for linear $B(z) = az + b$ and quadratic $B(z) = cz^2 + d$ moduli. Similar results obtained with PINNs in [Dersy et al, 2308.09451]



This is promising (the NO has indeed learnt non-trivial features of the unitarity constraint) but:

- Remember that we cannot use the unitarity equation to check. So, we do not really know when a prediction is good or not, or when a phase is not supposed to exist.

- The above NO has not learnt everything. Testing on higher-L finite partial wave expansions is very poor.

We need something more…

We need to show the NO what it means to fail…

Unique phases: Learning false predictions

- We trained on 400K samples with the following composition: 225K true samples with L=1,2,3, and 175K false samples

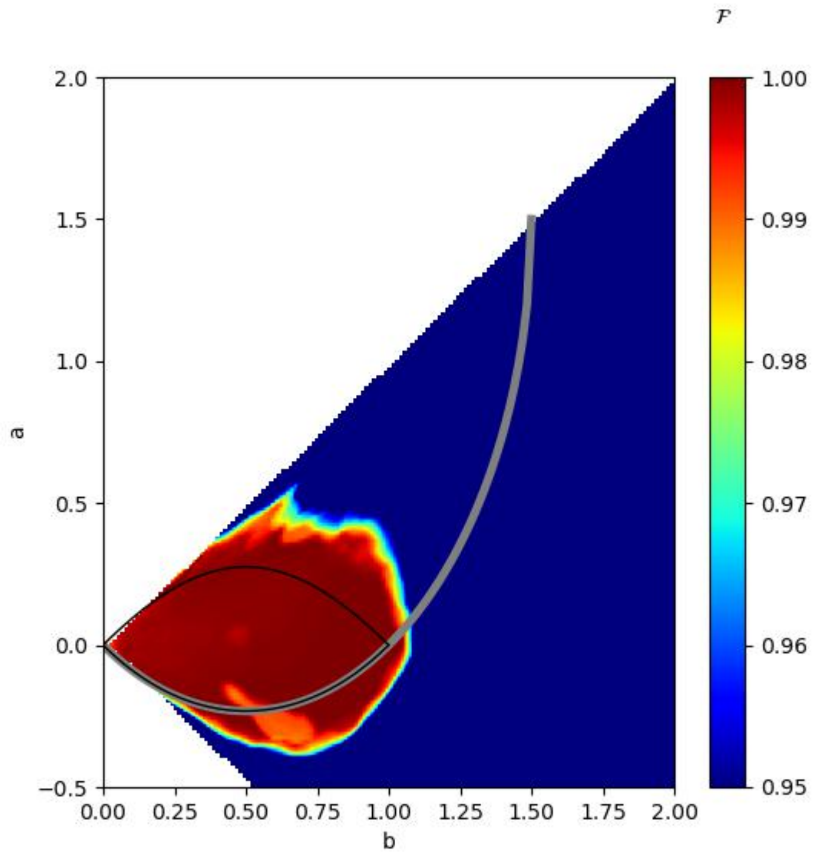
- The output is now a $\phi(z)$ prediction + an index $\mathscr{F}$ (fidelity index). During training, $\mathscr{F} = 0$ for false samples and $\mathscr{F} = 1$ for true samples.

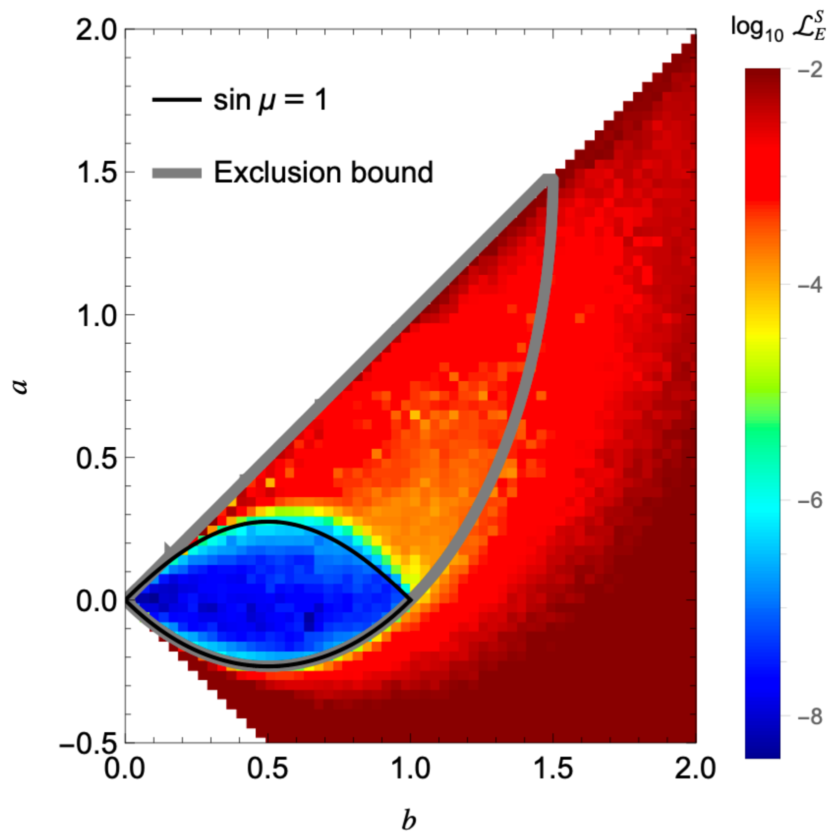
- We found that we get much better reliability from $\mathscr{F}$ when we average over several (independently-trained) NOs (we used a group of 56 NOs).

- Within the training-test dataset the average fidelity index $\overline{\mathscr{F}}$ was 98.6% correct when it predicted a good solution with the criterion $|\overline{\mathscr{F}} - 1| < 0.01$.

- We also found that $\overline{\mathscr{F}}$ correlated non-trivially in general with the unitarity loss.

Heatmaps of the mean fidelity index $\overline{\mathscr{F}}$ for the linear moduli $B(z)$ (right [Dersy et al, '23]).
Notice the sharp separation between acceptance and rejection.

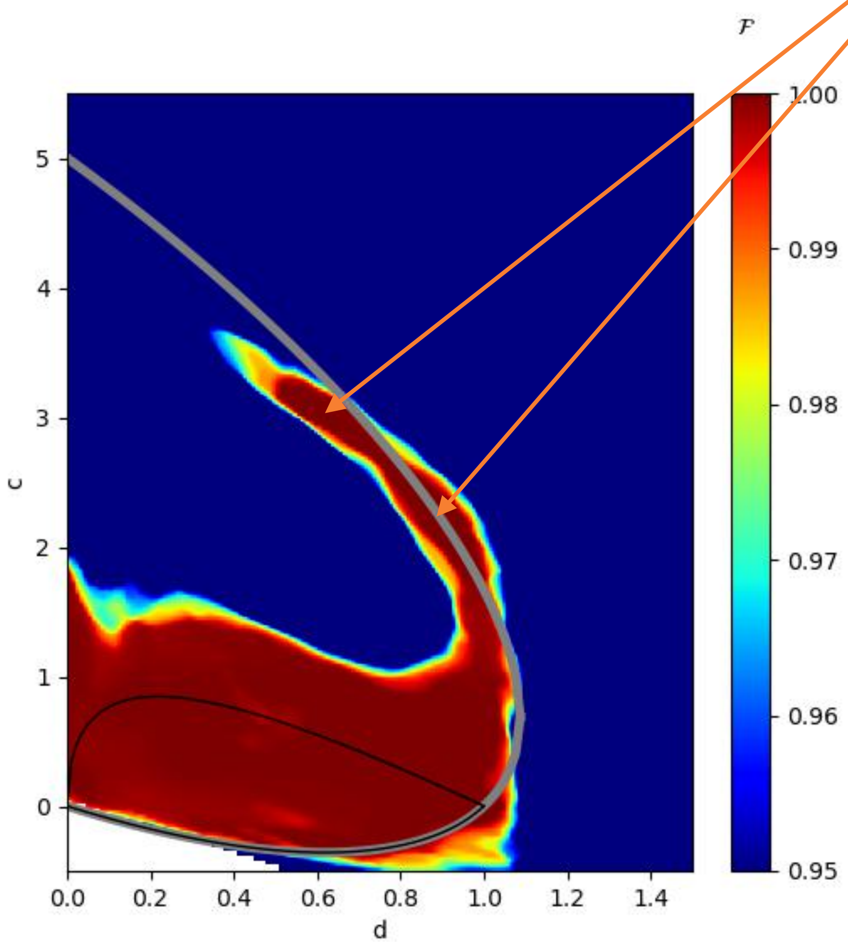


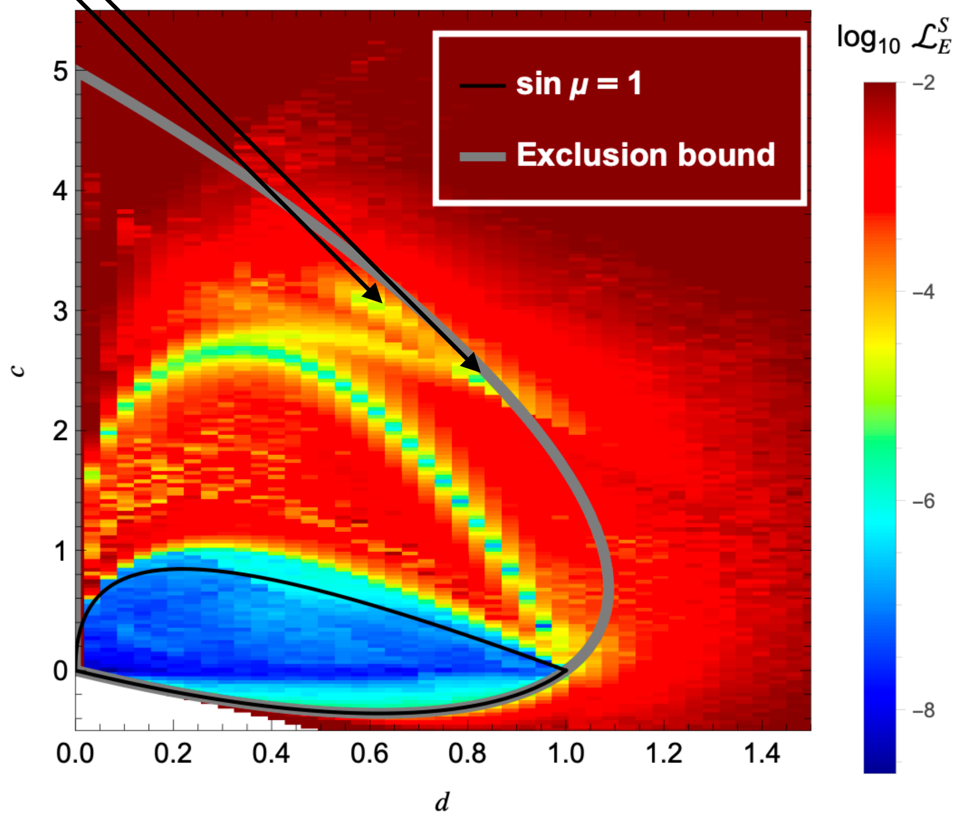


Heatmaps of the mean fidelity index $\overline{\mathscr{F}}$ for the quadratic moduli $B(z)$

(right [Dersy et al, '23])

there are two L=2 finite partial-wave solutions here!





Outlook

QFTs are a particularly interesting context characterized by challenging high-dimensional (∞-dim) structures

Question: Can we bring together the progress of recent years (in SUSY QFTs, AdS/CFT, numerical/analytic bootstrap…) to devise new strategies and to extend (generalize) the knowledge of tractable corners into generic inaccessible regimes?

Question: Can we use ML/AI to improve both numerical and analytical results? How will the interplay between numerics & analytics evolve?

Many interesting contexts and QFTs…

Ambiguous phases: NOs on the double cover

- To test the potential of reproducing ambiguous solutions, we also trained on a 100K-sample dataset composed of 30K unique L=3 solutions and 70K known classified L=2, 3 ambiguous solutions.

- In this case, the output function consisted of two $\phi(z)$ solutions.

- Training and generalization were much more subtle and difficult in this situation for obvious reasons.

Reproducing an infinite partial-wave solution by [Atkinson et al, 1978]. One NO predicts two solutions, another predicts one of the two solutions.