Using physics-informed quantum machine learning to solve nonlinear differential equations in latent spaces without costly grid evaluations

Annie E. Paine, Vincent E. Elfving, Oleksandr Kyriienko

We propose a physics-informed quantum algorithm to solve nonlinear and multidimensional differential equations (DEs) in a quantum latent space[1]. We suggest a strategy for building quantum models as state overlaps, where exponentially large sets of independent basis functions are used for implicitly representing solutions. By measuring the overlaps between states which are representations of DE terms, we construct a loss that does not require independent sequential function evaluations on grid points. In this sense, the solver evaluates the loss in an intrinsically parallel way, utilizing a global type of the model. A visualization of this algorithm is shown in Fig. 1. In particular, we present toolboxes for Chebyshev and Fourier basis sets, developing tools for automatic differentiation and multiplication, implementing nonlinearity, and describing multivariate extensions. The approach is compatible with, and tested on, a range of problems including linear, nonlinear and multidimensional differential equations.

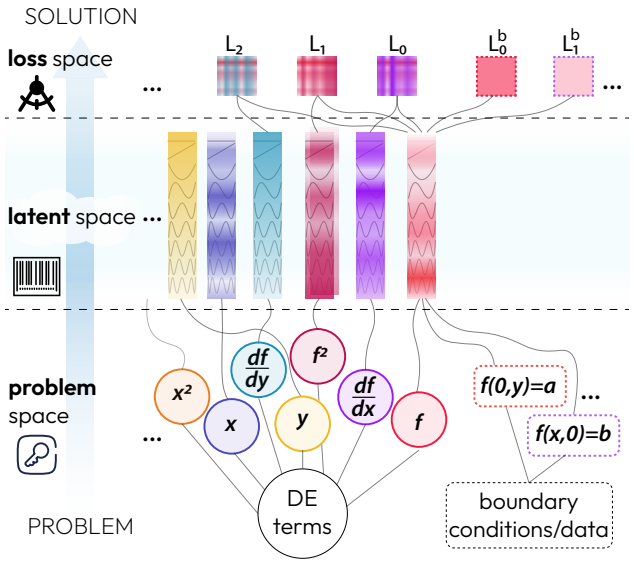

Figure 1: A conceptual visualization of the proposed physics-informed approach to solving nonlinear differential equations (bottom-up description). We start from the problem space, where differential equation (DE) terms, their relations, and the boundary terms are provided. Next, these terms are elevated to the latent space via feature map that prepares corresponding quantum states, with amplitudes associated to components of independent basis functions. These states are then compared via pairwise overlaps, forming corresponding distances in the loss space, and boundary terms are evaluated explicitly via basis mappings. The total loss can be minimized to learn a model that provides a solution to the problem.

Fast, flexible and efficient solvers of differential equations underpin the area of scientific computing. Applied to computationally hard problems in physics, fluid dynamics, geoscience, chemistry, and finance, differential equation solvers consume a significant portion of the world's high-performance computing resources. The quest for developing and improving DE solvers remains open. One of the problems for current DE solvers is known as the curse of dimensionality. As dimension scales so does the number of grid points, often exponentially, leading to poor scaling for the many solvers utilising grid points.

One approach to solving DEs is physics-informed neural networks (PINNs) — deep neural networks that include differential constraints introduced with automatic differentiation. PINN-based solvers have been

[1] A. E. Paine, V. E. Elfving, and O. Kyriienko, **Physics-informed quantum machine learning: Solving nonlinear differential equations in latent spaces without costly grid evaluations**, Aug. 2023. arXiv:2308.01827 [quant-ph]

applied to many use-cases. Although still remaining relatively niche, their flexibility allows them to compete with FEM/SEM solvers in cases involving data-based constraints, optimization loops, and can excel when dealing with stiff and multidimensional problems. However, PINN-based DE solvers require a costly training process that leads to large overheads as compared to conventional methods such as FEM, often needing $> 10,000$ epochs with many function and gradient evaluations over a training grid. While the required function evaluations on the grid may be parallelized, in practice the cost is often not considered worth it for many industrial applications, despite the potential of the ML setting and what the flexibility can offer in terms of utility in use cases. Reducing the grid evaluation cost for each epoch is crucial for making PINNs industrially practical.

We propose a physics-informed quantum algorithm to solve DEs which utilises quantum properties to avoid independent grid point evaluations. The goal of the protocol is to solve a system of differential equations with specified initial or boundary conditions. For example, simplified for clarity but generalisable,

$$\mathrm{DE}(x, f, df/dx) = \sum_{k=1}^{T} \mathrm{DE}^{[k]}(x, f, df/dx) = 0, \quad f(x_0) = f_0, \tag{1}$$

where $\{\mathrm{DE}^{[k]}(x, f, df/dx)\}_k$ denote the separate terms of the differential equation formed of products of f, and df/dx as well as arbitrary functions of x. We want to obtain a solution to Eq. (1) as function $f(x)$. For this, we use a quantum model $f_{\boldsymbol{\theta}}(x)$ parametrized by a vector of weights $\boldsymbol{\theta}$, and reformulate the problem as learning an optimal model $f_{\theta^*}(x)$ to represent $f(x)$.

We instantiate a quantum model in the general form

$$f_\theta(x) = \langle x | f_\theta \rangle\!\rangle, \tag{2}$$

represented by the overlap between a quantum state $|x\rangle$ and a classical sum of quantum states (or single quantum state) $|f_\theta\rangle\!\rangle$. $|f_\theta\rangle\!\rangle$ is prepared via variational circuits $\hat{\mathcal{V}}(\theta)$ such that the represented function $f_\theta(x)$ is altered by varying θ. The state $|x\rangle$ encodes the variable x as $|x\rangle = \hat{\mathcal{U}}(x)|0\rangle$ using feature map $\hat{\mathcal{U}}(x)$. This encoding provides an implicit fitting basis formed of the amplitudes of $|x\rangle$ considered as functions of x. Within the proposed quantum model definition, we posit that each term $\mathrm{DE}^{[k]}$ can be expressed as

$$\mathrm{DE}^{[k]}(x, f_\theta, df_\theta/dx) = \langle \tilde{x} | \mathrm{DE}^k_\theta \rangle\!\rangle, \tag{3}$$

where $|\tilde{x}\rangle$ is a quantum state corresponding to $|x\rangle$ and is the same for all DE terms.

The construction of the terms $|\mathrm{DE}^k_\theta\rangle\!\rangle$ depends on the encoding used for $|x\rangle$ and the resulting implicit basis. In general though we want to be able to represent our trial function $f_\theta(x)$, its derivatives, arbitrary functions $g(x)$ and multiplications of these terms. For $f_\theta(x)$ the representative state is simply our variational state $|f_\theta\rangle\!\rangle$. For the derivatives, for many bases (such as Chebyshev and Fourier) an x-independent operator $\check{G}$ can be found such that $\hat{\mathcal{U}}'(x) = \check{G}\hat{\mathcal{U}}(x)$. For arbitrary functions $g(x)$ either transformations are utilised or a pre-training step occurs. For multiplication we consider the basis formed by the set of products of the implicit basis formed by $|x\rangle$ and how to transform to this basis from the implicit basis. For Fourier and Chebyshev bases the resulting transformation is built from adder and subtractor circuits.

Substituting the expression in Eq. (3) into Eq. (1) we get

$$\mathrm{DE}(x, f_\theta, df_\theta/dx) = \langle \tilde{x} | \sum_{k=1}^{T} |\mathrm{DE}^k_\theta\rangle\!\rangle = 0, \tag{4}$$

for all x. We additionally restrict the state $|\tilde{x}\rangle$. Each amplitude of $|\tilde{x}\rangle$ can be considered as a function of x. We require that the set of these functions form an independent set. With this restriction the differential equation is satisfied if and only if

$$\sum_{k=1}^{T} |\mathrm{DE}^k_\theta\rangle\!\rangle = \mathbf{0}, \tag{5}$$

where the used notation implies that every element of the differential constraint vector has zero value. If the restriction on $|\tilde{x}\rangle$ is not in place then $\sum_{k=1}^{T} |\mathrm{DE}^k_\theta\rangle\!\rangle = \mathbf{0}$ is sufficient but not a necessary condition.

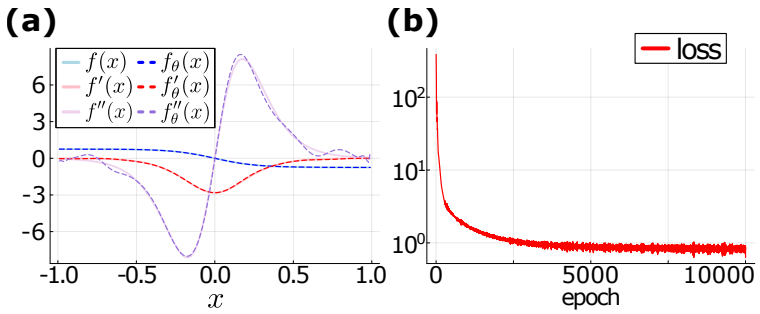

Figure 2: Results of solving stationary burgers equation $f(x)\frac{\partial f(x)}{\partial x} = \nu\frac{\partial^2 f(x)}{\partial x^2}$ with $\nu = 0.1$ and boundary values $f(-1) = -f(1) = 0.75\tanh(3.75) = f_{-1}$. **(a)** Plot of target function and first and second order derivatives (solid lines) versus resulting function and derivatives (dash lines). **(b)** Value of loss function over epoch iteration.

For solving (5) we consider a variational approach based on minimizing the loss function

$$\mathcal{L}_{\mathrm{DE}}(\theta) = \sum_{k=1}^{T} \langle\langle \mathrm{DE}_\theta^k | \sum_{m=1}^{T} | \mathrm{DE}_\theta^m \rangle\rangle \tag{6}$$

To evaluate the loss function $\mathcal{L}_{\mathrm{DE}}(\theta)$ the overlaps between prepared quantum states are measured and then post-processed classically to account for weights. Similarly, we need to evaluate derivative-based contributions. Once states are prepared, the overlaps are estimated using known methods, for instance the Hadamard test.

Next, we include the initial/boundary value by utilizing the access to our model via quantum feature map circuits. Namely, the function can be evaluated at some initial point x_0 as $f_\theta(x_0) = \langle 0|\hat{\mathcal{U}}(x_0)|f_\theta\rangle\rangle$, rather than globally. This can then be used in a loss term to minimise difference between current initial/boundary value and desired. For example

$$\mathcal{L}_{\mathrm{init}}(\theta) = \eta \left\{ f_\theta(x_0) - f_0 \right\}^2, \tag{7}$$

where η controls the importance of the initial value contribution. We note that separate grid evaluations are required at this stage though in future work we hope to reduce/remove this need.

The total loss is then

$$\mathcal{L}(\theta) = \mathcal{L}_{\mathrm{DE}}(\theta) + \mathcal{L}_{\mathrm{init}}(\theta), \tag{8}$$

which is to be minimized either by non-convex optimization methods, finding $\theta^* = \mathrm{argmin}_\theta \mathcal{L}(\theta)$, or applying other iterative methods to recover the optimal state $|f_{\theta^*}\rangle\rangle$.

We demonstrate by classical simulation the use of this algorithm on the stationary Burgers equation

$$f(x)\frac{\partial f(x)}{\partial x} = \nu\frac{\partial^2 f(x)}{\partial x^2}, \tag{9}$$

with $\nu = 0.1$ and boundary values $f(-1) = -f(1) = 0.75\tanh(3.75) = f_{-1}$. This is a second order, nonlinear DE and thus requires multiplication and derivatives. Using the five qubit Chebyshev basis, a depth ten hardware efficient ansatz and adam optimiser a good fit is able to be achieved in magnitude of 1000 epochs. The results of this is shown in Fig. 2. This algorithm has additionally been simulated for various other problems including multidimensional problems.

In summary, we have proposed a physics-informed quantum machine learning algorithm for solving differential equations able to avoid separate grid point evaluations during DE training. By avoiding grid point evaluations, negative scaling with grid size (such as the curse of dimensionality) is mitigated or reduced, improving on some of the limiting factors of PINNs. This is at the cost of overlap measurement and inclusion of some more expensive circuits such as multiplication. Particular future directions for this work include avoiding function evaluations when considering initial/boundary values, developing further bases to use, and scaling up to consider larger scale problems.