DLSC Project B

Operator Networks vs Parametric Approach

Yves Bicker

University of Zürich, Faculty of Science
Zürich, Switzerland

Contents

1 Introduction and problem description

In this project, the aim is to investigate and compare two strategies for forecasting the future behavior of the system at a specified time $t = T$. The first strategy employs a parametric approach, while the second strategy involves an operator-based approach, with particular emphasis on the Fourier Neural Operator (FNO). The comparison is conducted in two stages: initially, the applicability of both approaches is tested on the heat equation, followed by the wave equation.

Heat equation

The two-dimensional parametric heat equation is defined over the following domain with the given I.C.:

$$u_t = \Delta, \quad t \in [0, T], \quad (x_1, x_2) \in [-1, 1]^2, \quad (\mu_1, ..., \mu_d) \in [-1, 1]^d$$

$$u(t = 0, x_1, x_2, \mu) = -\frac{1}{d} \sum_{m=1}^{d} \frac{\mu_m}{\sqrt{m}} \sin(m\pi x_1) \sin(m\pi x_2) \tag{1}$$

and zero Dirichlet boundary conditions. The corresponding analytical solution is the following.

$$u(t, x_1, x_2, \mu_m) = -\frac{1}{d} \sum_{m=1}^{d} e^{-(m\pi)^2 t} \frac{\mu_m}{\sqrt{m}} \sin(m\pi x_1) \sin(m\pi x_2) \tag{2}$$

2 Surrogate model for the 2D heat equation

Parametric Neural Network Approach

The parametric approach predicts the values of $g(x_1, x_2, \mu_m) = u(T, x_1, x_2, \mu_m)$, using a Multi-Layer Perceptron (MLP). This approach tries to learn the underlying relationship between the input positions and parameters, and the corresponding output value. For the heat equation we study the system at T=0.1.

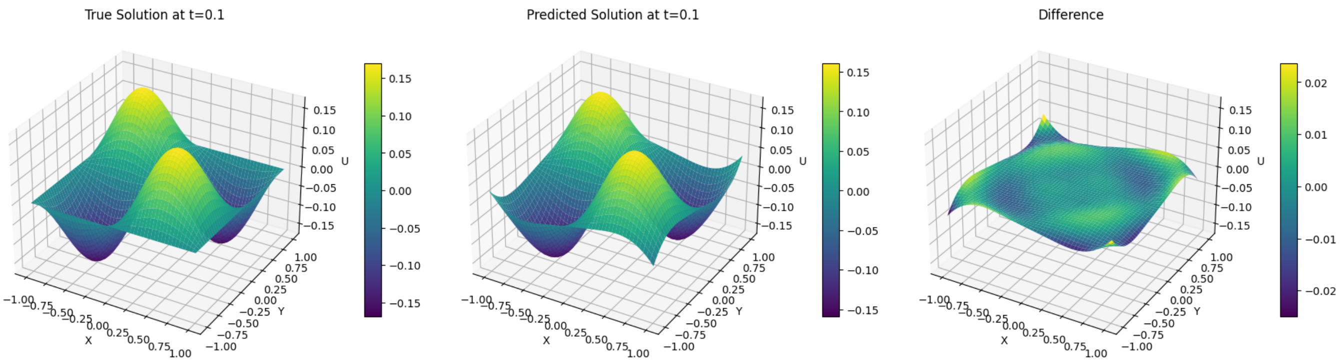

Figure 1: *Example output Parametric model with parameter $d = 2$*

The standard MLP uses an input size $2 + d$, output size 1 and N_{layers} fully connected hidden layers, each consisting of N_{neurons} neurons. S_{NN} defines the total number of learnable parameters, including weights and biases.

$$S_{\text{NN}} = (2 + d + 1)N_{\text{neurons}} + (N_{\text{layers}} - 1)(N_{\text{neurons}} + 1)N_{\text{neurons}} + (N_{\text{neurons}} + 1) \tag{3}$$

Fourier Neural Operator Approach

The FNO operator-based approach, involves a function-to-function mapping, where the initial condition function $u_0(x_1, x_2, \mu_m)$ specified across the entire domain, is mapped to the analytical solution $\mathcal{G}(u_0)(x_1, x_2) = u(T, x_1, x_2)$. This offers an alternative way to solve the parametric heat equation by approximating the operator.

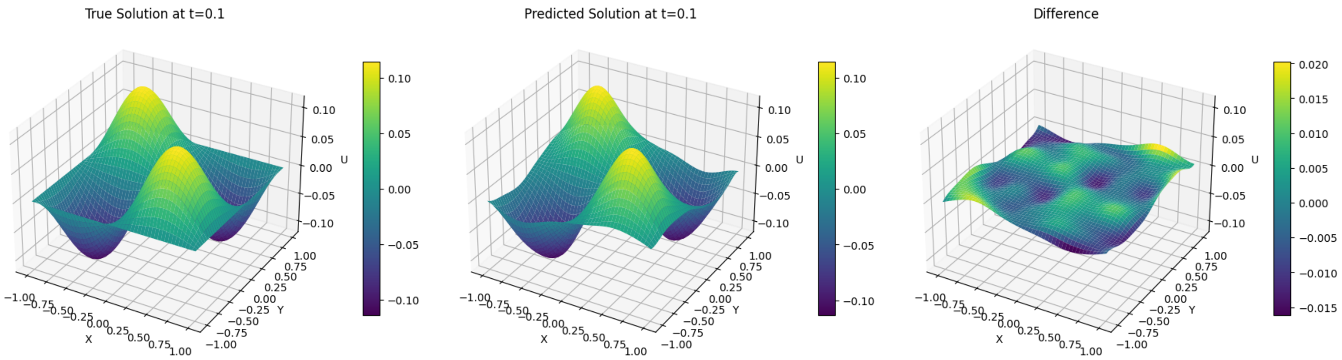

Figure 2: *Example output FNO model with parameter $d = 2$*

The FNO architecture uses a representation width w, number of modes m, N_{layers} Fourier layers and one fully connected step-down layer of size 128.

$$S_{\text{FNO}} = (3 + 1)w + N_{\text{layers}}((w + 1)w + 2w^2m^2) + (w + 1)(128) + 128. \tag{4}$$

3 Experiment and Results Heat Equation

Comparing two different network types introduces inherent complexities due to the varying structure and parameter size of each network. In order to conduct a fair comparison, efforts were made to match the parameter sizes as closely as possible. Both models were trained using the same quantity of data generated for the initial condition and the solution on a $50x50$ grid. This resulted in a sample size of 2500 elements. A total of 1000 samples were used for training, with a division of 80% for training and 10% each for validation and testing.

The parameters for each model were adjusted to achieve a similar size, with $S_{\text{NN}} = 8123$ for the Neural Network (NN) model and $S_{\text{FNO}} = 8197$ for the Fourier Neural Operator (FNO) model. Both models underwent training for a fixed number of 100 epochs using a Tesla T4 from Google Colab. The training process of the heat equation is illustrated for the parameter sizes $d \in \{1, 2, 3\}$.

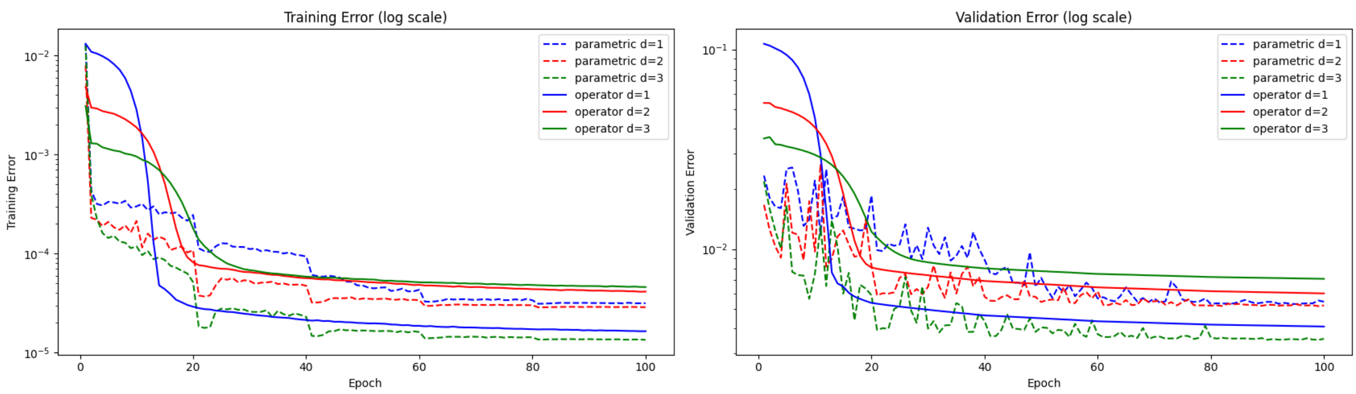

Figure 3: *Training and validation losses for the parametric NN and FNO models*

The computational resources required for model training were found to be significantly higher for the parametric approach in comparison to the Fourier Neural Operator (FNO) approach. To quantify model performance, a subset constituting 10% of the entire dataset, which corresponds to 100 samples, was designated as a hold-out test set. Errors for both model types were evaluated over a 50x50 grid, on a per sample basis. Given the significant variability in the scale at $T = 0.1$ for different dimensions (d) in the heat equation, relative log error was chosen as the error metric, facilitating a more equitable comparison between the models. The resultant test errors for the hold-out set are tabulated below

Approach	**Relative Log-Test-Error**		
Heat	d=1	d=2	d=3
FNO	-6.52	-4.26	-3.33
NN	-5.91	-4.63	-4.58

Table 1: *Relative Log-Test-Error for the heat equation*

4 Experiment and Results Wave Equation

The wave equation was chosen since it has a more complex evolution across the time scale, which should make it more difficult to capture the underlying behaviour for prediction.

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \tag{5}$$

The problem description is basically the same regarding the domain, parameter values ant the zero Dirichlet boundary condition. The only thing that changes is the initial analytical solution, without a decay term in time but a cosine instead.

$$u(t, x_1, x_2, \mu_m) = -\frac{1}{d} \sum_{m=1}^{d} \frac{\mu_m}{\sqrt{m}} \sin(m\pi x_1) \sin(m\pi x_2) \cos(m\pi t) \tag{6}$$

Training time

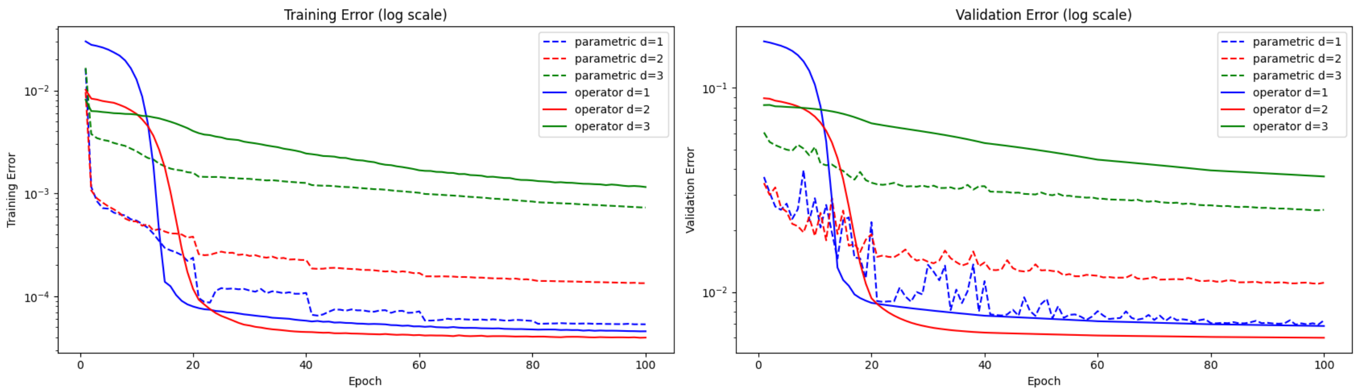

Figure 4: *Training and validation losses for the Wave Equation for the NN and FNO models*

The relative log-test-error shows a similar picture as for the heat equation.

Approach	Relative Log-Test-Error		
Wave	d=1	d=2	d=3
FNO	-6.42	-5.36	-1.72
NN	-6.30	-4.13	-2.18

Table 2: *Relative Log-Test-Error for the wave equation*

5 Discussion

The initial findings indicate that the parametric and FNO methodologies perform comparably well. However, it is important to recognize the substantial reduction in parameters required by the FNO model, alongside its expedited training times, which positions it as a seemingly more efficient approach. The constraints of this project, encompassing limited time and resources and man-power, restricted the depth of analysis that could be achieved. A compelling direction for future research lies in exploring predictions concerning temporal evolution, an area where the FNO method is predicted to offer significant advantages, as per literature [paper]. As such, it is suggested that this investigation be expanded to accommodate a time-series data set, facilitating the prediction of model evolution over time.