## A. Generation of the various heterogeneity profiles

This appendix lists the equations along with the relevant parameters used to create the heterogeneity profiles in the various demonstration cases shown in the paper.

*Figure 3:*

 

where  = (*a*) , (*b*) , and (*c*) .

*Figure 4(a):*

 

where  and .

*Figure 4(b):*

 

where  is a function created considering 32 randomly weighted Fourier harmonics in each direction.

*Figure 4(c):*

 

*Figure 6:* See equation (10), where  = (*a*) , (*b*) , (*c*) , (*d*) , (*e*) , and (*f*) .

*Figure 8:* Same as figure 6.

*Figure 9:* See equation (10), where  = (*a*) , (*b*) , and (*c*) .

*Figure 10(a):*

 

where .

*Figure 10(b):*

 

***Table 1.*** *The effect of the FNO hyper-parameters on the testing error under approach A, trained on striped heterogeneity profiles given by equation* (9)*. These parameters are the number of Fourier modes retained (), the width of the channel space after the lifting operator is applied () and the number of Fourier layers used (). The total number of model parameters is noted as . The hyper-parameter values adopted for the production runs are shaded in gray.*

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Error |  |    |  | error |  |    |  | error |  |
| 4 | 17.2% | 612,611 |  | 64 | 17.5% | 552,323 |  | 1 | 0.9% | 563,075 |
| 8 | 3.2% | 1,136,899 |  | 128 | 6.5% | 2,185,475 |  | 2 | 0.9% | 1,103,875 |
| 16 | 6.5% | 2,185,475 |  | 256 | 18.3% | 8,695,811 |  | 4 | 6.5% | 2,185,475 |

***Table 2.*** *The effect of the FNO hyper-parameters on the testing error under approach A, trained on heterogeneity profiles given by equation* (10)*. The table notation is the same as that of table 1.*

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | error |  |  |  | error |  |  |  | error |  |
| 4 | 0.7% | 612,611 |  | 64 | 0.8% | 552,323 |  | 1 | 0.8% | 563,075 |
| 8 | 0.6% | 1,136,899 |  | 128 | 0.7% | 2,185,475 |  | 2 | 0.5% | 1,103,875 |
| 16 | 0.7% | 2,185,475 |  | 256 | 0.7% | 8,695,811 |  | 4 | 0.7% | 2,185,475 |

***Table 3.*** *The effect of the FNO hyper-parameters on the testing error under approach B, trained on heterogeneity profiles given by equation* (10)*. The table notation is the same as that of table 1.*

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | error |  |  |  | error |  |  |  | error |  |
| 4 | 31% | 608,003 |  | 32 | 30% | 140,099 |  | 1 | 32% | 558,467 |
| 8 | 30% | 1,132,291 |  | 64 | 27% | 550,019 |  | 2 | 26% | 1,099,267 |
| 16 | 24% | 2,180,867 |  | 128 | 24% | 2,180,867 |  | 4 | 24% | 2,180,867 |

## B. Exploration of the FNO hyper-parameters

To better appreciate the specific learning architecture adopted, this appendix presents the results of a parametric exploration of the FNO hyper-parameters, which include the number of Fourier modes retained (), the width of the channel space after the lifting operator is applied () and the number of Fourier layers used (). The various models presented here were trained for 500 epochs, using a learning rate of  which was halved every 50 epochs. To facilitate the comparison, a reference model with ,  and  was considered using a dataset with  and  cases, allowing a single hyperparameter to vary across different runs.

The parametric exploration for each of the different training scenarios within approach A (described in 3.1.1 and 3.1.2) and approach B (described in 3.2) are listed in tables 1–3. Furthermore, the errors obtained when using different activation functions are shown in table 4. The parameters adopted for the best performing model for each approach are as follows:

• Approach A, trained on varied striped heterogeneity profiles given by equation (14) (3.1.1): , , , ReLU activation.

• Approach A, trained on heterogeneity profiles given by equation (15) (3.1.2): , , , ReLU activation.

• Approach B (3.2): , , , GELU activation.

***Table 4.*** *The effect of the different activation functions on the testing error under approaches A and B, trained on heterogeneity profiles given by equation* (10)*. The activation functions adopted for the production runs are shaded in gray.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| function |  | error app. A |  | error app. B |
| ReLU |  | 0.7% |  | 28% |
| GELU |  | 9.6% |  | 24% |
| leaky ReLU |  | 0.8% |  | 28% |
| sigmoid |  | 17.7% |  | 97% |
| tanh |  | 21.0% |  | 32% |