IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Projects on Recent Advances in Machine Learning

Neural Operators

Gonzalo BECKER
Alex SZAPIRO
Álvaro SCARRAMBERG

29/03/24

Index

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Neural Operators

Neural Operator: Learning Maps Between Function Spaces With Applications to PDEs - Kovachki et al. (2022)

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Neural networks try to learn mappings between functions.

They satisfy the property of being universal approximators.

Universal approximation: a model that can uniformly approximate any continuous operator.

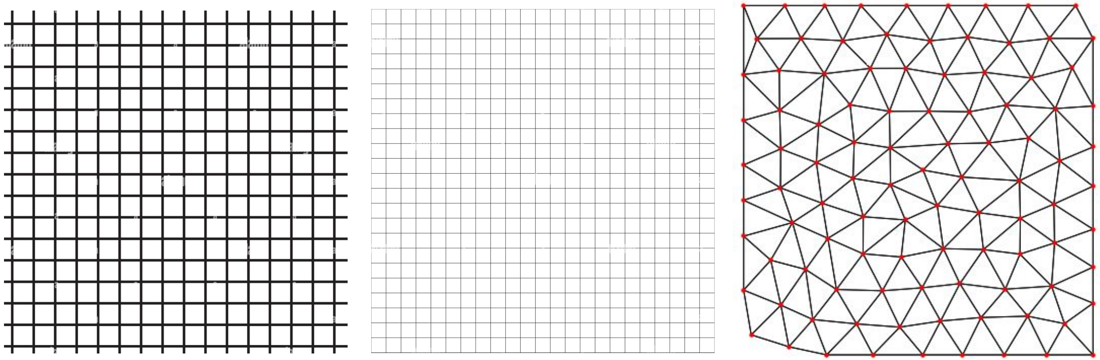
Caveat: Typical neural networks do so for finite dimensional spaces.

What if we wanted instead to learn an operator such as a PDE ?

Neural Operators: the only models that satisfy both universal approximation and discretization invariance.

A discretization invariant model satisfies:

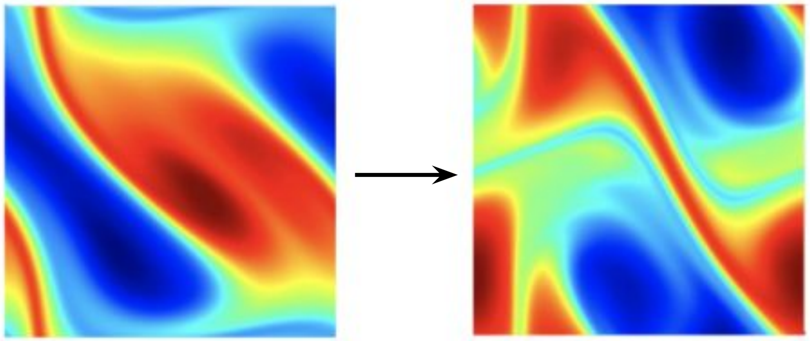
1. can act on any discretization of the input function, i.e. accepts any set of points in the input domain,
2. can be evaluated at any point of the output domain,
3. converges to a continuum operator as the discretization is refined.



Why is this important?

Burgers, Darcy subsurface flow, Navier-Stokes equations, etc.

Numerical solvers are expensive! This method allows to obtain solutions in a manner that is several orders of magnitude faster.



Navier-Stokes equation

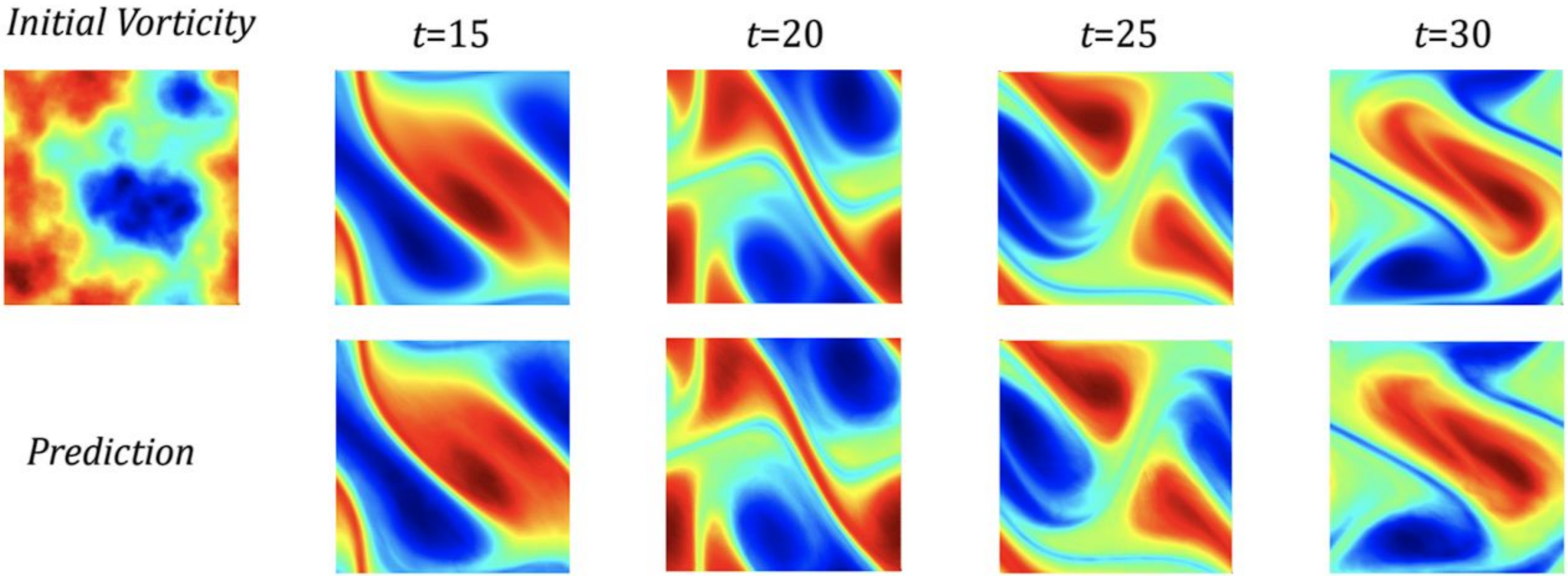
We aim to estimate an operator (PDE), using a parametric map between infinite dimension spaces (Neural Network):

$$G_\theta : \mathcal{A} \to \mathcal{U}, \quad \theta \in \Theta$$

The inputs and outputs of the network are functions.

In practical terms they are equivalent to data points, or samplings of those functions.

Neural Operators
Example



Neural Operators
Variants

There many variants of neural operators:

- graph-based operators
- low-rank operators
- multipole graph-based operators
- **Fourier operators**

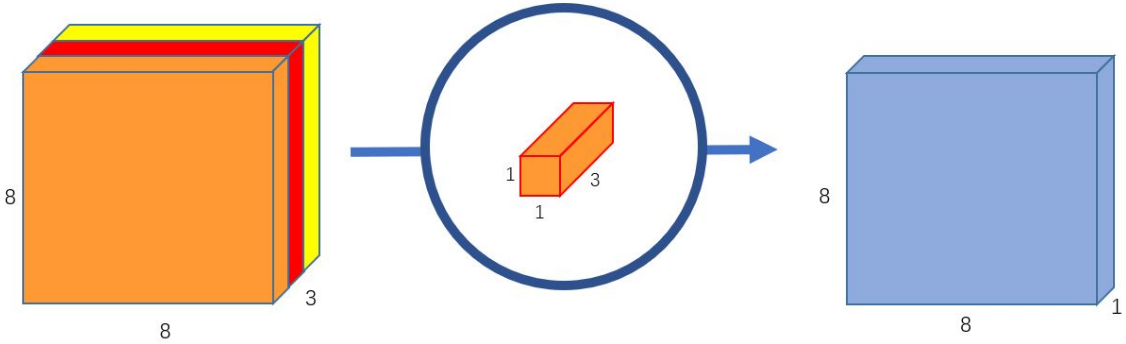
Fourier Neural Operator

Fourier Neural Operator for parametric partial differential equations - Li et al.(2021)

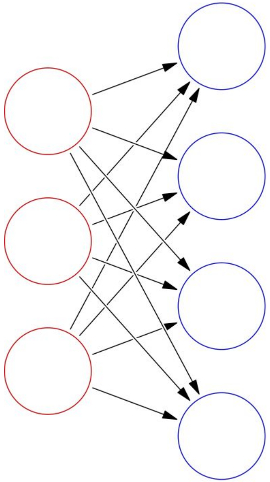
IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

What is a pixel-wise or 1x1 convolution?

Only pixels from different channels are combined.

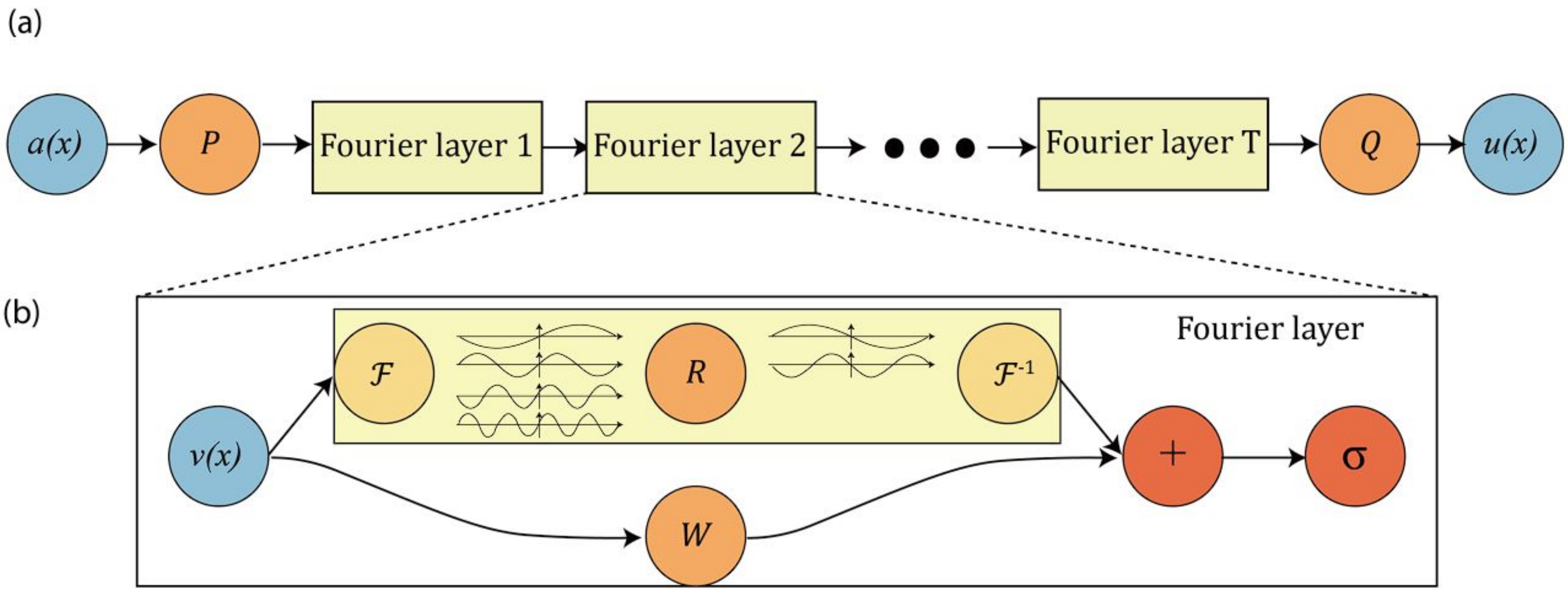


The operation remains the same no matter the resolution.

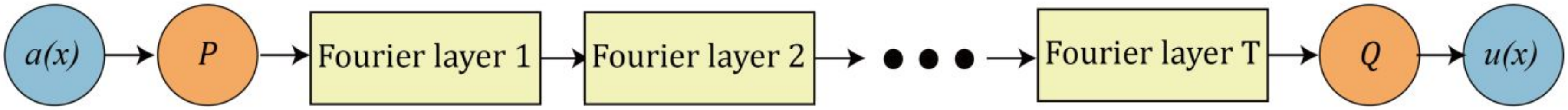


The Fourier Neural operator achieves this while also being *Resolution Invariant.*

How?



Fourier Neural Operator



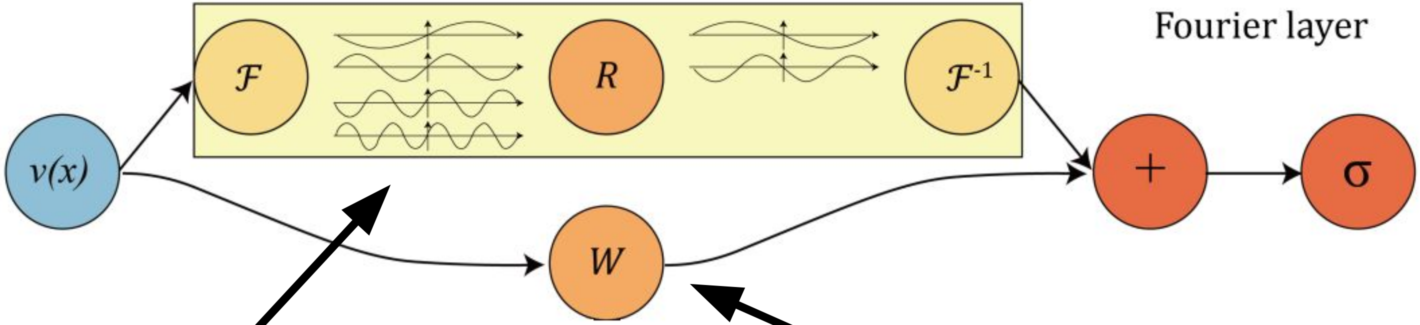
$$v_0(x) = P(a(x))$$

$$u(x) = Q(v_T(x))$$

The input is translated into a latent space with a higher dimensionality

In this case they use 1x1 convolutions

Fourier Neural Operator



This is a 1x1 convolution.

How is this block resolution invariant?

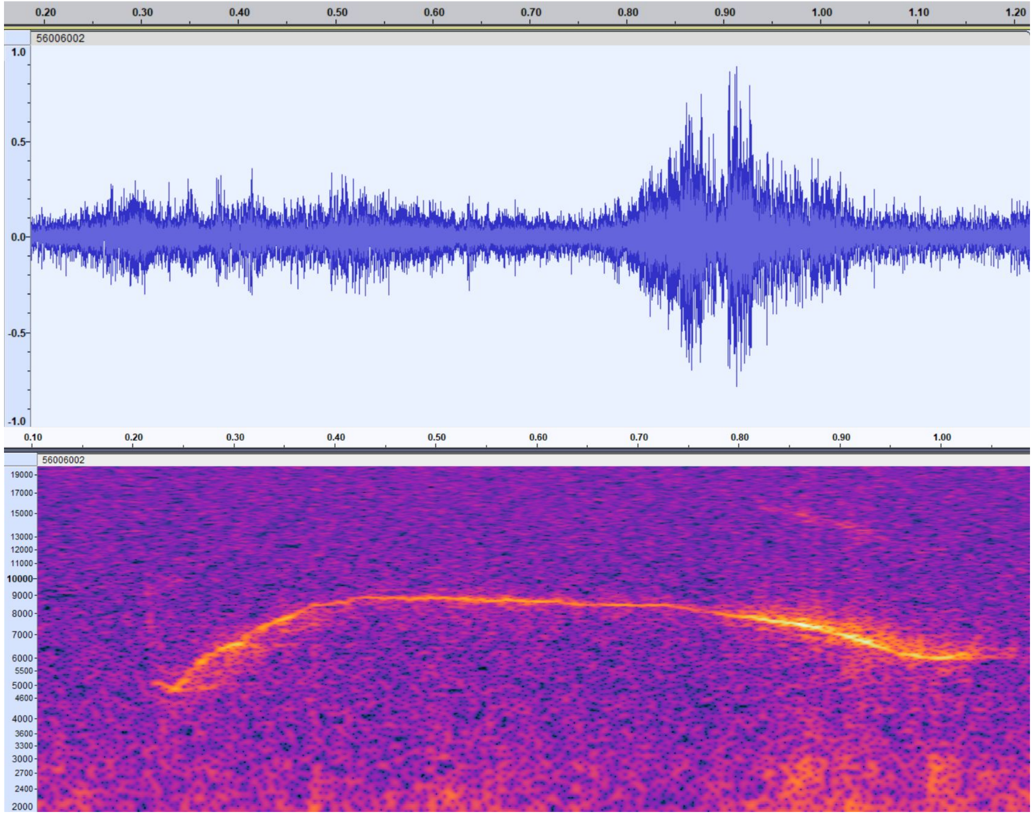
This layer emulates convolutions as matrix products in the fourier space

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Fourier Neural Operator

Signal sampled at 66 kHz

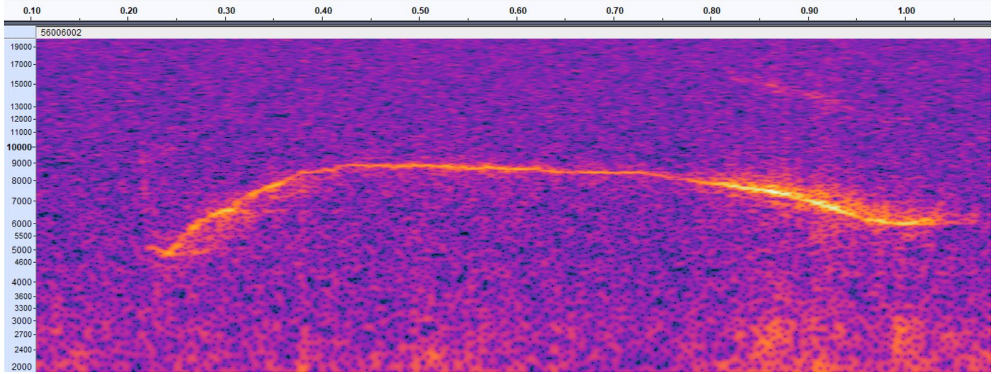
The maximum frequency obtained in the FFT is 33 kHz



Fourier Neural Operator

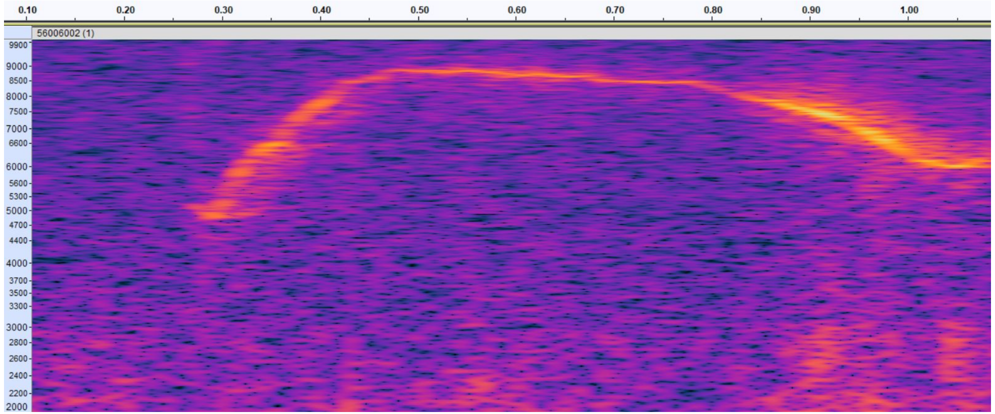
Sample Rates:

66k



Lowering the SR lowers the maximum frequency...

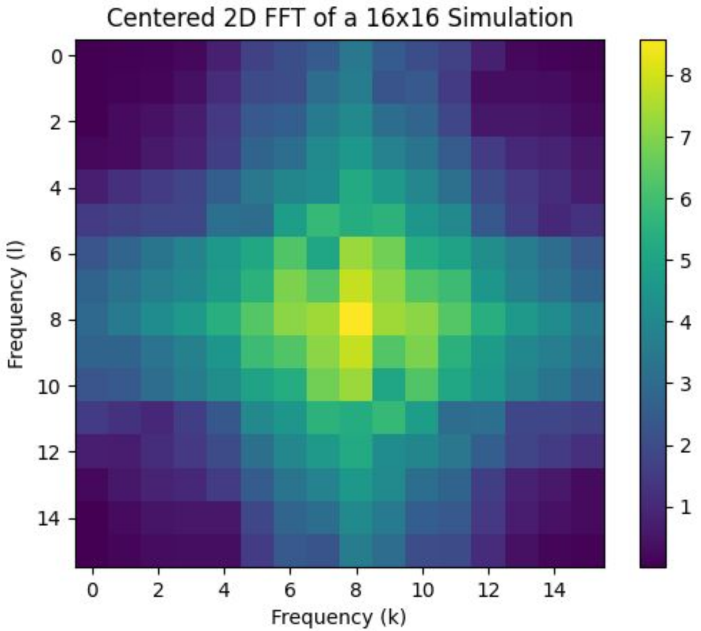
21k

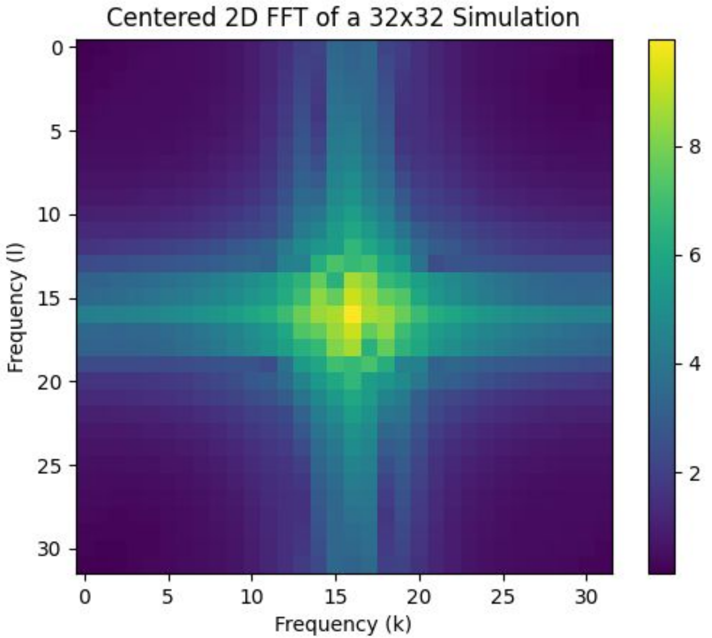


But it will not change the lower frequency values of the FFT!

Resolution invariance in the Fourier Space

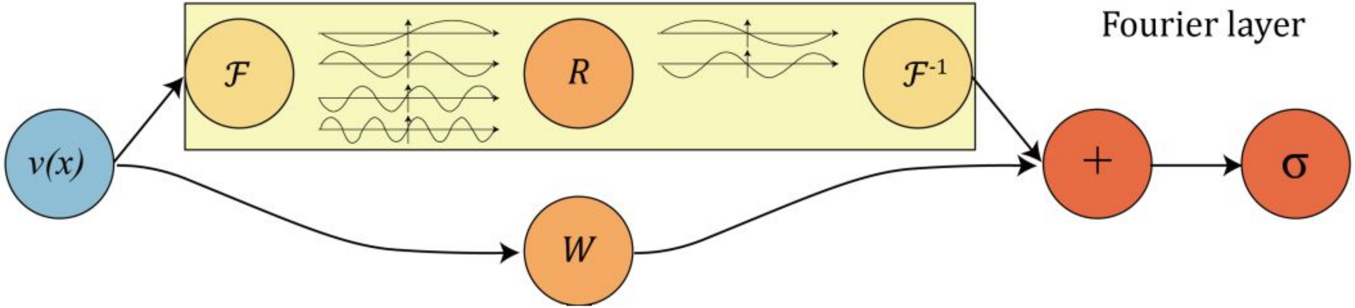
The FNO uses only the base modes of the input spectrum.





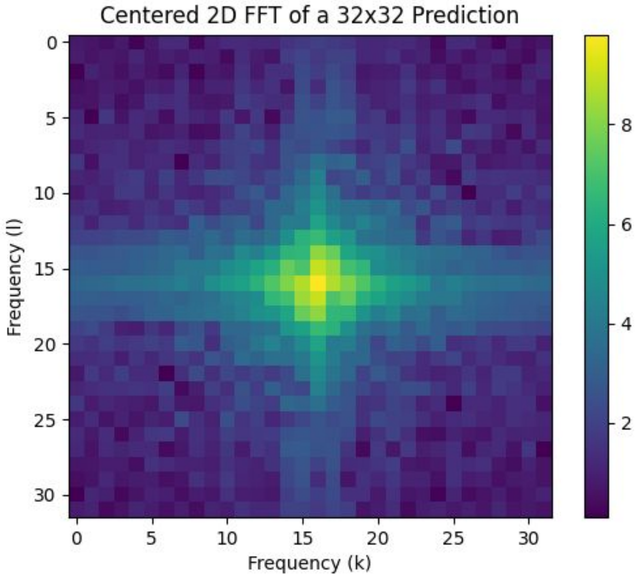
Doubling the spatial sampling frequency should yield the same values for the base frequencies.

Fourier Neural Operator



From our perspective the main drawback of the FNO is:

Since the FNO discards higher frequency modes its output will only recreate the lower frequency components of the solution.

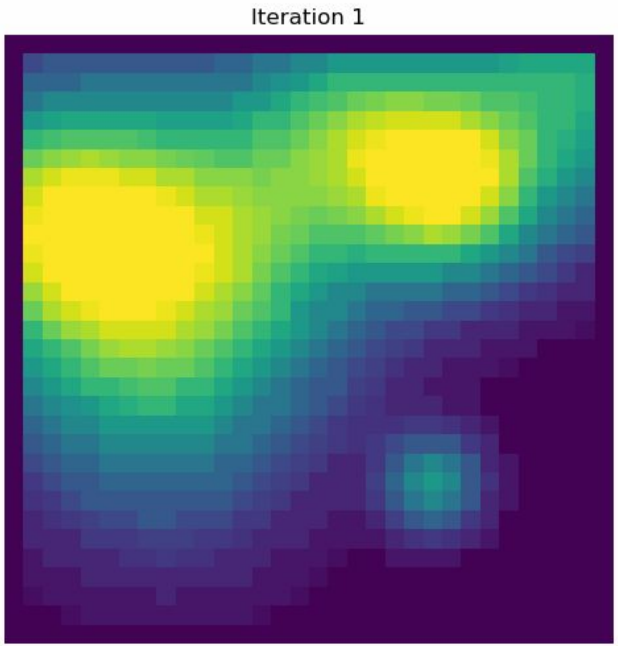


Dataset definition

for the heat equation

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Time evolution



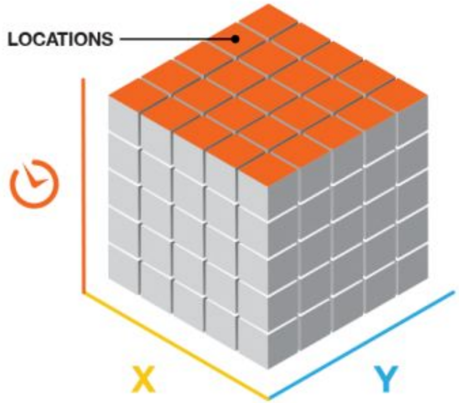
Boundary condition:
T = 0 for borders

Problem statement

$$\boxed{\frac{\partial u}{\partial t} = \Delta u}$$

Differential equation relating time and spatial variables

Spatial and temporal discretization



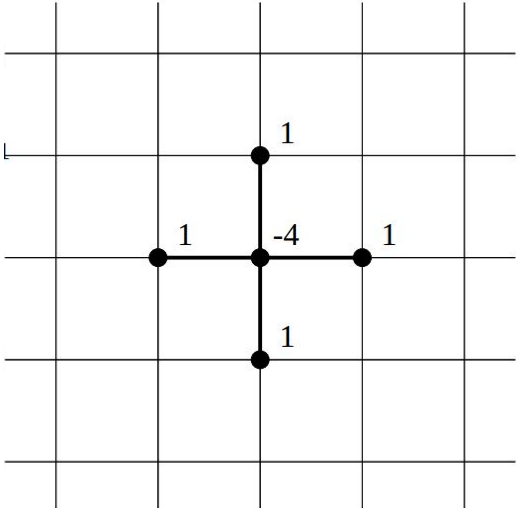
Discretization of the problem

$$\boxed{\frac{\partial u}{\partial t} = \Delta u} \Longrightarrow \frac{u_{i,j}^{k+1} - u_{i,j}^{k}}{\Delta t} - \alpha\left(\frac{u_{i+1,j}^{k} - 2u_{i,j}^{k} + u_{i-1,j}^{k}}{\Delta x^2} + \frac{u_{i,j+1}^{k} - 2u_{i,j}^{k} + u_{i,j-1}^{k}}{\Delta y^2}\right) = 0$$

$$u_{i,j}^{k+1} = \boxed{\gamma}\left(u_{i+1,j}^{k} + u_{i-1,j}^{k} + u_{i,j+1}^{k} + u_{i,j-1}^{k} - 4u_{i,j}^{k}\right) + u_{i,j}^{k}$$

$$\boxed{\gamma = \alpha\frac{\Delta t}{\Delta x^2}}$$

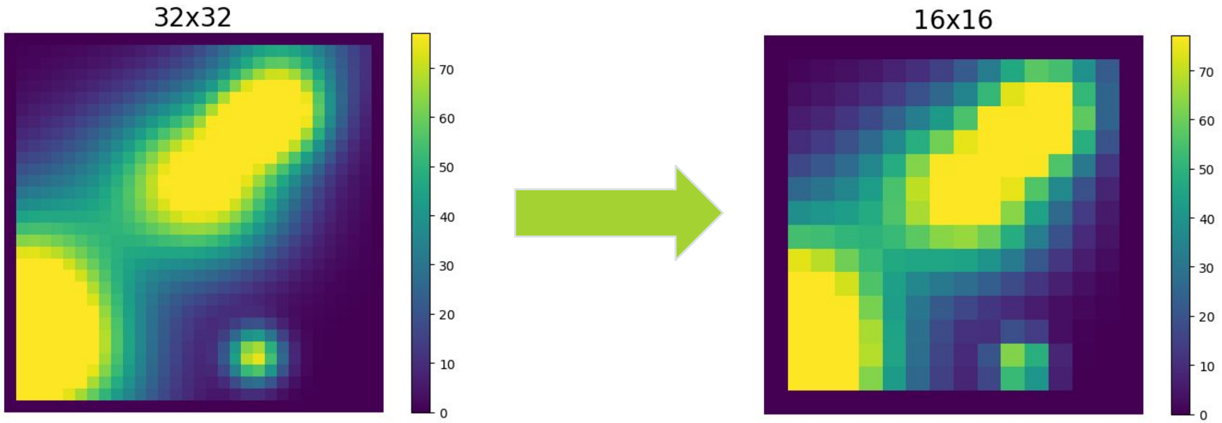
Convergence condition: $\gamma \leq \frac{1}{4}$



IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

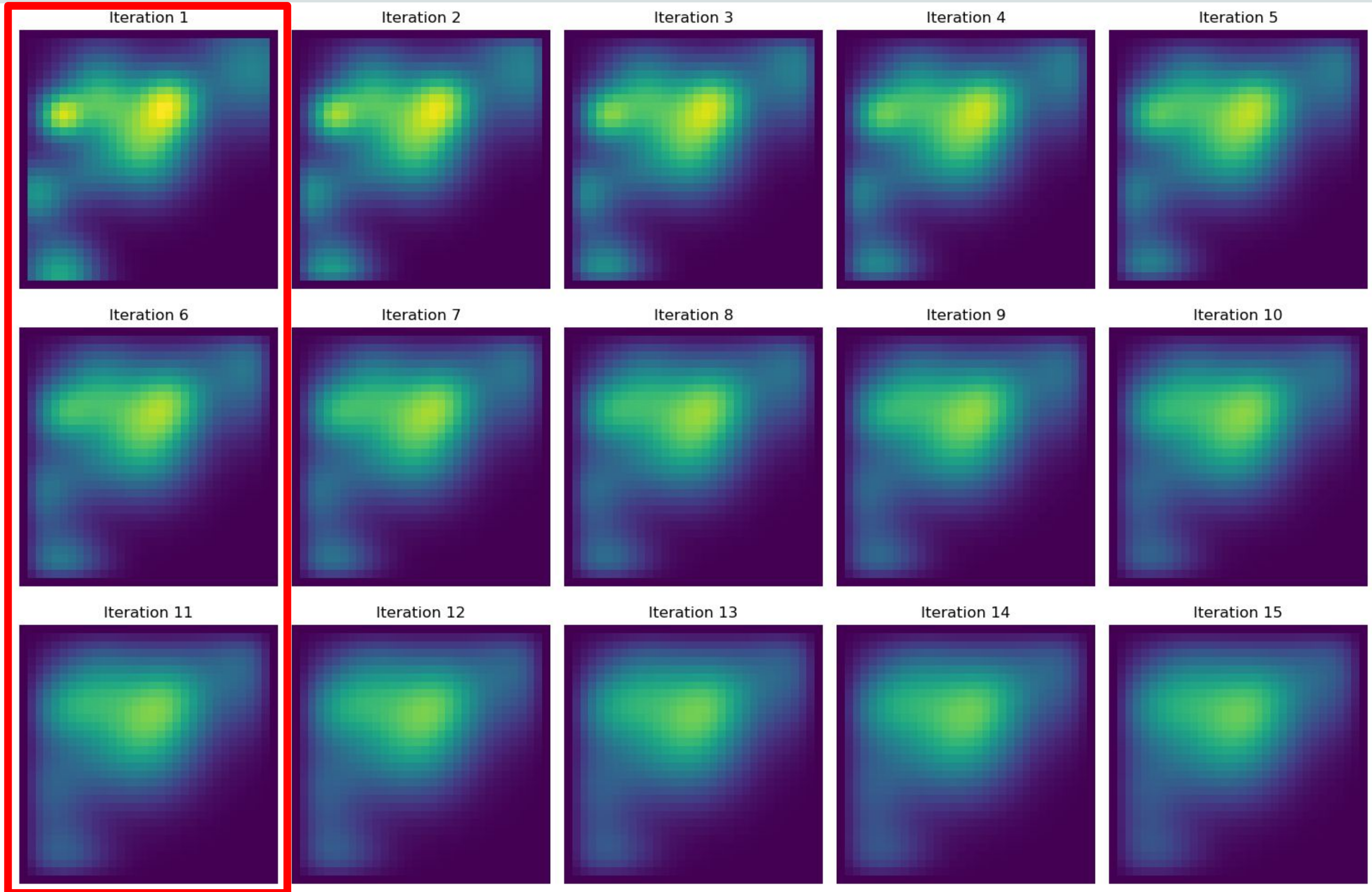
Initial Conditions

- ➢ Boundary conditions: All grid borders are set to a fixed value (0°)
- ➢ Random gaussian shaped heat areas are set in the first iteration
- ➢ To simulate lower resolutions, a subsampling is performed



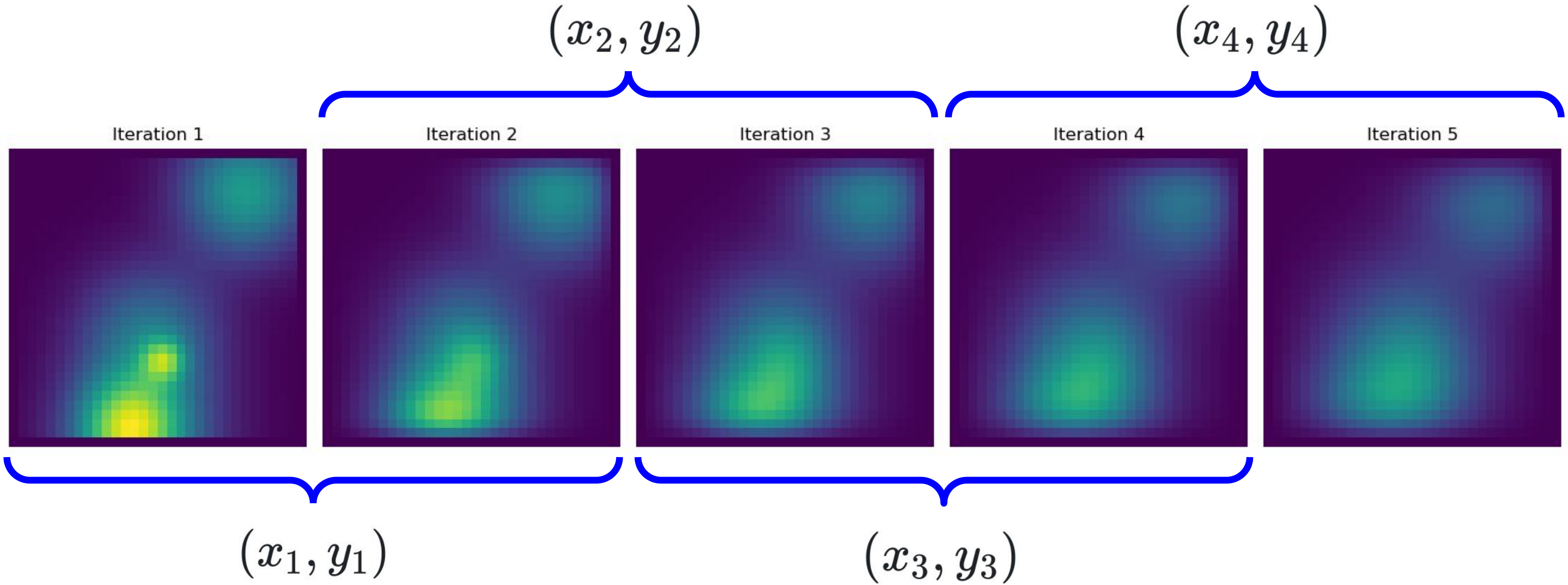
Dataset Generation

Idea: jump every 5 iterations to accelerate diffusion



IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Dataset Generation

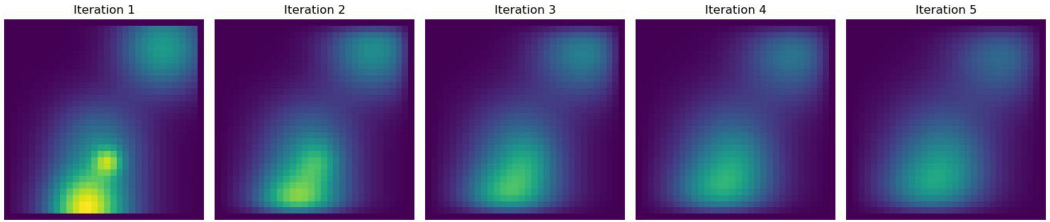


IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

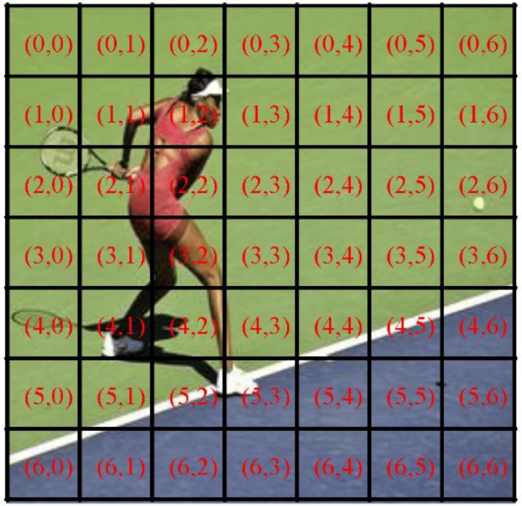
Parametric Encoding

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Time evolution prediction



Positional Embedding

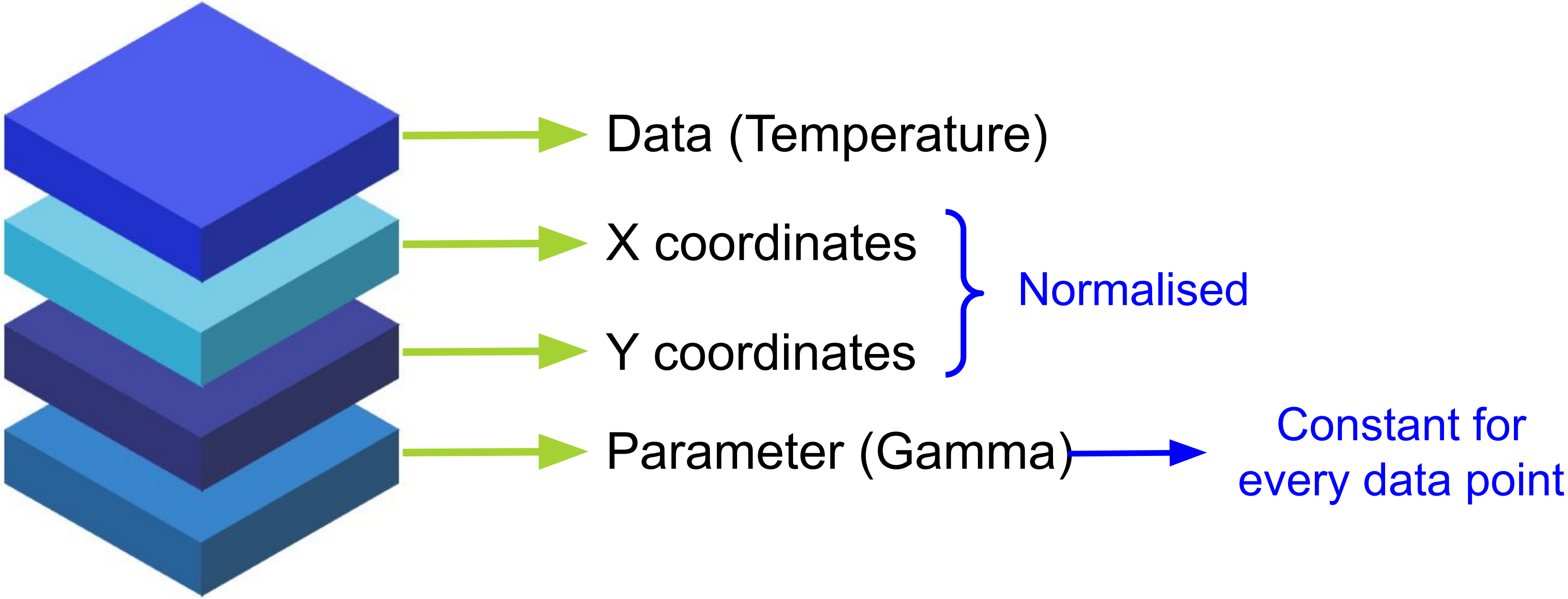


Coordinate-based Spatial Position Encoding

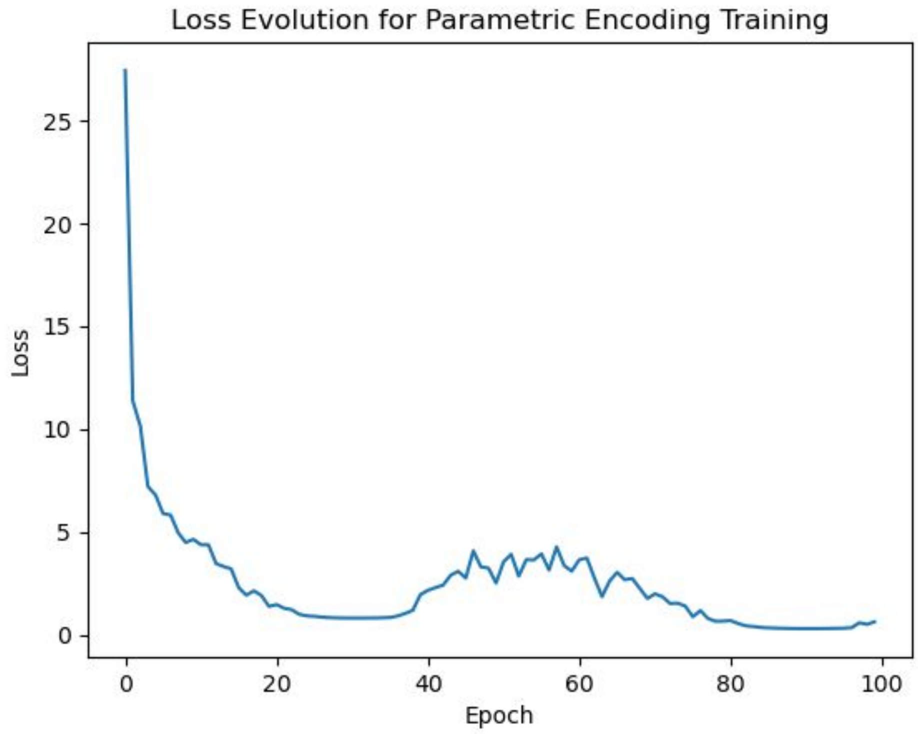
https://www.mdpi.com/2227-7390/11/21/4550

Spatial Understanding: Spatial positional embedding aids deep learning models in comprehending the spatial relationships between elements in input data.

Parametric Embedding



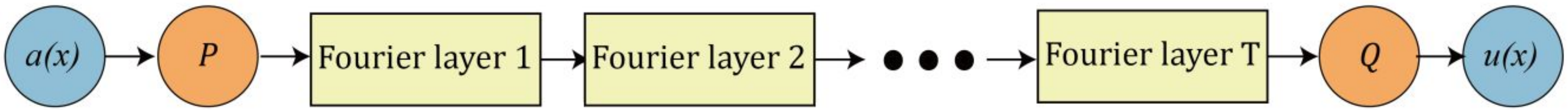
Training



- ➢ Training **γ** values: 5 values between 0.01 and 0.25
- ➢ Training duration: 2 min. on GPU
- ➢ Training size: 2386 input-output pairs
- ➢ Batch size: 32
- ➢ Number of epochs: 100

Hyperparameters for the architecture

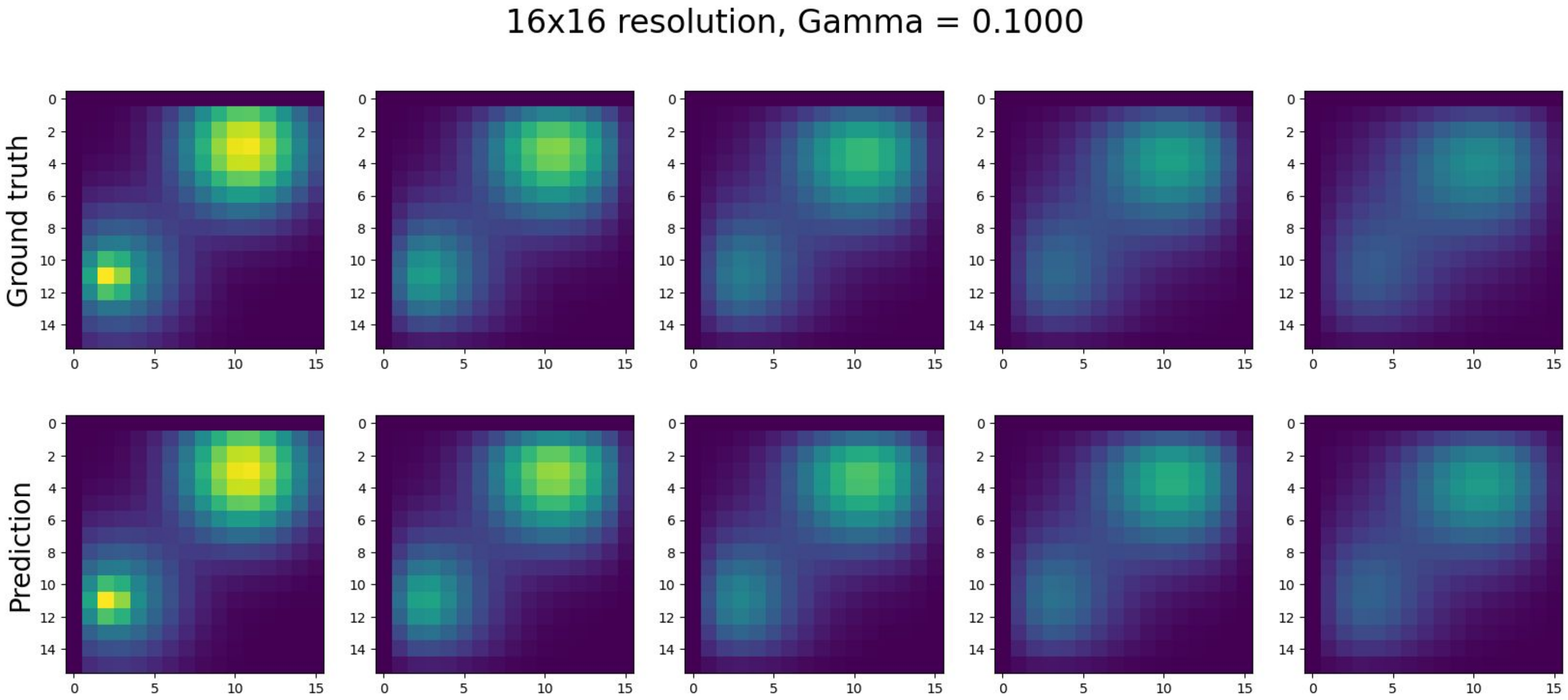
- 8 Fourier layers
- 32 hidden channels
- 64 projection channels
- 4 input channels
- Initial learning rate: 0.008
- Adam optimizer



IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

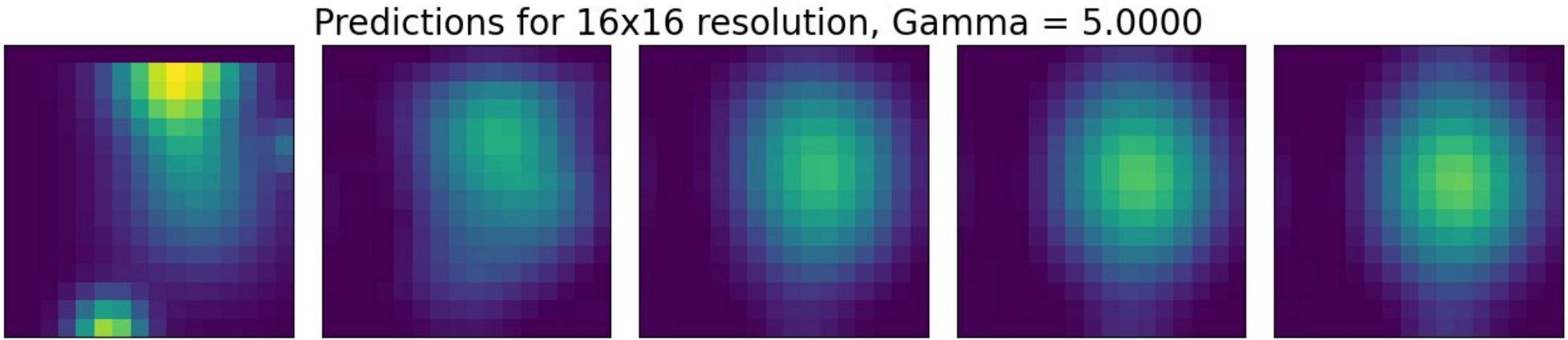
Results

16x16



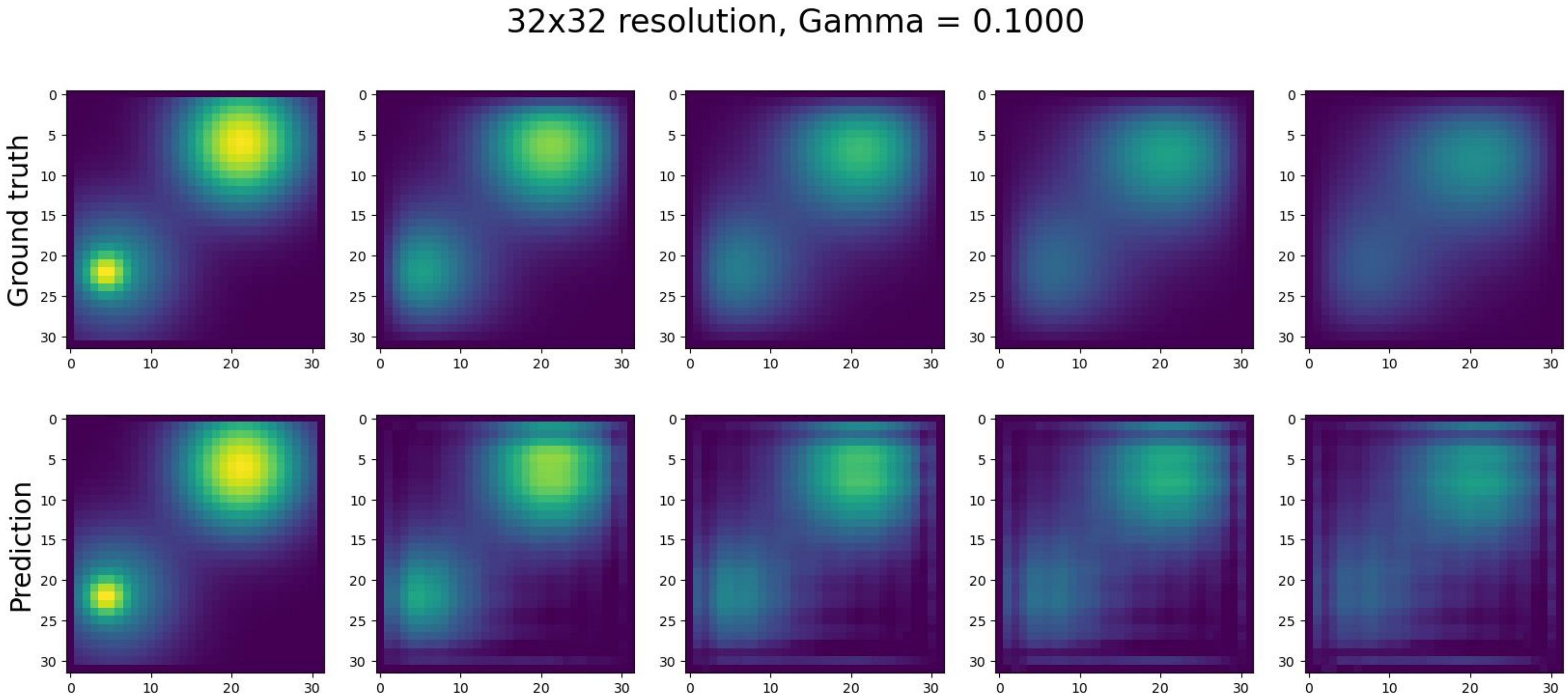
Results

What if γ is outside the training range?



IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Results
32x32



Exploring higher dimensions

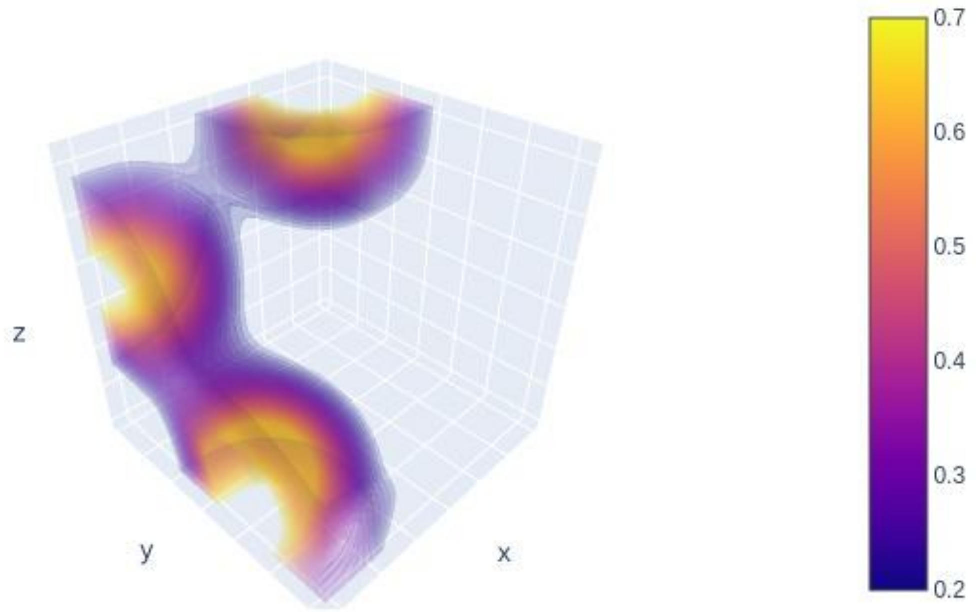
IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

3D heat equation

Initial condition

There are no constraints on the dimensionality of input data!

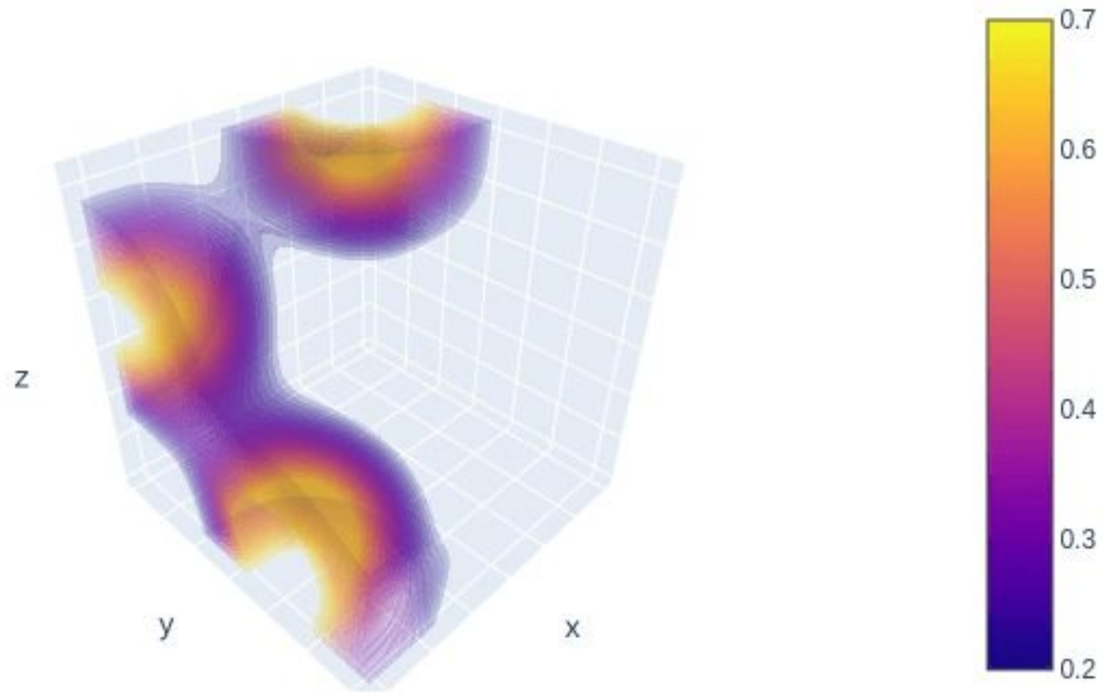
(Only the computational cost…)



$$\mathcal{F}f(\xi_1, \xi_2, \ldots, \xi_n) = \int_{\mathbf{R}^n} e^{-2\pi i(x_1\xi_1 + \cdots + x_n\xi_n)} f(x_1, \ldots, x_n)\, dx_1 \ldots dx_n\,,$$

Results

Time propagation of the 3D heat equation



Future improvements

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Future improvements:

- Regression to find the equation parameter.
- Upgrade the model to be N-dimensional.

Conclusion

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

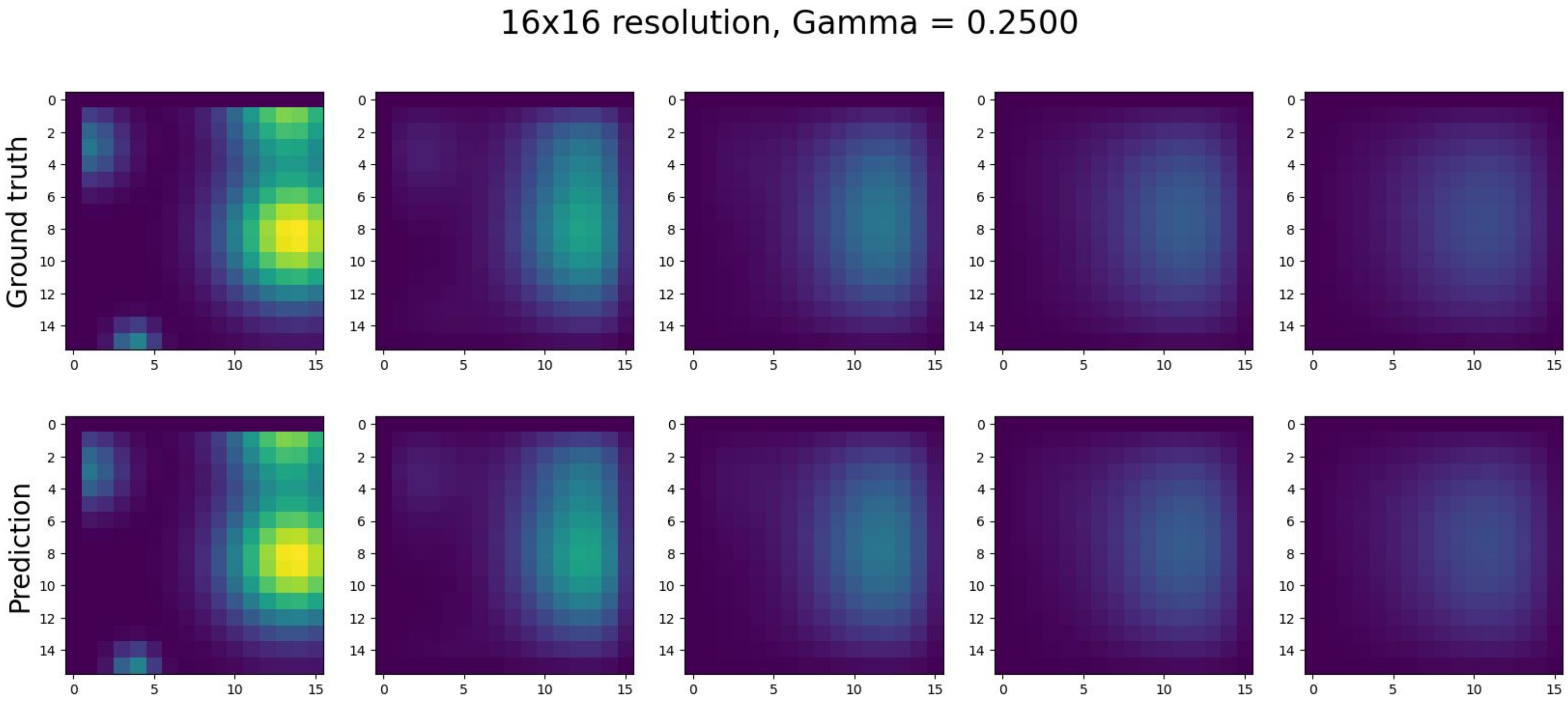
Conclusion

- We reviewed the literature on Neural Operators.
- We implemented and tested the Fourier Neural Operators for the 2D and 3D Heat Diffusion equation.
- We implemented parameter embedding into the model.

Appendix

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Results


16x16 resolution, Gamma = 0.2500
Ground truth
Prediction
0
2
4
6
8
10
12
14
0
5
10
15

Results