0.1 ConvNP

For a 1 million parameter model ConvNP we used the following configurations for the 1D and 2D datasets:

0.1.1 1D

- Encoder and Decoder: MLPs with 2 layers with hidden units 64, 64
- UNet: 5 layers with hidden units 64, 64, 64, 64, 64
- Discretization: 64 points per unit length
- SetConv initial lengtscale: 0.1

0.1.2 2D

- Encoder and Decoder: MLPs with 2 layers with hidden units 32, 32
- UNet: 5 layers with hidden units 32, 16, 8, 8, 16
- Discretization: 20 points per unit length
- SetConv initial lengtscale: 0.1

0.2 TNP/TETNP

For a 1 million parameter model TNP/TETNP we used the following configuration which was found through hyperparameter optimization:

- Encoder and Decoder: MLP with 2 layers with hidden units 100, 100 (D_{em}, D_{em})
- Number of Transformer layers: 3
- Number of heads: 24
- Head dimension: 100