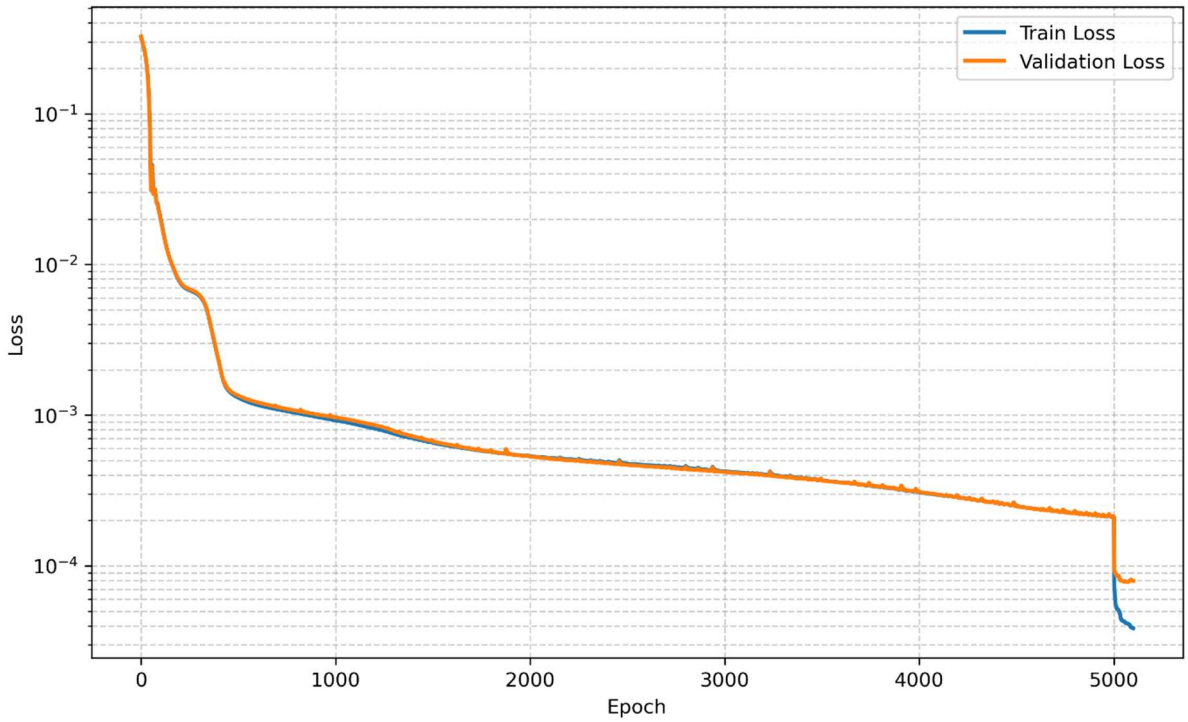


Figure S1: Loss curve for model trained on α = 0.5 and 1.5, and number of samples = 10 000. Both the training and validation loss converge during the first segment of the training (epochs< 5000) with the Adam optimiser. When the optimiser is switched to L-BFGS at the 5000's epoch, both losses go down and converge with no signs of overfitting.

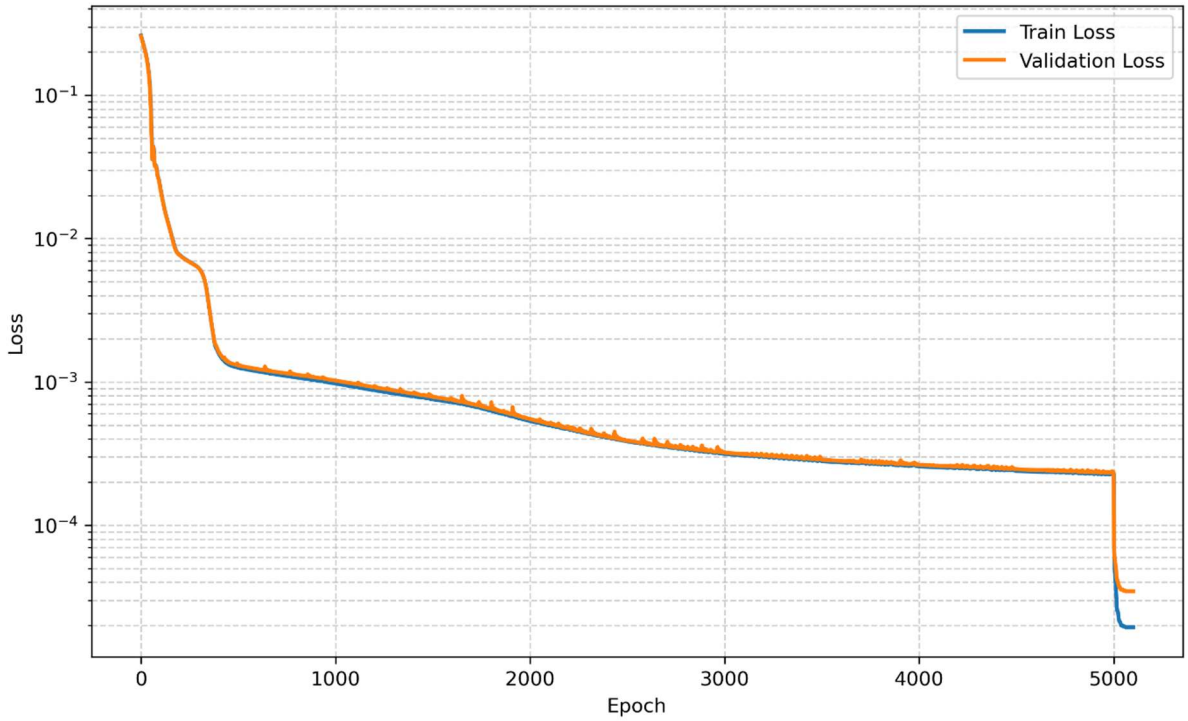


Figure S2: Loss curve for model trained on α = 0.5 and 1.5, and number of samples = 26 000. Both the training and validation loss converge during the first segment of the training (epochs< 5000) with the Adam optimiser. When the optimiser is switched to L-BFGS at the 5000's epoch, both losses go down and converge with no signs of overfitting.