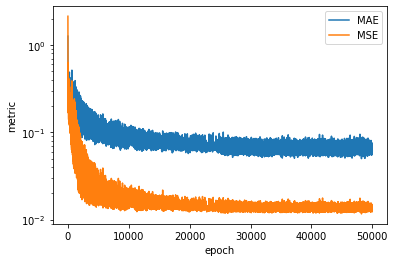
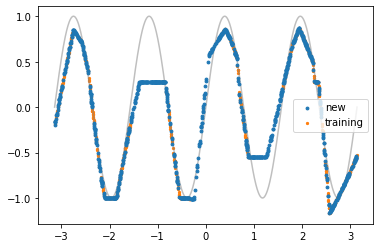
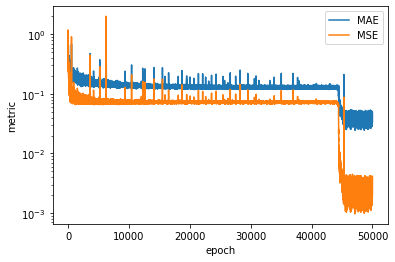
Sine curve tests:

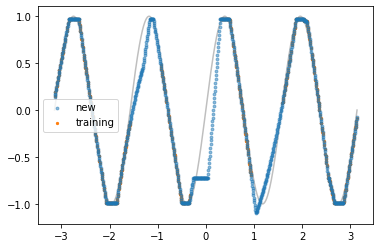
8 nodes x 3 layers + l1 0.001 regularization = evaluation loss 0.1448





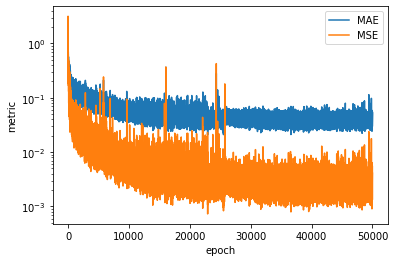
16nodes x 4layers +l1 0.001 regularization = evaluation loss 0.11440694963932037

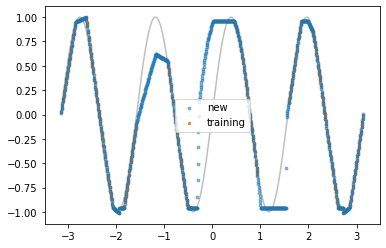




No improve after further training

16nodes x 4layers + no regularization = evaluation loss 0.14572595487535

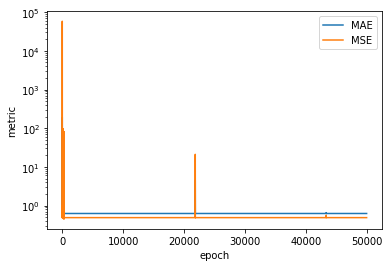


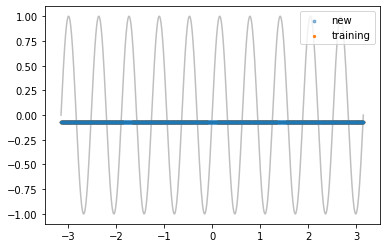


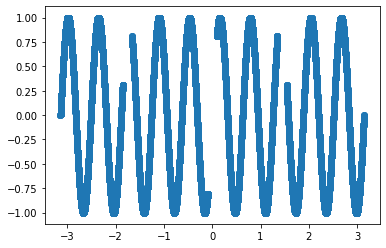
No improve after further training

Increasing the frequency of the sine wave under the theory that it will be more able to approximate the wave if given more periods of the wave, it yielded these results:

128 nodes, 6 layers, no regularization = flat fit, bad

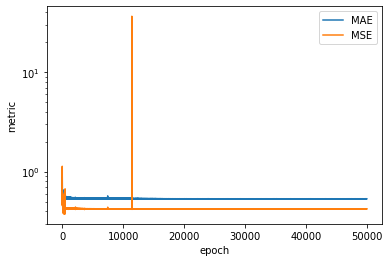


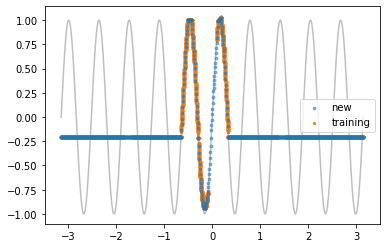




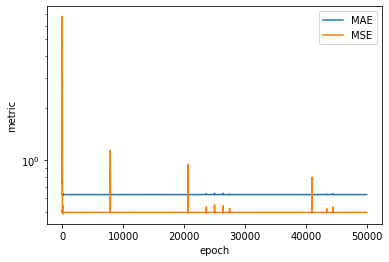
Trying to increase the number of training data points instead:

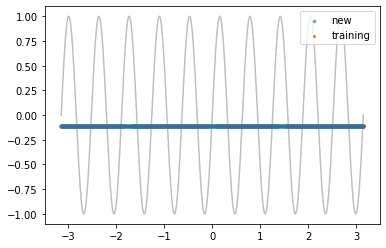
16 nodes, 4 layers, no regularization, 50k datapoints = evaluation loss 0.5464872899055481



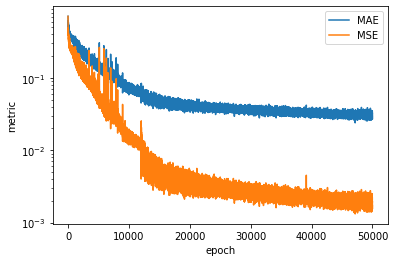


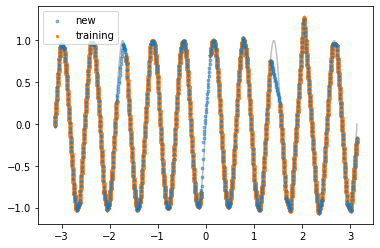
Still no luck, increasing learning rate instead, from 0.05 to 0.1:



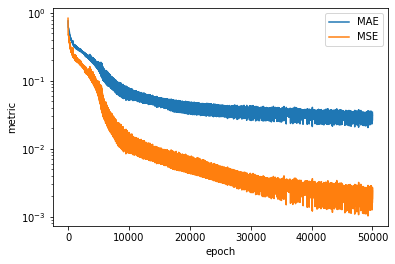


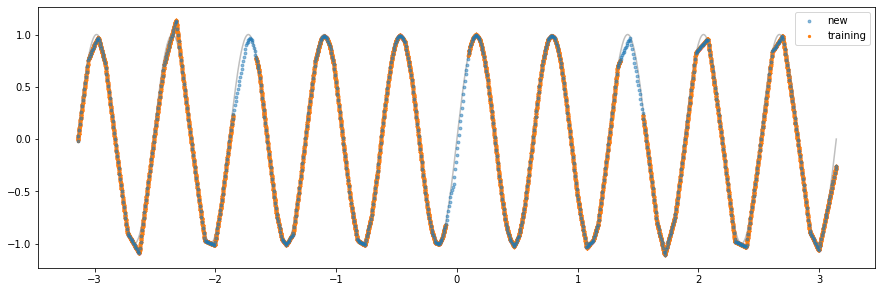
No luck, reducing learning rate instead 0.05->0.01: evaluation loss = 0.041717229545116426





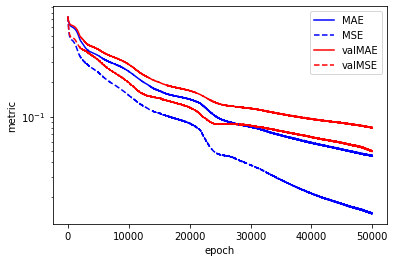
And now this is right direction! Further lowering the learning rate: 0.01->0.001: evaluation loss = 0.034996628925204276

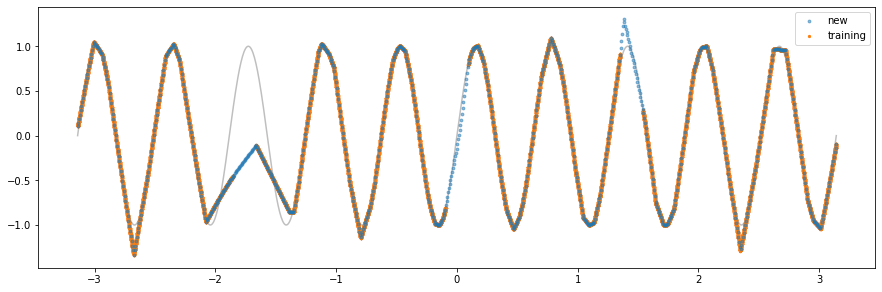




Not huge improvement in numbers but the approximation of the gaps is better,

Trying one more lower order of magnitude: lr 0.001->0.0001: evaluation loss = 0.07989501693844796





Way too slow and haven’t converge, back to lr = 0.01