Figure 1: The predicted variance is computed for all points in a test set, for both NUQLS and the NTK-GP. This figure then plots the l_2 norm of the difference between the two variance sets, which we term MSE(Variance). We plot this against (left) number of epochs of training for NUQLS and (right) number of NUQLS realizations S. These values are averaged over 5 random realizations. Note that the MSE(Variance) cannot be zero, due to the high condition number of the NTK ($\approx 10^6$); this affects the solution of the linear system $K_{XX}\mathbf{y} = K_{Xx}$ in the computation of the variance for the NTK-GP. However, we see clear convergence of our method to the distribution of an NTK-GP.