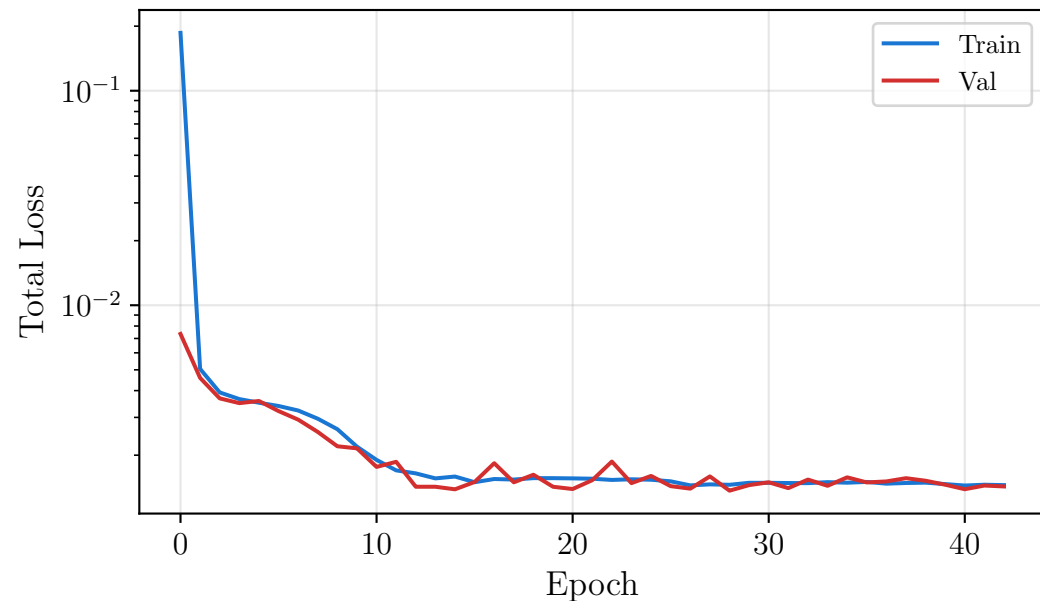
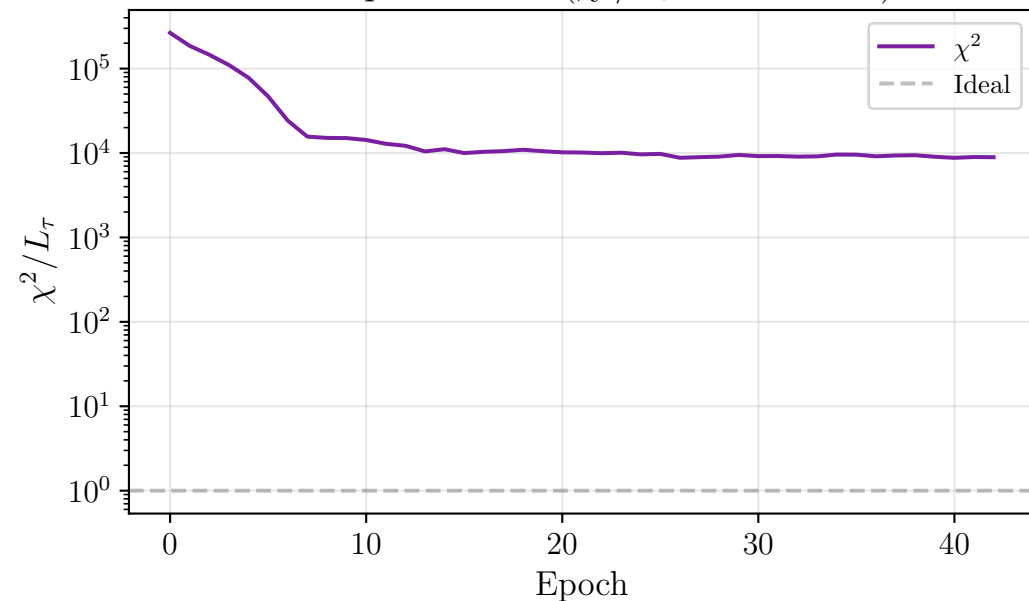


Training Curves (pretrain)

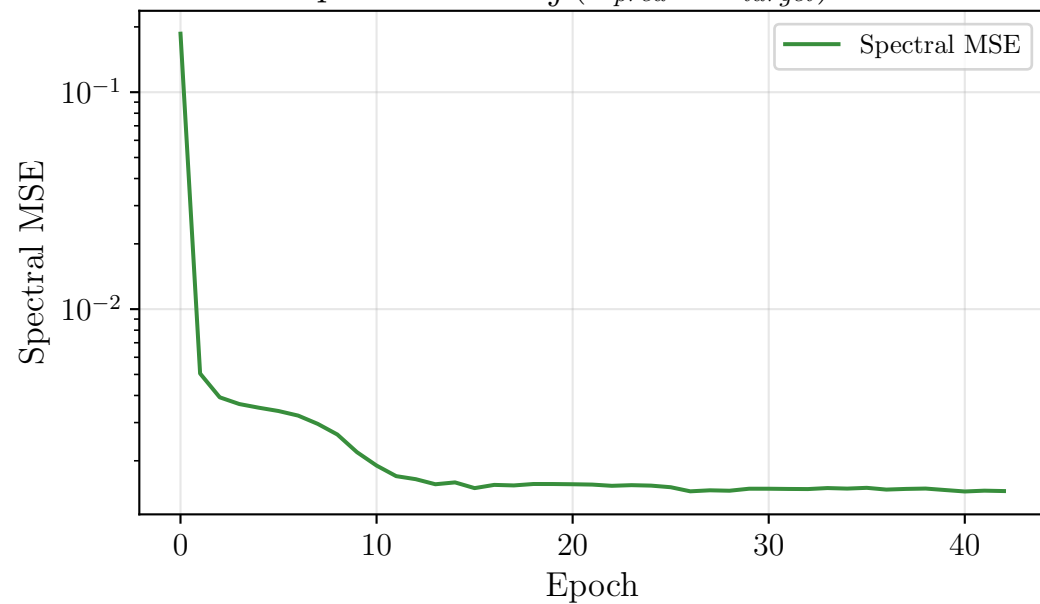
Total Loss



Chi-squared Loss ($\chi^2/L_\tau \rightarrow 1$ is ideal)



Spectral MSE: $\int (A_{pred} - A_{target})^2 d\omega$



Smoothness + KL

