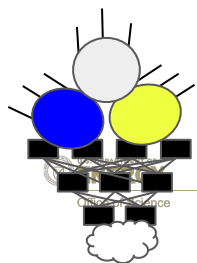




# GELATO Update 5/17/24

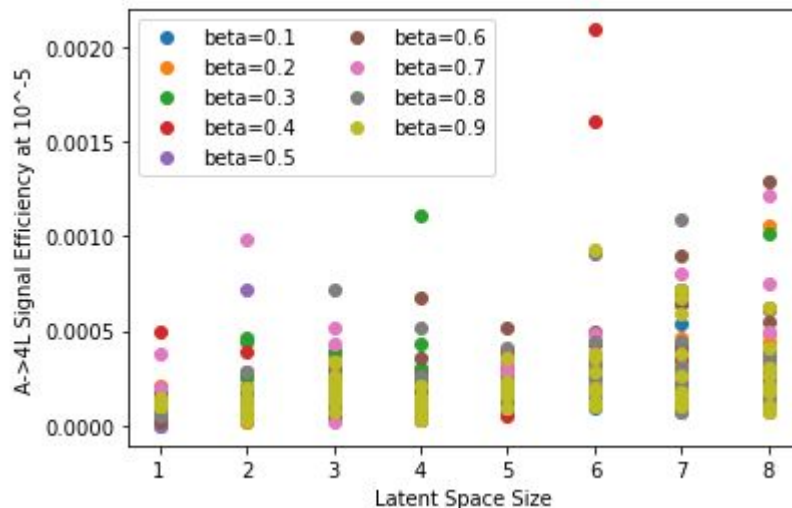
Liam Sherman



# Hyper Parameter Dependencies/Stability



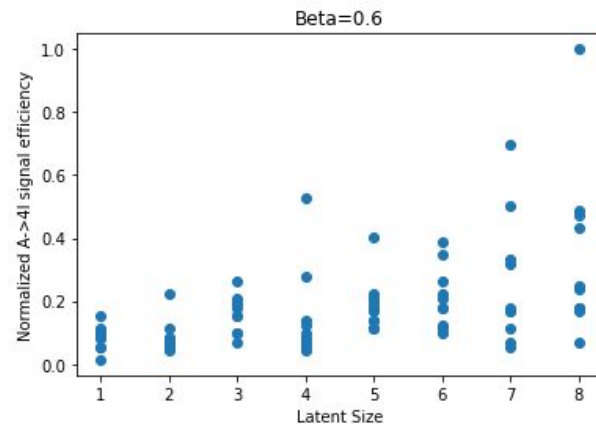
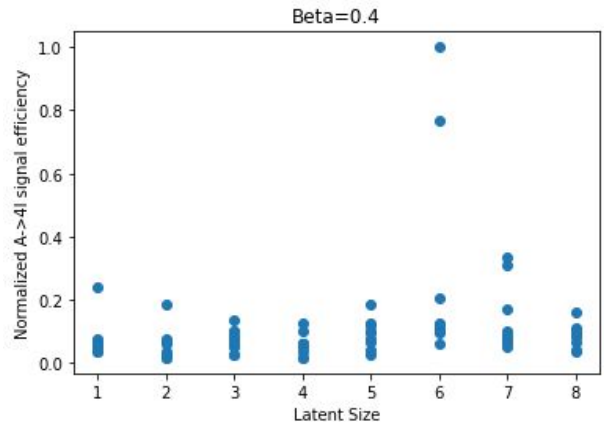
- I trained 10 models at each latent size/beta combination of latent sizes [1,2,3,4,5,6,7,8] and beta values [0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9]
- Higher beta  $\rightarrow$  KL is a higher proportion of the loss



# Hyper Parameter Dependencies/Stability

SLAC

- Here's two easier to read slices of that table
- Beta 0.4 had the best overall signal efficiency, but they are two significant outliers
- Beta 0.6 is more the norm, they both show a general trend of being better at higher latent size but still having lots of variance



# Hyper Parameter Dependencies/Stability



- To try to visualize the drift in model performance at each beta/latent size I calculated for each latent size and beta,  $(\max(\text{efficiency}) - \min(\text{efficiency})) / \max(\text{efficiency})$
- For nearly every model, the efficiency varies by more than 50% between the maximum and minimum values!
- I think this tells us we need to seriously think about our model-averaging strategy

