

t-SNE embeddings of error functions  $\mathcal{E}_i$   
(wikitext\_bert-large-uncased\_params-random\_dense-on\_layers-24, sample 2)

