

t-SNE embeddings of error functions  $\mathcal{E}_i$   
(wikitext\_albert-xlarge-v2\_params-trained\_dense-off\_layers-48, sample 9)

