KL Divergence vs Threshold (ChaosNLI Original) 2.4 before llama 8b original llama 70b original 2.2 gwen 7b original gwen 72b original KL Divergence 2.0 1.6 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 Threshold