divergence across pairs 1.0 sample_id cosine similarity empirical baseline 2 repetition