Disentanglement vs. OOD Generalization: RMTS 0.95 0.90 Generalization 0.85 0.80 0.75 IID OOD 0.70 Compositional 0.82 0.84 0.86 0.90 0.88 Counterfactual Intervention Acc.