Experiment 17 1e-01 experiment_name: more_optimizers 1e-02 -S: 1 initialization_method: identity regularization_parameter: 0.001 max_epochs: 10000 est_B_L1_norm batch_size: 10000 optimizer _{1e-03} -do_early_stopping: True num_controls: 10 **ADAM** gradient_clip_val: NA L-BFGS do shuffle: False do line search: True _{1e-04} stopping_threshold: 0.1 divergence_threshold: 10000 profiler: advanced 1e-05 -10 2 20 50 100 200 500 dimension