

Figure 10: Results on real-world datasets measured in SHD. Darker and lighter lines are for  $\alpha=0.05$  and 0.01, respectively. Error bars represent 95% confidence interval.

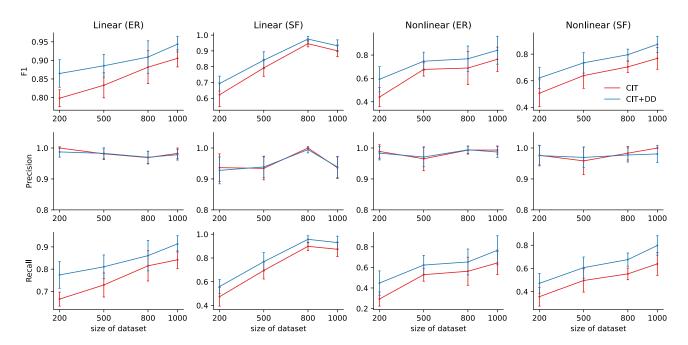


Figure 11: Results of correction experiments on continuous data with linear/non-linear relationships and different graph topologies, Erdös-Renyi (ER) and Scale-Free (SF). Error bars represent 95% confidence interval.  $|\mathbf{V}|=10$  and  $|\mathbf{E}|=24$ . For CIT, we use partial correlation and kernel-based CIT (KCI) for linear and non-linear settings, respectively.

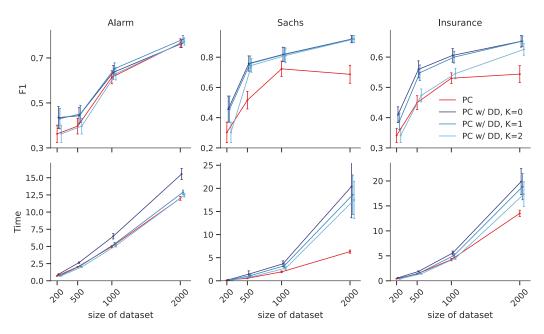


Figure 12: Ablation on K. (**Top**) F1 score. (**Bottom**) Computational performances measured as wall-clock time (seconds). Error bars represent 95% confidence interval.