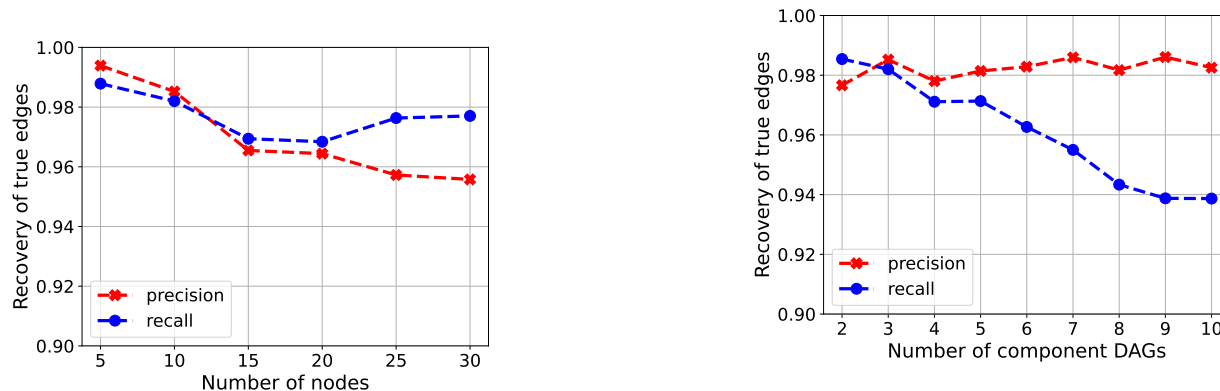


Interventional Causal Discovery in a Mixture of DAGs: Additional Experiments



(a) **Varying the number of nodes:** Precision and recall rates for recovering true edges when varying the number of nodes. $K = 3$ component DAGs are considered. The performance remains strong up to $n = 30$ nodes.

(b) **Varying the number of mixture components:** Precision and recall rates for recovering true edges when varying the number of component edges. Each DAG consists of $n = 10$ nodes. The decrease in performance is only marginal even for $K = 10$ component DAGs.

Figure 1: Precision and recall rates of the algorithm for recovering the true edges. The reported results show an average of 50 runs. $s = 5000$ samples are used.

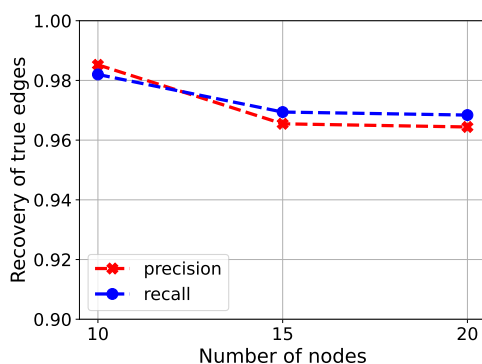


Figure 2: **Changing the true edge weights:** Precision and recall rates for recovering true edges where the edge weights of the true edges can change across the component DAGs. For instance, for $(1 \rightarrow 2) \in E_1$, $(1 \rightarrow 2) \notin E_2$, and $(1 \rightarrow 2) \in E_3$ relations, the edge weight from node 1 to node 2 can be different in \mathcal{G}_1 and \mathcal{G}_3 . The results are reported for $K = 3$ component DAGs and $s = 5000$ samples are used. The performance is virtually the same as the cases shown in Figure 1.