

An efficient search-and-score algorithm for ancestral graphs using multivariate information scores

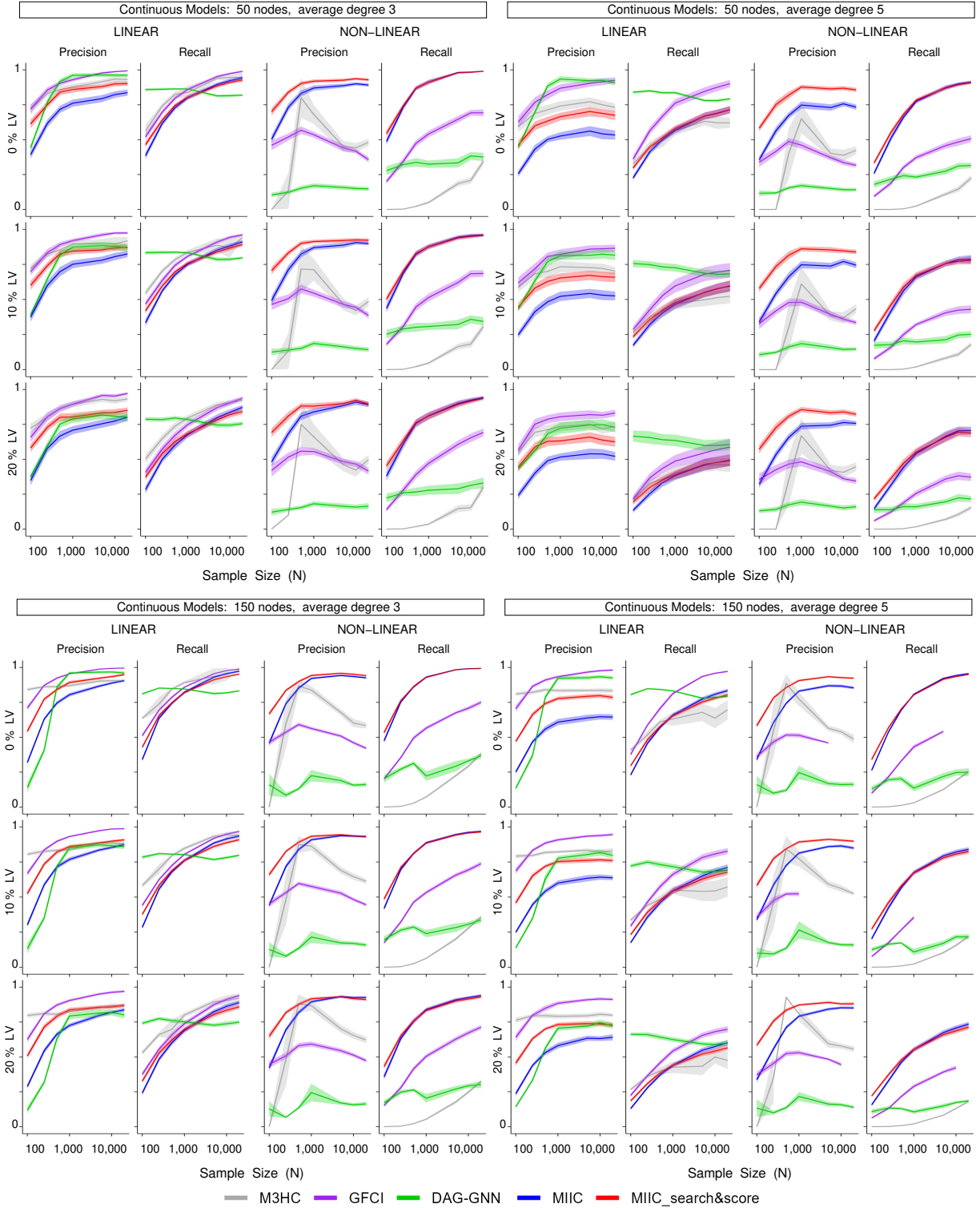


Figure E.2. Benchmark results on continuous datasets. Benchmark results are averaged over 30 independent ancestral graph models obtained by hiding 0%, 10% or 20% of variables in linear Gaussian models or more complex non-linear multimodal models including 50 or 150 nodes of average degree 3 or 5. MIIC_search&score results are compared to MIIC results used as starting point for MIIC_search&score, M3HC (Triantafillou & Tsamardinos, 2016), GFCI (Ogarrio et al., 2016), and DAG-GNN (Yu et al., 2019). Causal discovery performance is assessed in terms of *Precision* and *Recall* relative to the theoretical PAGs, while counting as false positive all correctly predicted edges but with a different orientation as the directed or bidirected edges of the PAG. Error bars: 95% confidence interval.