HCD Simulations Write Up

Audrey Fu Lab

2024-03-07

Data Simulation

We simulate hierarchical networks in a top-down approach. We consider several parameters of simulation such as sparsity, noise, and the architecture of the super level graph(s), namely small-world, and scale-free networks (Watts and Strogatz 1998; Barabási and Bonabeau 2003). We simplify our simulations by focusing on basic hierarchies with just one or two hierarchical layers.

In each hierarchy, we start by simulating the top-level nodes in topological order, using either a small world or scale-free network structure (Watts and Strogatz 1998; Barabási and Bonabeau 2003). We define origin nodes as nodes in the topological graph that have no parental input. All origin nodes are simulated from a standard normal distribtion. After generating an initial graph corresponding to the top most layer of the hierarchy, we simulate the middle and bottom layers of the hierarchy by creating groups of new nodes for each parent super-node in the upper level(s).

Each hierarchical level contains community structure nested in the previous layer. The number of offspring nodes generated for each parent node in the level above is chosen from a uniform $\text{unif}(a, b)$ distribution. We also control the connection probabilities both within and between the communities of each hierarchical layer. Once a hierarchical graph is simulated we use the hierarchy to generate the node-feature matrix which represents the expression of N genes in p samples. The number N represents the number of nodes in the observed (bottom) layer of the hierarchy and ranges between $a^{\ell+1} < N < a \times b^{\ell}$ where ℓ represents the number of hierarchical layers.

We consider three sets of hierarchical networks which represent varying difficulty levels for inference:

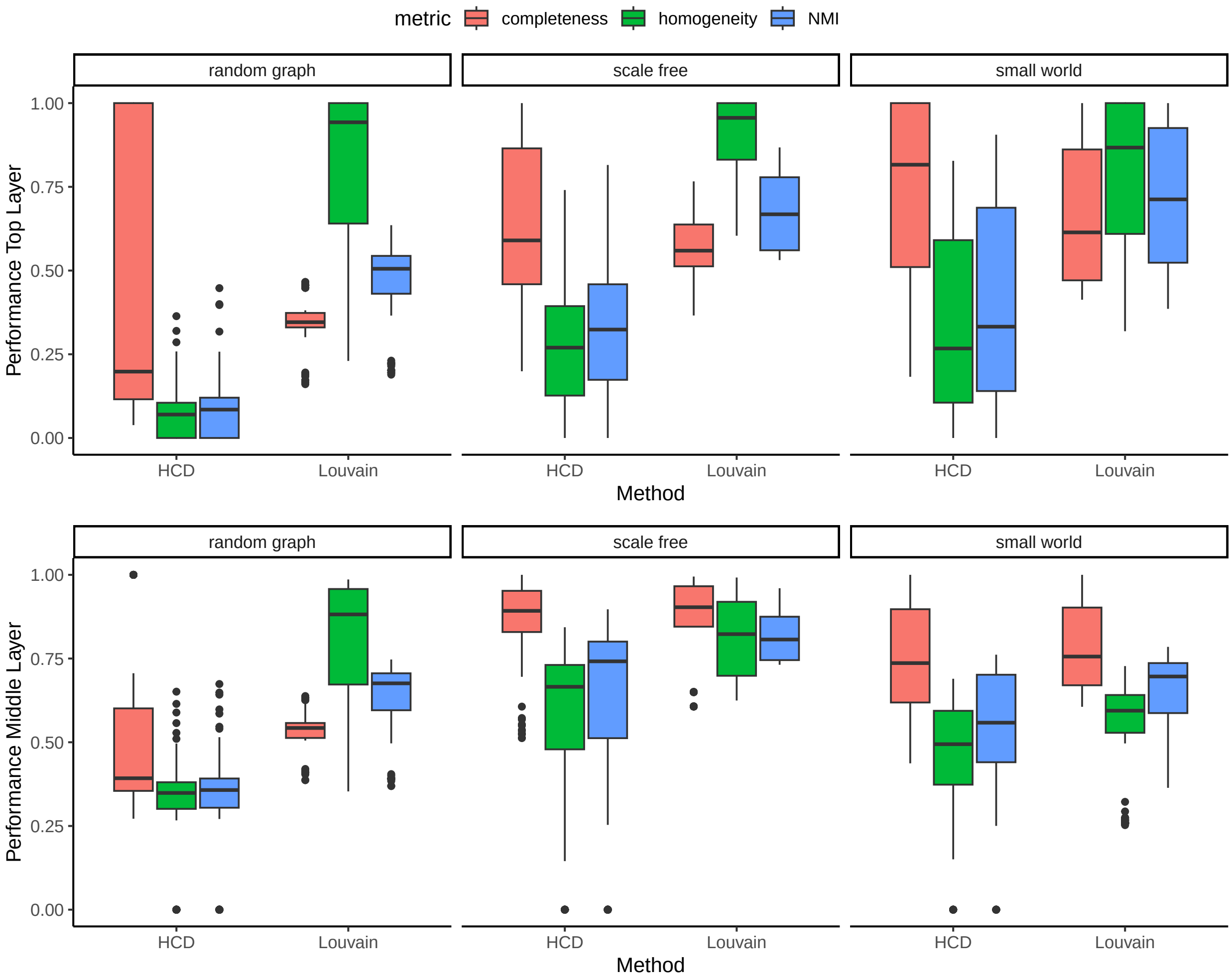
1. Complex networks - - used for final simulation assessment
2. Intermediate networks - used for investigative model tuning and performance assessment
3. Simple networks - used for code implementation and debugging

Application to Intermediate Networks

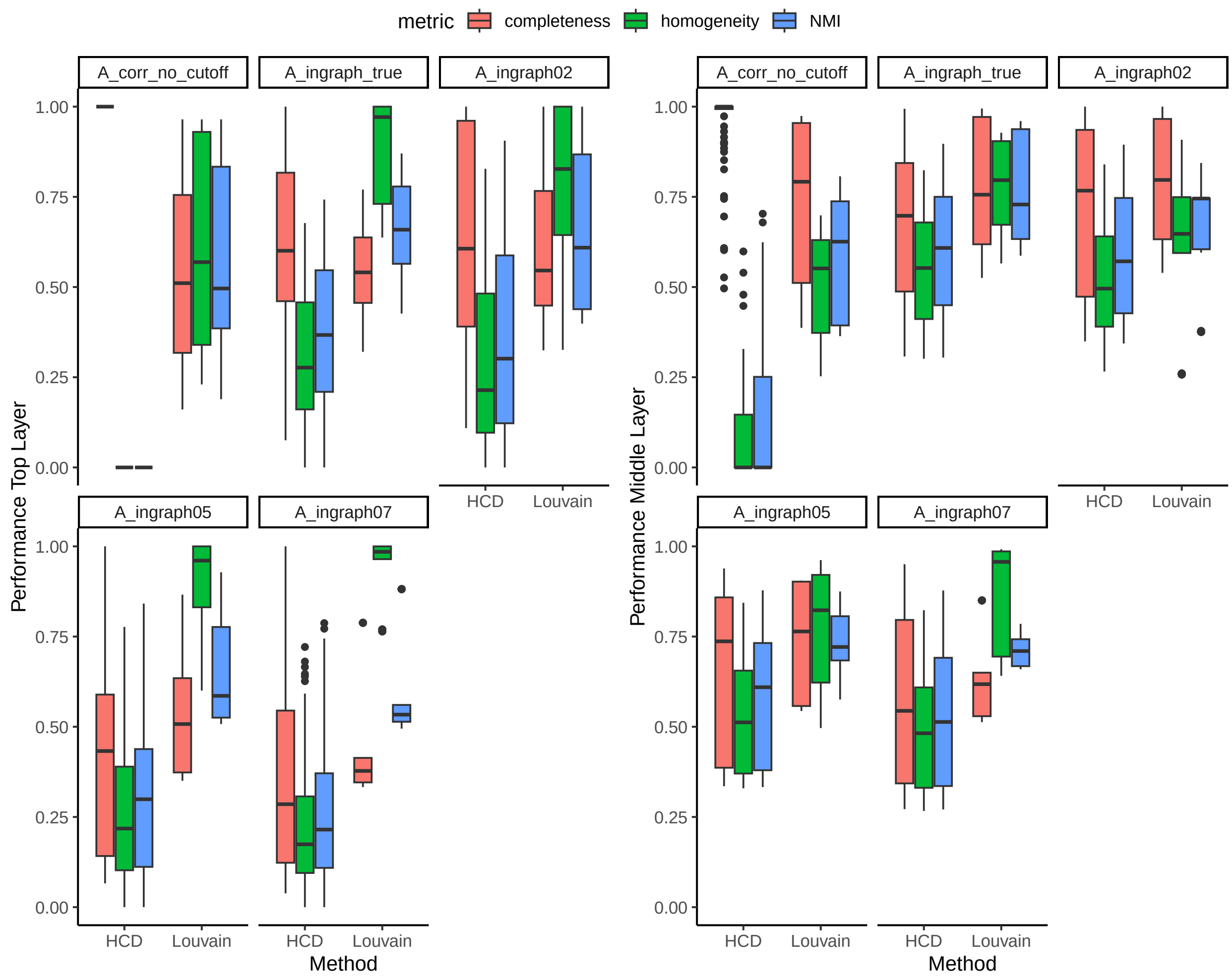
A summary of the intermediate networks can be found in **Table 1**. These

Preliminary Findings

Figures



metric
completeness
homogeneity
NMI
random graph
scale free
small world
Performance Top Layer
1.00
0.75
0.50
0.25
0.00
HCD
Louvain
Method
random graph
scale free
small world
Performance Middle Layer
1.00
0.75
0.50
0.25
0.00
HCD
Louvain
Method


metric
completeness
homogeneity
NMI
A_corr_no_cutoff
A_ingraph_true
A_ingraph02
A_ingraph05
A_ingraph07
1.00
0.75
0.50
0.25
0.00
Performance Top Layer
Performance Middle Layer
HCD
Louvain
Method

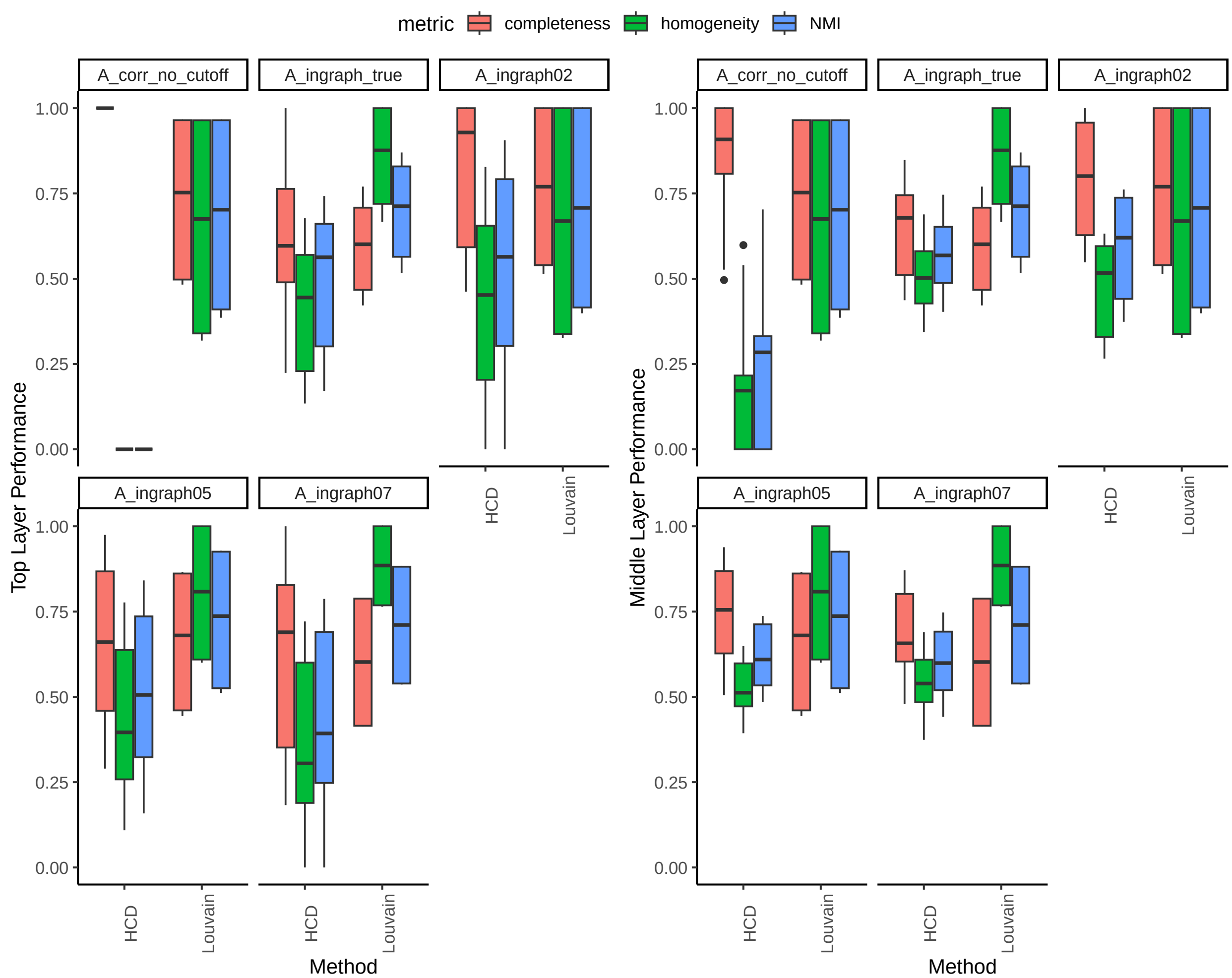


Figure 1: Small world graphs

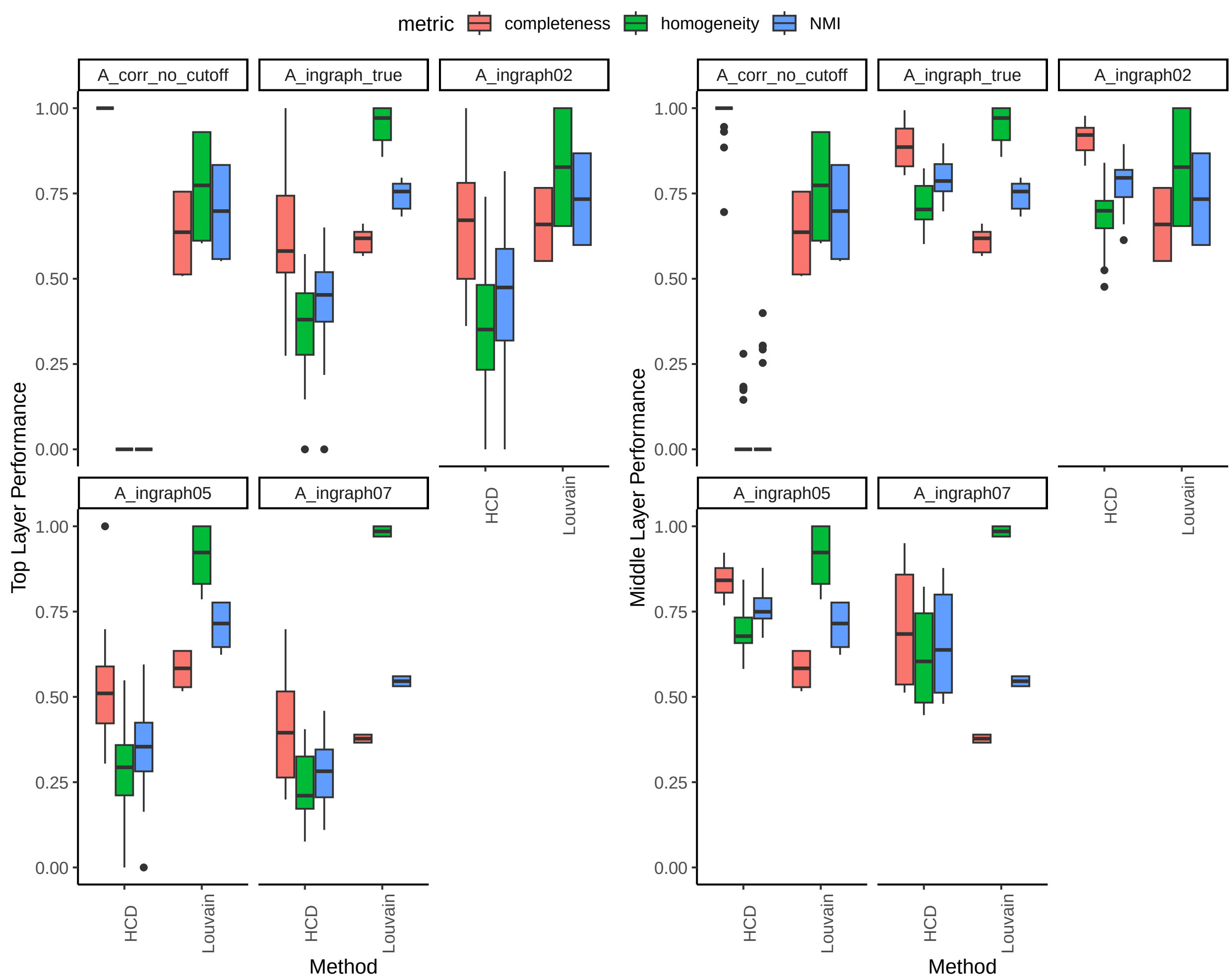


Figure 2: Scale free graphs

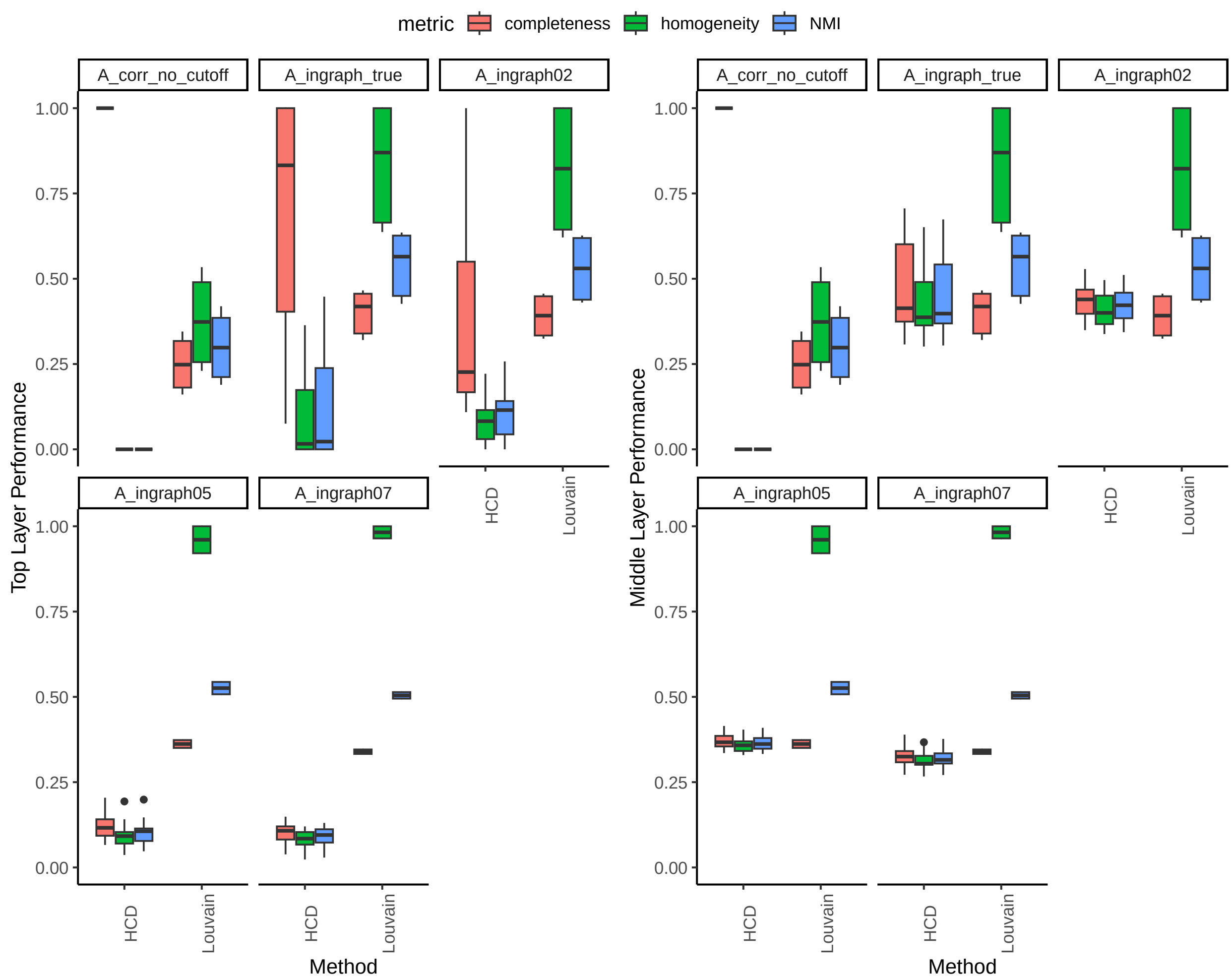


Figure 3: random graphs

Table 1: Summary statistics for intermediate difficulty simulated networks.

Value	Network1	Network2	Network3	Network4	Network5	Network6
Subgraph type	small world	small world	scale free	scale free	random graph	random graph
Connection type	disc	full	disc	full	disc	full
Layers	3	3	3	3	3	3
Standard deviation	0.1	0.1	0.1	0.1	0.1	0.1
Nodes per layer	(5, 15, 300)	(5, 15, 300)	(5, 15, 300)	(5, 15, 300)	(5, 12, 167)	(5, 12, 167)
Edges per layer	(0, 15, 358)	(10, 25, 300)	(0, 10, 965)	(10, 20, 300)	(0, 7, 129)	(10, 17, 167)
Subgraph probability	0.05	0.05	0.05	0.05	0.05	0.05
Sample size	500	500	500	500	500	500
Modularity (top)	0.8	0.686	0.781	0.739	0.789	0.663
Average node degree top	1.193	1.38	3.217	3.337	0.772	0.886
Avg connections within top communities	71.6	73.4	193	191.6	25.8	25.8
Avg. connections between top communities	0	2.35	0	2.15	0	0.95
Modularity (middle)	0.771	0.658	0.875	0.841	0.813	0.697
Average node degree middle	1.193	1.38	3.217	3.337	0.772	0.886
Avg connections within middle communities	20	20	61.333	61.333	9.667	9.667
Avg connections between middle communities	0.276	0.543	0.214	0.386	0.098	0.242

Tables

Table 2: Simulation settings for intermediate difficulty networks. Each row represents a single simulation scenario applied to all 6 simulated networks given in Table 1

Scenario	Input Graph	Graph Recon. Loss	Attr. Recon. Loss	Modularity Weigth	Clust. Weight
1	A_ingraph_true	1 = on	False (on)	1 = on	1 (middle), 1 (top)
2	A_corr_no_cutoff	1 = on	False (on)	1 = on	1 (middle), 1 (top)
3	A_ingraph02	1 = on	False (on)	1 = on	1 (middle), 1 (top)
4	A_ingraph05	1 = on	False (on)	1 = on	1 (middle), 1 (top)
5	A_ingraph07	1 = on	False (on)	1 = on	1 (middle), 1 (top)
6	A_ingraph_true	0 = off	False (on)	1 = on	1 (middle), 1 (top)
7	A_corr_no_cutoff	0 = off	False (on)	1 = on	1 (middle), 1 (top)
8	A_ingraph02	0 = off	False (on)	1 = on	1 (middle), 1 (top)
9	A_ingraph05	0 = off	False (on)	1 = on	1 (middle), 1 (top)
10	A_ingraph07	0 = off	False (on)	1 = on	1 (middle), 1 (top)
11	A_ingraph_true	1 = on	True (off)	1 = on	1 (middle), 1 (top)
12	A_corr_no_cutoff	1 = on	True (off)	1 = on	1 (middle), 1 (top)
13	A_ingraph02	1 = on	True (off)	1 = on	1 (middle), 1 (top)
14	A_ingraph05	1 = on	True (off)	1 = on	1 (middle), 1 (top)
15	A_ingraph07	1 = on	True (off)	1 = on	1 (middle), 1 (top)
16	A_ingraph_true	0 = off	True (off)	1 = on	1 (middle), 1 (top)
17	A_corr_no_cutoff	0 = off	True (off)	1 = on	1 (middle), 1 (top)
18	A_ingraph02	0 = off	True (off)	1 = on	1 (middle), 1 (top)
19	A_ingraph05	0 = off	True (off)	1 = on	1 (middle), 1 (top)
20	A_ingraph07	0 = off	True (off)	1 = on	1 (middle), 1 (top)
21	A_ingraph_true	1 = on	False (on)	0 = off	1 (middle), 1 (top)
22	A_corr_no_cutoff	1 = on	False (on)	0 = off	1 (middle), 1 (top)
23	A_ingraph02	1 = on	False (on)	0 = off	1 (middle), 1 (top)
24	A_ingraph05	1 = on	False (on)	0 = off	1 (middle), 1 (top)
25	A_ingraph07	1 = on	False (on)	0 = off	1 (middle), 1 (top)
26	A_ingraph_true	0 = off	False (on)	0 = off	1 (middle), 1 (top)
27	A_corr_no_cutoff	0 = off	False (on)	0 = off	1 (middle), 1 (top)
28	A_ingraph02	0 = off	False (on)	0 = off	1 (middle), 1 (top)
29	A_ingraph05	0 = off	False (on)	0 = off	1 (middle), 1 (top)
30	A_ingraph07	0 = off	False (on)	0 = off	1 (middle), 1 (top)
31	A_ingraph_true	1 = on	True (off)	0 = off	1 (middle), 1 (top)
32	A_corr_no_cutoff	1 = on	True (off)	0 = off	1 (middle), 1 (top)
33	A_ingraph02	1 = on	True (off)	0 = off	1 (middle), 1 (top)
34	A_ingraph05	1 = on	True (off)	0 = off	1 (middle), 1 (top)
35	A_ingraph07	1 = on	True (off)	0 = off	1 (middle), 1 (top)
36	A_ingraph_true	0 = off	True (off)	0 = off	1 (middle), 1 (top)
37	A_corr_no_cutoff	0 = off	True (off)	0 = off	1 (middle), 1 (top)
38	A_ingraph02	0 = off	True (off)	0 = off	1 (middle), 1 (top)
39	A_ingraph05	0 = off	True (off)	0 = off	1 (middle), 1 (top)
40	A_ingraph07	0 = off	True (off)	0 = off	1 (middle), 1 (top)
41	A_ingraph_true	1 = on	False (on)	1 = on	0.1 (middle), 1e-4 (top)
42	A_corr_no_cutoff	1 = on	False (on)	1 = on	0.1 (middle), 1e-4 (top)
43	A_ingraph02	1 = on	False (on)	1 = on	0.1 (middle), 1e-4 (top)
44	A_ingraph05	1 = on	False (on)	1 = on	0.1 (middle), 1e-4 (top)

Scenario	Input Graph	Graph Recon. Loss	Attr. Recon. Loss	Modularity Weigth	Clust. Weight
45	A_ingraph07	1 = on	False (on)	1 = on	0.1 (middle), 1e-4 (top)
46	A_ingraph_true	0 = off	False (on)	1 = on	0.1 (middle), 1e-4 (top)
47	A_corr_no_cutoff	0 = off	False (on)	1 = on	0.1 (middle), 1e-4 (top)
48	A_ingraph02	0 = off	False (on)	1 = on	0.1 (middle), 1e-4 (top)
49	A_ingraph05	0 = off	False (on)	1 = on	0.1 (middle), 1e-4 (top)
50	A_ingraph07	0 = off	False (on)	1 = on	0.1 (middle), 1e-4 (top)
51	A_ingraph_true	1 = on	True (off)	1 = on	0.1 (middle), 1e-4 (top)
52	A_corr_no_cutoff	1 = on	True (off)	1 = on	0.1 (middle), 1e-4 (top)
53	A_ingraph02	1 = on	True (off)	1 = on	0.1 (middle), 1e-4 (top)
54	A_ingraph05	1 = on	True (off)	1 = on	0.1 (middle), 1e-4 (top)
55	A_ingraph07	1 = on	True (off)	1 = on	0.1 (middle), 1e-4 (top)
56	A_ingraph_true	0 = off	True (off)	1 = on	0.1 (middle), 1e-4 (top)
57	A_corr_no_cutoff	0 = off	True (off)	1 = on	0.1 (middle), 1e-4 (top)
58	A_ingraph02	0 = off	True (off)	1 = on	0.1 (middle), 1e-4 (top)
59	A_ingraph05	0 = off	True (off)	1 = on	0.1 (middle), 1e-4 (top)
60	A_ingraph07	0 = off	True (off)	1 = on	0.1 (middle), 1e-4 (top)
61	A_ingraph_true	1 = on	False (on)	0 = off	0.1 (middle), 1e-4 (top)
62	A_corr_no_cutoff	1 = on	False (on)	0 = off	0.1 (middle), 1e-4 (top)
63	A_ingraph02	1 = on	False (on)	0 = off	0.1 (middle), 1e-4 (top)
64	A_ingraph05	1 = on	False (on)	0 = off	0.1 (middle), 1e-4 (top)
65	A_ingraph07	1 = on	False (on)	0 = off	0.1 (middle), 1e-4 (top)
66	A_ingraph_true	0 = off	False (on)	0 = off	0.1 (middle), 1e-4 (top)
67	A_corr_no_cutoff	0 = off	False (on)	0 = off	0.1 (middle), 1e-4 (top)
68	A_ingraph02	0 = off	False (on)	0 = off	0.1 (middle), 1e-4 (top)
69	A_ingraph05	0 = off	False (on)	0 = off	0.1 (middle), 1e-4 (top)
70	A_ingraph07	0 = off	False (on)	0 = off	0.1 (middle), 1e-4 (top)

Scenario	Input Graph	Graph Recon. Loss	Attr. Recon. Loss	Modularity Weigth	Clust. Weight
71	A_ingraph_true	1 = on	True (off)	0 = off	0.1 (middle), 1e-4 (top)
72	A_corr_no_cutoff	1 = on	True (off)	0 = off	0.1 (middle), 1e-4 (top)
73	A_ingraph02	1 = on	True (off)	0 = off	0.1 (middle), 1e-4 (top)
74	A_ingraph05	1 = on	True (off)	0 = off	0.1 (middle), 1e-4 (top)
75	A_ingraph07	1 = on	True (off)	0 = off	0.1 (middle), 1e-4 (top)
76	A_ingraph_true	0 = off	True (off)	0 = off	0.1 (middle), 1e-4 (top)
77	A_corr_no_cutoff	0 = off	True (off)	0 = off	0.1 (middle), 1e-4 (top)
78	A_ingraph02	0 = off	True (off)	0 = off	0.1 (middle), 1e-4 (top)
79	A_ingraph05	0 = off	True (off)	0 = off	0.1 (middle), 1e-4 (top)
80	A_ingraph07	0 = off	True (off)	0 = off	0.1 (middle), 1e-4 (top)

References

Barabási, Albert-László, and Eric Bonabeau. 2003. "Scale-Free Networks." *Scientific American* 288 (5): 60–69.

Watts, Duncan J, and Steven H Strogatz. 1998. "Collective Dynamics of 'Small-World'networks." *Nature* 393 (6684): 440–42.