Supplementary Material: SoftBlock: Transferable Soft Blocks for Overlapping Protein Complex Recovery under a Frozen Protocol

Haziq Jeelani (Claremont Graduate University)

A   Seed Variability

Table 1: Seed variability for SoftBlock under the frozen protocol (three seeds: 42/43/44). We report mean $\pm$ std across seeds for averaged best-match F1, OS Accuracy, and greedy MMR (each averaged over CORUM and Complex Portal).

Graph	#PC	BestMatch	Acc	MMR
STRING	8000 ± 0	0.534 ± 0.003	0.215 ± 0.000	0.290 ± 0.003
BioPlex	8000 ± 0	0.430 ± 0.007	0.189 ± 0.002	0.236 ± 0.003
HuRI	4832 ± 96	0.259 ± 0.005	0.105 ± 0.000	0.117 ± 0.001
BioGRID	4533 ± 26	0.341 ± 0.004	0.158 ± 0.003	0.175 ± 0.003
ComPPI	8000 ± 0	0.400 ± 0.006	0.222 ± 0.003	0.250 ± 0.002
IntAct	8000 ± 0	0.245 ± 0.002	0.168 ± 0.000	0.157 ± 0.001
hu.MAP2	2605 ± 39	0.530 ± 0.002	0.228 ± 0.002	0.196 ± 0.004

Table 2: Seed variability at the frozen operating point (cap $N = 2000$, overlap suppression with max Jaccard < 0.5). We report mean $\pm$ std across seeds for averaged OS Accuracy/MMR and averaged best-match F1.

Graph	#PC	BestMatch	Acc	MMR
STRING	2000 ± 0	0.444 ± 0.001	0.206 ± 0.001	0.162 ± 0.001
BioPlex	2000 ± 0	0.323 ± 0.003	0.161 ± 0.001	0.121 ± 0.001
HuRI	2000 ± 0	0.217 ± 0.011	0.108 ± 0.001	0.085 ± 0.001
BioGRID	2000 ± 0	0.304 ± 0.005	0.151 ± 0.004	0.125 ± 0.003
ComPPI	2000 ± 0	0.282 ± 0.004	0.185 ± 0.003	0.138 ± 0.002
IntAct	1900 ± 6	0.181 ± 0.006	0.137 ± 0.001	0.089 ± 0.001
hu.MAP2	1380 ± 15	0.486 ± 0.006	0.209 ± 0.001	0.144 ± 0.002

B   Full Baseline Landscape

Full baseline results on CORUM and Complex Portal are provided in machine-readable format in the code repository.

Table 3: One-sided significance checks over seeds (42/43/44). For each graph, we test whether SoftBlock improves averaged best-match F1 over the strongest baseline under the same table's protocol.

Protocol	Graph	Baseline	ΔBestMatch	p
Uncapped	STRING	Link communities	0.034 ± 0.003	0.000975
Uncapped	BioPlex	Link communities	0.048 ± 0.007	0.00372
Uncapped	HuRI	ClusterONE	0.141 ± 0.005	0.000226
Uncapped	BioGRID	Link communities	0.163 ± 0.004	0.000125
Uncapped	ComPPI	Link communities	0.039 ± 0.006	0.00415
Uncapped	IntAct	Link communities	0.032 ± 0.002	0.000601
Uncapped	hu.MAP2	Link communities	0.094 ± 0.002	9.04e-05
Cap $N = 2000$	STRING	Link communities	0.028 ± 0.001	0.000374
Cap $N = 2000$	BioPlex	Link communities	0.020 ± 0.003	0.00359
Cap $N = 2000$	HuRI	Link communities	0.200 ± 0.011	0.000484
Cap $N = 2000$	BioGRID	Link communities	0.153 ± 0.005	0.000185
Cap $N = 2000$	ComPPI	Link communities	0.029 ± 0.004	0.00375
Cap $N = 2000$	IntAct	Link communities	0.048 ± 0.006	0.00266
Cap $N = 2000$	hu.MAP2	Link communities	0.060 ± 0.006	0.00167

Table 4: Method-level averages and mean ranks across graphs. Ranks are computed per graph using the average of CORUM and Complex Portal mean best-match F1 ("BestMatch rank") and the average of CORUM and Complex Portal OS Accuracy ("Acc rank"); lower is better.

Method	BM rank	Acc rank	$F1_C$	$F1_{CP}$	Acc_C	Acc_{CP}
SoftBlock+LinkComm (auto)	1.14	1.71	0.360	0.425	0.182	0.187
SoftBlock (multi-K)	2.29	3.00	0.308	0.400	0.137	0.160
SoftBlock (posthoc-OS)	3.14	1.71	0.315	0.370	0.177	0.182
Link communities	3.43	3.57	0.281	0.329	0.163	0.157
ClusterONE	5.00	5.71	0.174	0.223	0.082	0.099
GreedyDense	6.00	5.57	0.142	0.176	0.084	0.090
OSLOM2	7.29	7.00	0.102	0.136	0.058	0.080
SLPA	7.71	7.71	0.071	0.105	0.034	0.049
BigCLAM	8.71	8.71	0.006	0.007	0.008	0.008

C Cap Capability at $N = 2000$

D BioPlex Operating-Point Diagnosis

E Pipeline Extras Ablation

F Operating-Point Sensitivity

G Pipeline Diagram

The SoftBlock pipeline consists of the following stages:

1. **Reference PPI** (G_{ref}, STRING) $\rightarrow$ SoftBlock soft memberships $R \in [0,1]^{|V_{\text{ref}}| \times K}$

2. **Target PPI** (G: BioPlex / HuRI / ...) $\rightarrow$ Overlapping blocks $\{B_k\}_{k=1}^K$

Table 5: Cap capability under the frozen operating-point protocol ($N = 2000$, overlap suppression with max Jaccard < 0.5). We report the uncapped candidate pool size (#PC pool) and the selected count (#PC out). "Cap-limited" indicates #PC pool $< N$; "NMS-limited" indicates #PC pool $\geq N$ but #PC out $< N$ due to overlap suppression.

Graph	Method	#PC pool	#PC out	Status
BioGRID	SoftBlock	4560	2000	Hits cap
BioGRID	Link communities	7940	2000	Hits cap
BioGRID	OSLOM2	144	144	Cap-limited
BioPlex	SoftBlock	8000	2000	Hits cap
BioPlex	Link communities	7838	2000	Hits cap
BioPlex	OSLOM2	342	342	Cap-limited
ComPPI	SoftBlock	8000	2000	Hits cap
ComPPI	Link communities	7984	1976	NMS-limited
ComPPI	OSLOM2	266	266	Cap-limited
HuRI	SoftBlock	4733	2000	Hits cap
HuRI	Link communities	1781	417	Cap-limited
HuRI	OSLOM2	55	55	Cap-limited
IntAct	SoftBlock	8000	1905	NMS-limited
IntAct	Link communities	7974	1681	NMS-limited
IntAct	OSLOM2	188	188	Cap-limited
hu.MAP2	SoftBlock	2641	1385	NMS-limited
hu.MAP2	Link communities	1227	836	Cap-limited
hu.MAP2	OSLOM2	311	311	Cap-limited
STRING	SoftBlock	8000	2000	Hits cap
STRING	Link communities	7983	2000	Hits cap
STRING	OSLOM2	464	464	Cap-limited

3. Local solver in each block (MCL / wrappers)

4. Candidate pool (union + dedupe)

5. Graph-only scoring (+ optional learned rerank)

6. Top-N selection + overlap suppression $\rightarrow$ Predicted complexes

H Hybrid Auto-Gate Triggers

I Runtime and Scalability

J Biological Plausibility: GO-Slim Enrichment Case Studies

To complement benchmark scores, we perform a conservative functional/localization coherence check on STRING predictions. We select top-ranked predicted clusters (weighted-density score) that are unmatched under our CORUM/Complex Portal overlap-score criterion (max OS < 0.2) and restrict to size ≤ 50 with at least three GO-annotated members.

Table 6: BioPlex operating-point diagnosis ($N = 2000$, density ranking, overlap suppression with max Jaccard < 0.5). Under blocks-only selection, density on an unweighted AP-MS graph over-selects size-3 cliques (83.7%), fragmenting large complexes. The official `hybrid_auto` recipe unions filtered link-communities candidates on unweighted graphs (frozen rule), reducing the size-3 fraction to 44.6%.

Method	Acc	MMR	F1	%($|C|$=3)	Large recall
SoftBlock (multi-K; blocks-only)	0.125	0.103	0.256	83.7	0.481 / 0.618
SoftBlock (hybrid_auto)	0.161	0.121	0.322	44.6	0.815 / 0.971
Link communities	0.157	0.114	0.303	57.0	0.778 / 0.971

Table 7: Pipeline-extras ablation (uncapped). We report averaged best-match F1, OS Accuracy, and greedy MMR to isolate the contribution of (i) the transferable SoftBlock blocks (core), (ii) graph-only reranking, and (iii) hybrid union with link communities.

Graph	Variant	#PC	F1	Acc	MMR
STRING	Link communities (baseline)	7983	0.500	0.202	0.267
STRING	SoftBlock blocks+MCL (core)	9117	0.433	0.135	0.208
STRING	SoftBlock blocks+MCL (+rerank)	8000	0.421	0.139	0.203
STRING	SoftBlock (hybrid_auto)	8000	0.533	0.215	0.292
BioPlex	Link communities (baseline)	7838	0.382	0.173	0.193
BioPlex	SoftBlock blocks+MCL (core)	6426	0.405	0.159	0.205
BioPlex	SoftBlock (hybrid_auto)	8000	0.431	0.191	0.239
HuRI	Link communities (baseline)	1781	0.063	0.038	0.025
HuRI	SoftBlock blocks+MCL (core)	4196	0.265	0.110	0.116
HuRI	SoftBlock (hybrid_auto)	4733	0.265	0.105	0.117
hu.MAP2	Link communities (baseline)	1227	0.437	0.188	0.113
hu.MAP2	SoftBlock blocks+MCL (core)	1533	0.485	0.234	0.159
hu.MAP2	SoftBlock (hybrid_auto)	2641	0.532	0.230	0.200

K Node Overlap with STRING

Table 8: Sensitivity of the frozen operating point to the cap N and overlap-suppression threshold. We summarize the rank of SoftBlock among all evaluated methods under OS Accuracy and greedy MMR across the sweep settings.

Graph	Sweep	#settings	Top1 Acc	Med rank	Worst	Top1 MMR	Worst
BioGRID	cap sweep	4	4	1.0	1	4	1
BioGRID	Jaccard sweep	4	3	1.0	2	4	1
HuRI	cap sweep	4	0	2.0	2	1	2
HuRI	Jaccard sweep	4	0	2.0	2	1	2

Table 9: Hybrid auto-gate triggers. The dataset-agnostic `auto_confidence_weighted` policy unions link-communities candidates on unweighted graphs (with a frozen min-size filter) and on confidence-like weighted graphs (scaled weights in $[0, 1]$), otherwise falling back to SoftBlock true-multiK.

Graph	Weighted	Gate uses linkcomm	Reason
BioGRID	yes	no	weights not in $[0, 1]$
BioPlex	no	yes	unweighted (min $|C| \geq 4$ filter)
ComPPI	yes	yes	confidence-like weights in $[0, 1]$
HuRI	no	yes	unweighted (min $|C| \geq 4$ filter)
IntAct	yes	yes	confidence-like (all weights $= 1$)
hu.MAP2	yes	yes	confidence-like weights in $[0, 1]$
STRING	yes	yes	confidence-like weights in $[0, 1]$

Table 10: Runtime and memory footprint (peak RSS) for SoftBlock (`hybrid_auto`) and the strongest baseline per graph. We report wall-clock time in minutes for the main pipeline stages and the total, along with peak resident memory (GiB). Timings exclude CORUM/ComplexPortal evaluation.

Graph	Nodes	Edges	MCL	LC	Ens	Hyb	Total	Peak GiB
BioGRID	27590	1002631	2.25	0.00	0.13	0.08	2.45	2.76
IntAct	17733	527860	1.42	0.24	0.09	0.24	1.99	1.91
STRING	15882	236712	0.86	0.20	0.22	0.26	1.54	0.95
ComPPI	15277	170728	1.03	0.17	0.17	0.36	1.73	2.10
BioPlex	13923	118144	0.53	0.12	0.10	0.16	0.90	0.78
HuRI	8109	51686	0.30	0.06	0.06	0.08	0.51	0.75
hu.MAP2	7824	19631	0.32	0.06	0.04	0.05	0.47	0.74

Table 11: GO-slim enrichment for representative high-scoring predictions on STRING that are unmatched under our CORUM/Complex Portal overlap-score criterion (max OS < 0.2; size ≤ 50). We report the most significant GO-slim cellular-component (CC) and biological-process (BP) terms (BH-corrected q).

Rank	$|C|$	max OS	Top CC term	Top BP term
1	31	0.073	nucleosome ($q = 3.0 \times 10^{-20}$)	heterochromatin formation
2	30	0.033	nucleosome ($q = 1.1 \times 10^{-9}$)	nucleosome assembly
3	30	0.188	ribosome ($q = 3.8 \times 10^{-38}$)	mitochondrial translation
4	29	0.034	nucleosome ($q = 5.0 \times 10^{-19}$)	nucleosome assembly
5	29	0.046	nucleosome ($q = 2.5 \times 10^{-12}$)	heterochromatin formation
6	28	0.143	nucleosome ($q = 1.2 \times 10^{-11}$)	chromatin organization
7	26	0.158	ribosome ($q = 2.7 \times 10^{-37}$)	translation
8	26	0.038	nucleosome ($q = 3.9 \times 10^{-6}$)	telomere organization
9	25	0.164	cytosolic ribosome ($q = 6.9 \times 10^{-34}$)	cytoplasmic translation
10	24	0.150	nucleolus ($q = 4.8 \times 10^{-24}$)	maturation of SSU-rRNA

Table 12: Node overlap with the reference graph (STRING) and frozen-protocol performance. Overlap is $|V \cap V_{\mathrm{ref}}|/|V|$ where V_{ref} are STRING nodes. Performance is shown for SoftBlock (`hybrid_auto`) using mean best-match F1 and mean greedy MMR (averaged over CORUM and Complex Portal).

Graph	Nodes	Overlap	Mean F1	Mean MMR
BioGRID	27590	53.2%	0.336	0.172
BioPlex	13923	79.0%	0.431	0.239
IntAct	17733	82.7%	0.246	0.158
HuRI	8109	83.6%	0.265	0.117
hu.MAP2	7824	84.6%	0.532	0.200
ComPPI	15277	86.8%	0.405	0.251
STRING	15882	100.0%	0.533	0.292