A Practical, Progressively-Expressive GNN

Lingxiao Zhao[1], Louis Hartel[2], Neil Shah[3], Leman Akoglu[1]

[1]Carnegie Mellon University [2]RWTH Aachen University [3]Snap Inc.

Paper here!

Overview

Highly expressive GNNs becomes a hot topic:

- Node Identifier, Random Feature (Subgraph counts, RNI, LE)
- Subgraph-Enhanced GNNs (GNN-AK, NestedGNN, ESAN, …)
- Higher Order GNNs (PPGN, K-IGN, K-GNN , TokenGT …)
- Others (random-walk, individualization, drop node, …)

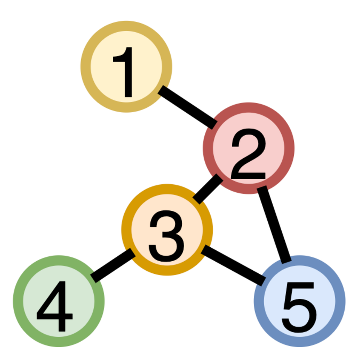
Should we chase expressivity?

How to study the impact of expressivity systematically?

We need a **practical**, **progressively** expressive GNN!

Contribution: a highly expressive GNN with desired properties:

- Practical - Scales significantly better than k-WL
- Theoretically Powerful - Retains theoretical connection to k-WL
- Progressively Expressive - Fine-grained "ruler" of expressivity
- New SOTA – Achieves SOTA on ZINC



Background: k-WL

1-WL:

1-tuples

$$c^{(t+1)}(v) \leftarrow \text{HASH}\big(c^{(t)}(v), \{\!\{c^{(t)}(u) \mid u \in \mathcal{N}_v\}\!\}\big)$$

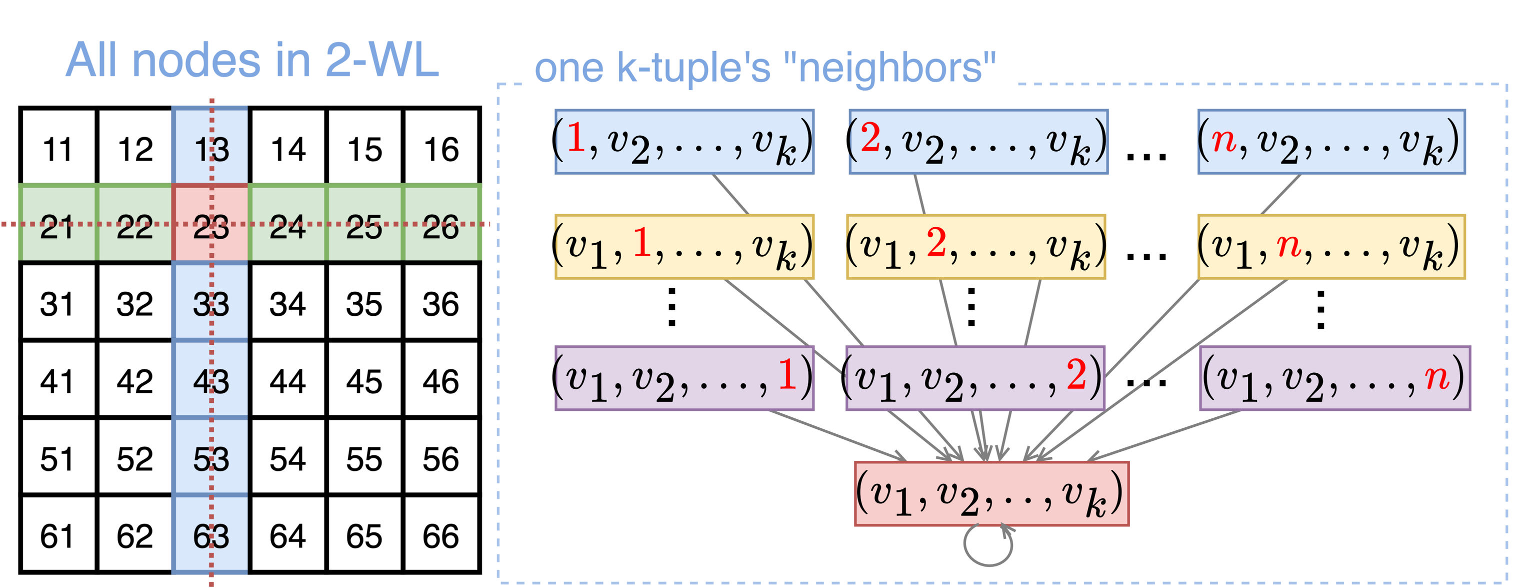
When $v = 2$: ② {{①, ③, ⑤}}

k-WL: (k=2)

k-tuples

When $u = 2, v = 3$

$$c^{(t+1)}\{(u,v)\} \leftarrow \text{HASH}\Big(c^{(t)}\{(u,v)\}, \{\!\{c^{(t)}\{(i,v)\} \mid i \in V(G)\}\!\}, \{\!\{c^{(t)}\{(u,i)\} \mid i \in V(G)\}\!\}\Big)$$



Impractical for k>3! $O(kn^{k+1})$ for 1 step

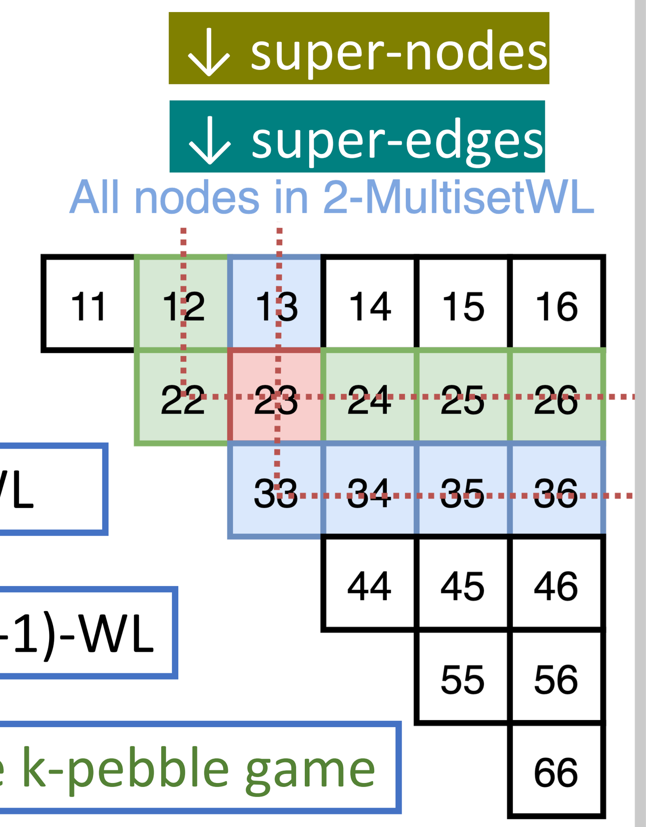
Scale k-WL

From tuples to multisets

↓ super-nodes ↓ super-edges

Remove ordering information

- # Super-nodes: $n^k \rightarrow \binom{n+k-1}{k}$ ratio≈ k!
- # Super-edges: $kn^{k+1} \rightarrow\approx n^2\binom{n+k-3}{k-1}$



Thm. 1: Upper-bounded by k-WL

Thm. 2: No less powerful than (k-1)-WL

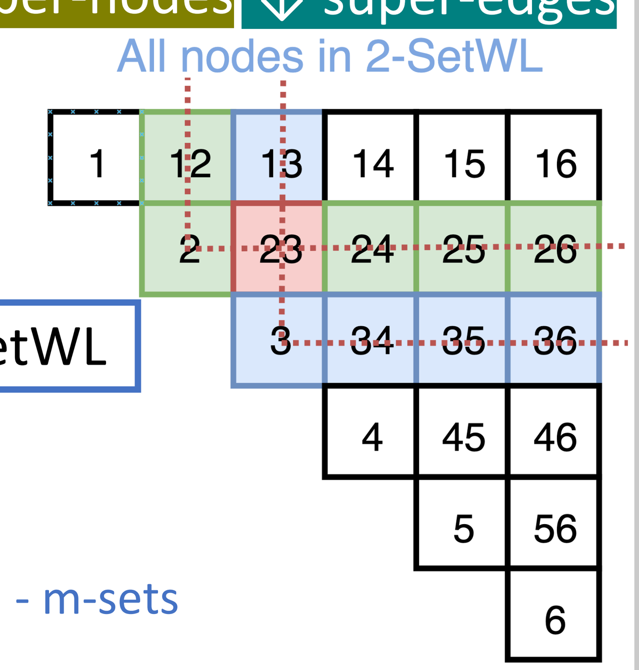
Thm. 3: Same power as doubly bijective k-pebble game

From multisets to sets

↓ super-nodes ↓ super-edges

Remove repetitions

- # Super-nodes: $\binom{n+k-1}{k} \rightarrow \sum_{i=1}^{k}\binom{n}{i}$
- # Super-edges: $n^2\binom{n+k-3}{k-1} \rightarrow \sum_{i=2}^{k} i\binom{n}{i}$



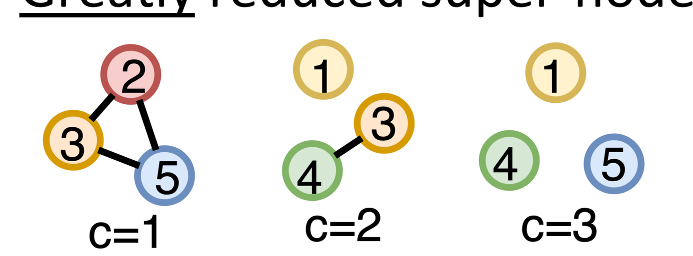
Thm. 4: Upper-bounded by k-MultisetWL

- Super-nodes: m-sets with 1≤m≤k
- For each m-set, its neighbors include:
  - (m-1)-sets - (m+1)-sets - m-sets

From full sets to restricted sets

↓ super-nodes ↓ super-edges

- Only include sets whose induced subgraph having ≤c connected components
- Greatly reduced super-nodes&-edges for sparse graph

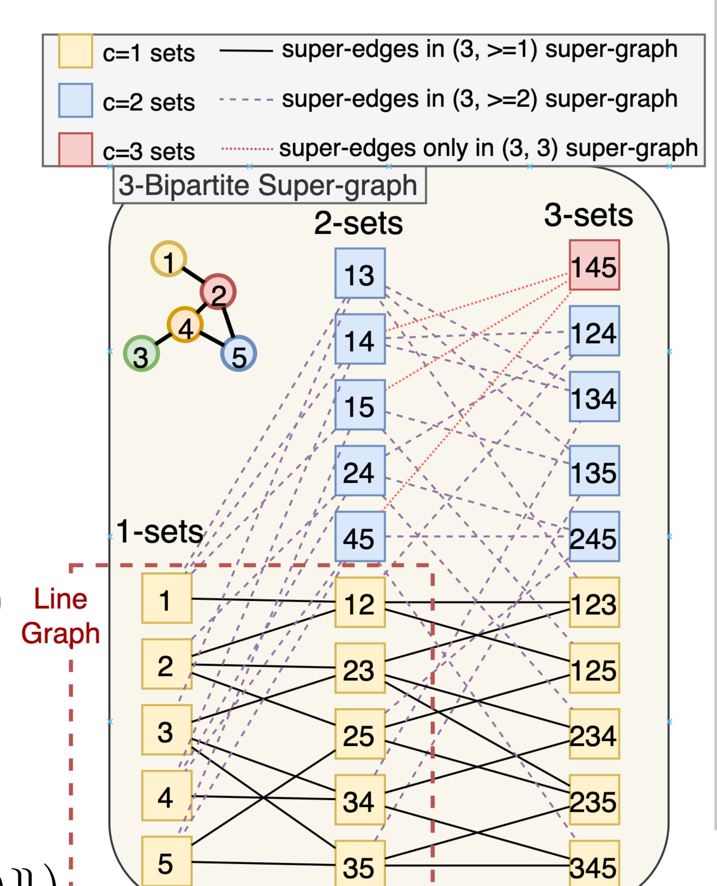


Thm. 5: Progressively Expressive with ↑k and ↑c

K-bipartite super-graph

↓ super-edges

- Introduced bidirectional parallel propagation on super-graph
- Connections from m-sets to m-sets can be safely ignored.



Parallel Backward Propagation

$$c^{(t+\frac{1}{2})}(\hat v) \leftarrow \text{HASH}\big(\{\!\{c^{(t)}(\hat u) \mid \hat u \in \mathcal{N}_{right}(\hat v)\}\!\}\big)$$

Parallel Forward Propagation

$$c^{(t+1)}(\hat v) \leftarrow \text{HASH}\big(c^{(t)}(\hat v),\ c^{(t+\frac{1}{2})}(\hat v), \{\!\{c^{(t)}(\hat u) \mid \hat u \in \mathcal{N}_{left}(\hat v)\}\!\}, \{\!\{c^{(t+\frac{1}{2})}(\hat u) \mid \hat u \in \mathcal{N}_{left}(\hat v)\}\!\}\big)$$

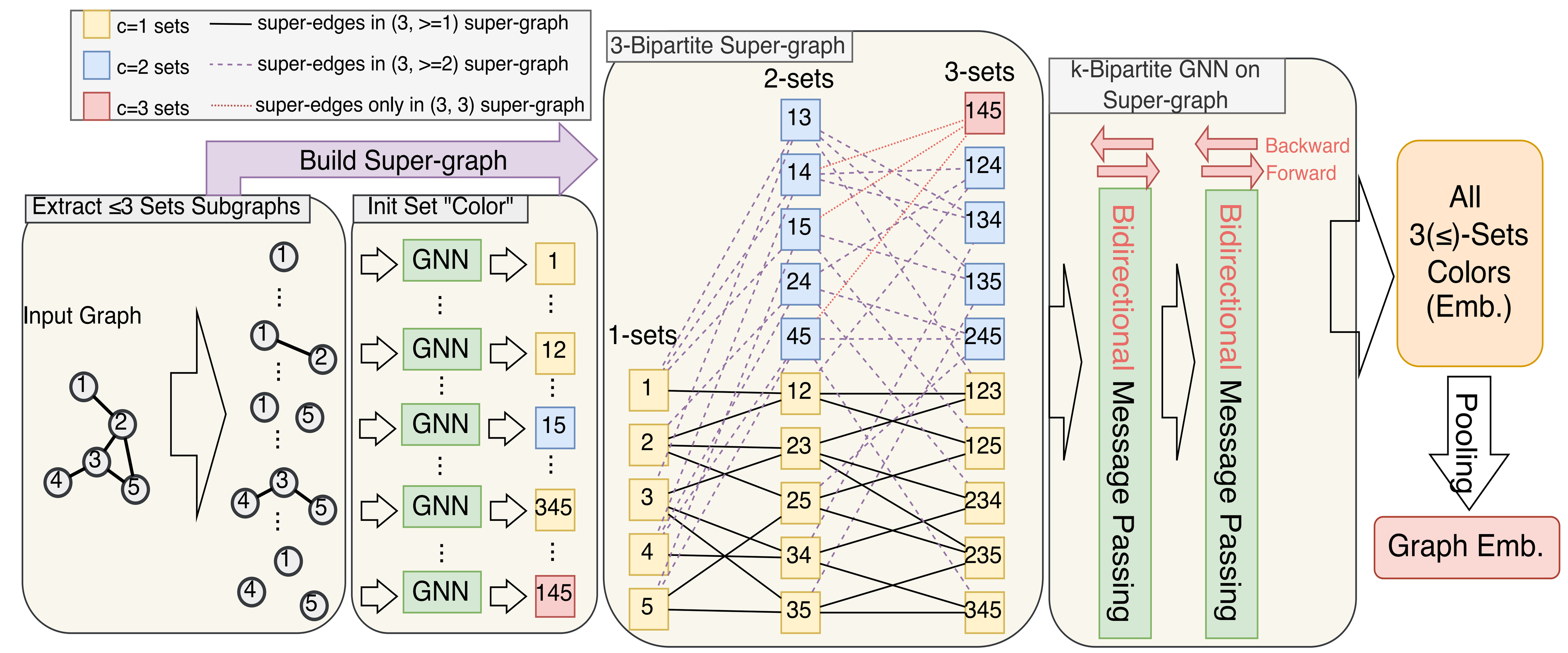
Final Model: (k,c)(≤)-SetGNN*

- Phase 1 - "Color" Initialization: encoding isomorphism type with a BaseGNN
- Phase 2 - Bidirectional Sequential Message Passing on k-bipartite Super-graph

$$m = k-1 \text{ to } 1, \forall m\text{-set } \hat v,\ h^{(t+\frac{1}{2})}(\hat v) = \text{MLP}^{(t)}_{m,1}\Big(h^{(t)}(\hat v), \sum_{\hat u \in \mathcal{N}^G_{\text{right}}(\hat v)} \text{MLP}^{(t)}_{m,2}\big(h^{(t+\frac{1}{2})}(\hat u)\big)\Big)$$ **Sequential Backward**

$$m = 2 \text{ to } k, \forall m\text{-set } \hat v,\ h^{(t+1)}(\hat v) = \text{MLP}^{(t+\frac{1}{2})}_{m,1}\Big(h^{(t+\frac{1}{2})}(\hat v), \sum_{\hat u \in \mathcal{N}^G_{\text{left}}(\hat v)} \text{MLP}^{(t+\frac{1}{2})}_{m,2}\big(h^{(t+1)}(\hat u)\big)\Big)$$ **Sequential Forward**

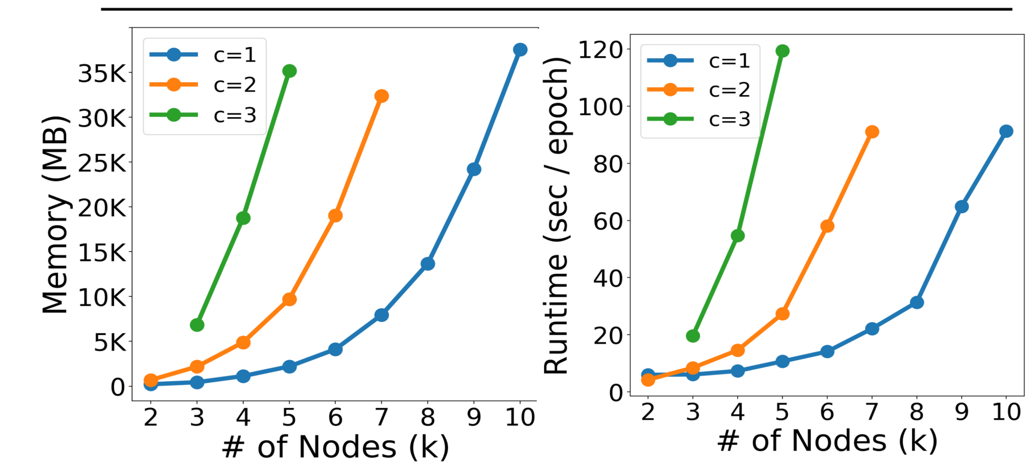
Thm. 7: t-layer Sequential MP is more expressive than t-layer Parallel MP



Experiments

ZINC-12k

k	c	Train loss	Val. MAE	Test MAE
2	1	0.1381 ± 0.0240	0.2429 ± 0.0071	0.2345 ± 0.0131
3	1	0.1172 ± 0.0063	0.2298 ± 0.0060	0.2252 ± 0.0030
4	1	0.0693 ± 0.0111	0.1645 ± 0.0052	0.1636 ± 0.0052
5	1	0.0643 ± 0.0019	0.1593 ± 0.0051	0.1447 ± 0.0013
6	1	0.0519 ± 0.0064	0.0994 ± 0.0093	0.0843 ± 0.0048
7	1	0.0543 ± 0.0048	0.0965 ± 0.0061	0.0747 ± 0.0022
8	1	0.0564 ± 0.0152	0.0961 ± 0.0043	0.0732 ± 0.0037
9	1	0.0817 ± 0.0274	0.0909 ± 0.0094	0.0824 ± 0.0056
10	1	0.0894 ± 0.0266	0.1060 ± 0.0157	0.0950 ± 0.0102
2	2	0.1783 ± 0.0602	0.2913 ± 0.0102	0.2948 ± 0.0210
3	2	0.0640 ± 0.0072	0.1668 ± 0.0078	0.1391 ± 0.0102
4	2	0.0499 ± 0.0043	0.1029 ± 0.0033	0.0836 ± 0.0010
5	2	0.0483 ± 0.0017	0.0899 ± 0.0056	**0.0750 ± 0.0027**
6	2	0.0530 ± 0.0064	0.0927 ± 0.0050	0.0737 ± 0.0006
7	2	0.0547 ± 0.0036	0.0984 ± 0.0047	0.0784 ± 0.0043
3	3	0.0798 ± 0.0062	0.1881 ± 0.0076	0.1722 ± 0.0086
4	3	0.0565 ± 0.0059	0.1121 ± 0.0066	0.0869 ± 0.0026
5	3	0.0671 ± 0.0156	0.1091 ± 0.0097	0.0920 ± 0.0054



Substructure Count

Method	EXP (ACC)	SR25 (ACC)	Counting Substructures (MAE)			
			Triangle	Tailed Tri.	Star	4-Cycle
GCN	50%	6.67%	0.4186	0.3248	0.1798	0.2822
GIN	50%	6.67%	0.3569	0.2373	0.0224	0.2185
PNA*	50%	6.67%	0.3532	0.2648	0.1278	0.2430
PPGN	**100%**	6.67%	0.0089	0.0096	0.0148	0.0090
GIN-AK+	**100%**	6.67%	0.0123	0.0112	0.0150	0.0126
PNA*-AK+	**100%**	6.67%	0.0118	0.0138	0.0166	0.0132
$(k,c)(\leq)$	**100%** (≥3, ≥2)	**100%** (≥4, ≥1)	**0.0073** (3, 2)	**0.0075** (4, 1)	**0.0134** (3, 2)	**0.0075** (4, 1)

QM 9

k	c	Train loss	Val. MAE	Test MAE
2	1	0.0376 ± 0.0005	0.0387 ± 0.0007	0.0389 ± 0.0008
3	1	0.0308 ± 0.0010	0.0386 ± 0.0017	0.0379 ± 0.0010
4	1	0.0338 ± 0.0003	0.0371 ± 0.0005	0.0370 ± 0.0006
5	1	0.0299 ± 0.0017	0.0343 ± 0.0008	0.0341 ± 0.0009
6	1	0.0226 ± 0.0004	0.0296 ± 0.0007	0.0293 ± 0.0007
7	1	0.0208 ± 0.0005	0.0289 ± 0.0007	0.0269 ± 0.0003
2	2	0.0367 ± 0.0007	0.0398 ± 0.0004	0.0398 ± 0.0004
3	2	0.0282 ± 0.0013	0.0358 ± 0.0009	0.0356 ± 0.0007
4	2	0.0219 ± 0.0004	0.0280 ± 0.0008	0.0278 ± 0.0008
5	2	0.0175 ± 0.0003	0.0267± 0.0005	**0.0251± 0.0006**
3	3	0.0391 ± 0.0107	0.0428 ± 0.0057	0.0425 ± 0.0052
4	3	0.0219 ± 0.0011	0.0301 ± 0.0010	0.0286 ± 0.0004
GINE (L=4)		0.0507 ± 0.0014	0.0478 ± 0.0003	0.0479 ± 0.0004
GINE (L=6)		0.0440 ± 0.0009	0.0440 ± 0.0009	0.0451 ± 0.0009