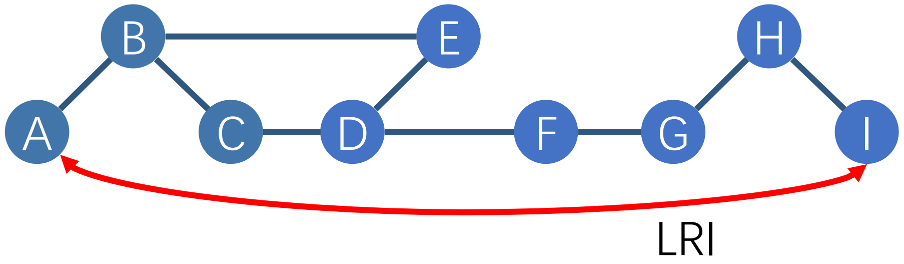
MeGraph: Capturing Long-Range Interactions by Alternating Local and Hierarchical Aggregation on Multi-Scaled Graph Hierarchy

Honghua Dong*, Jiawei Xu*, Yu Yang*, Rui Zhao, Shiwen Wu, Chun Yuan, Xiu Li, Chris J. Maddison, Lei Han

Project page: https://sites.google.com/view/megraph

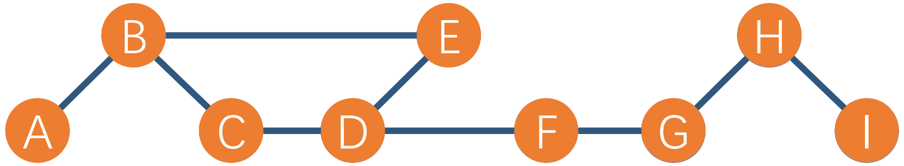
Challenge: Long-Range Interactions

- Local-aware message passing Graph Neural Networks (GNNs)
  - Struggle to capture Long-Range Interactions (LRIs)



The receptive field of **A** after 4 layers:

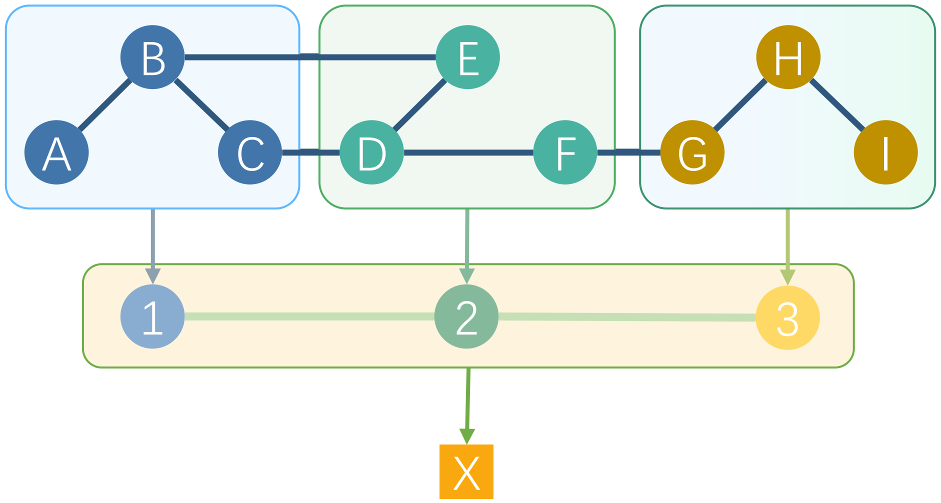
This number could be $\boldsymbol{O}(|\mathcal{V}|)$



Graph Hierarchy

- Try to capture LRIs over the hierarchy
  - Build the hierarchy using (learnable) **graph pooling** methods

Select
Reduce
Connect

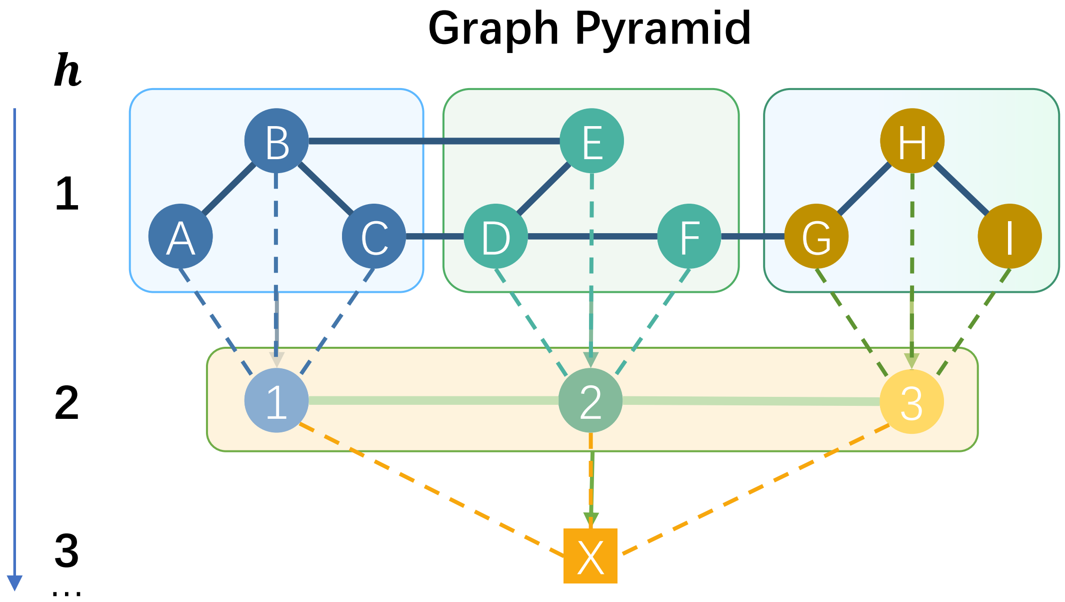


Pool:
Select &
Reduce &
Connect

- Path travelling the hierarchy is of length $O(log\,|\mathcal{V}|) \ll O(|\mathcal{V}|)$
- But, what's the cost? **Local Structure!**

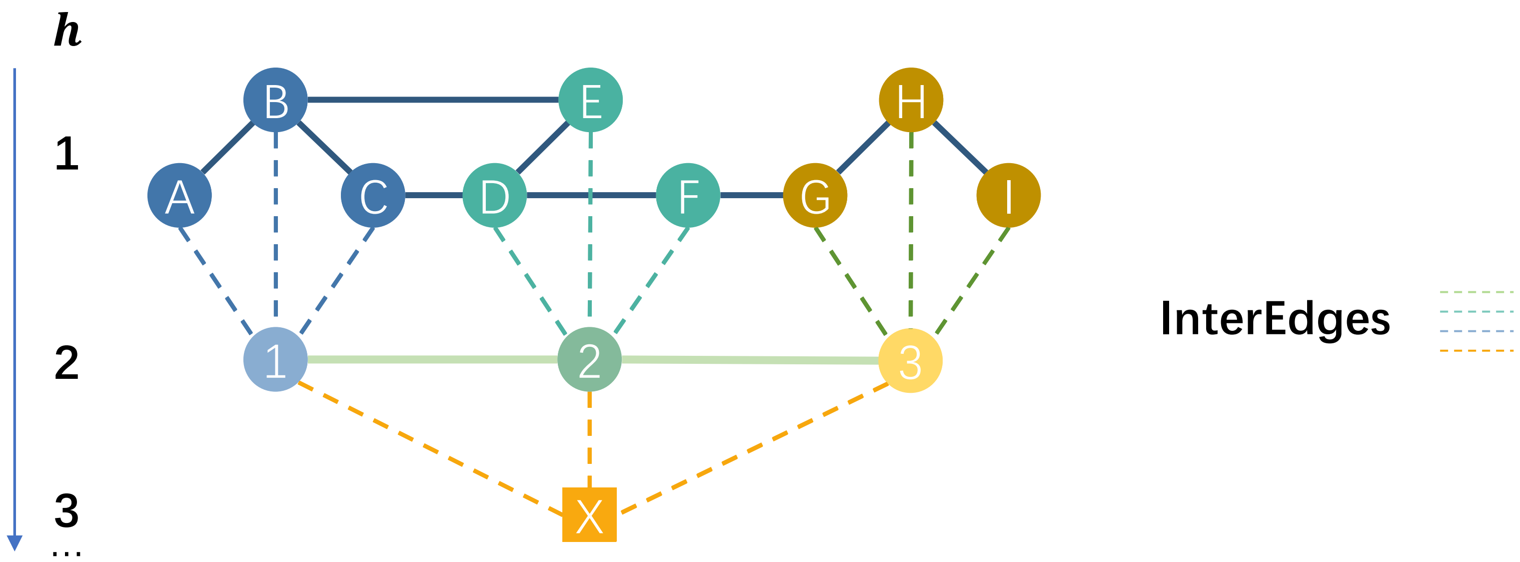
Connect the Graph Hierarchy

Let's connect the nodes before and after pooling



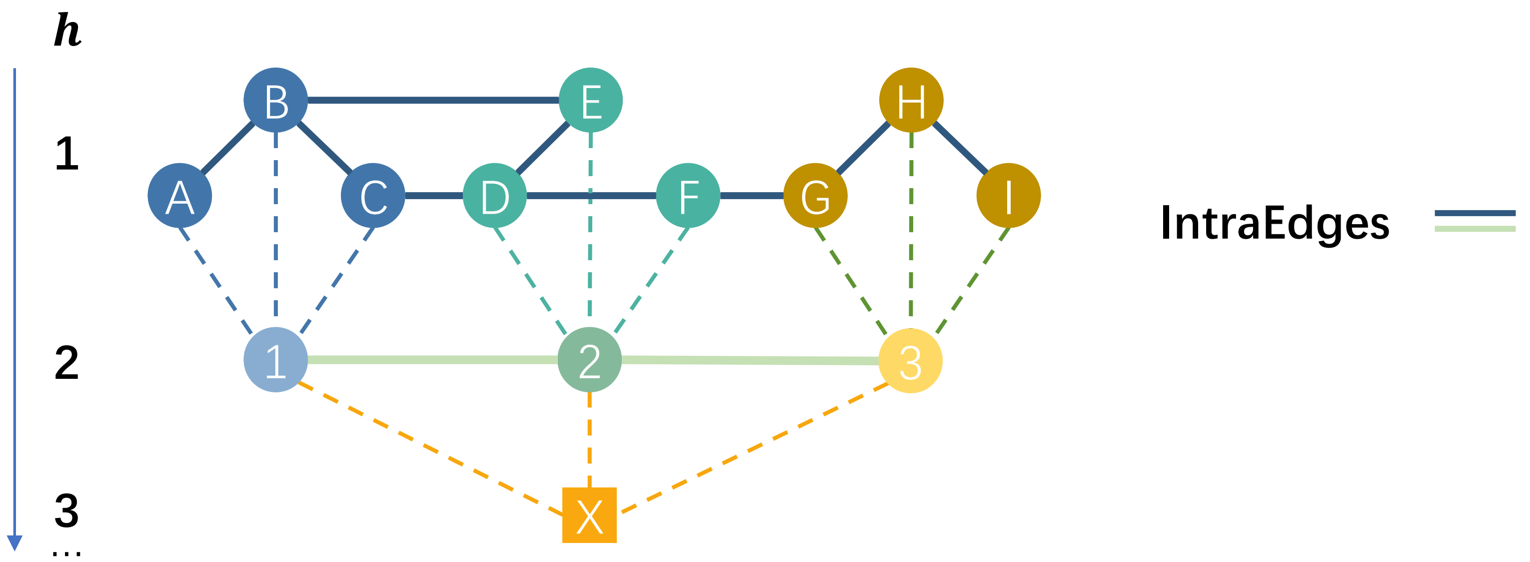
Connect the Graph Hierarchy

We call such edges as inter-graph edges (InterEdges),



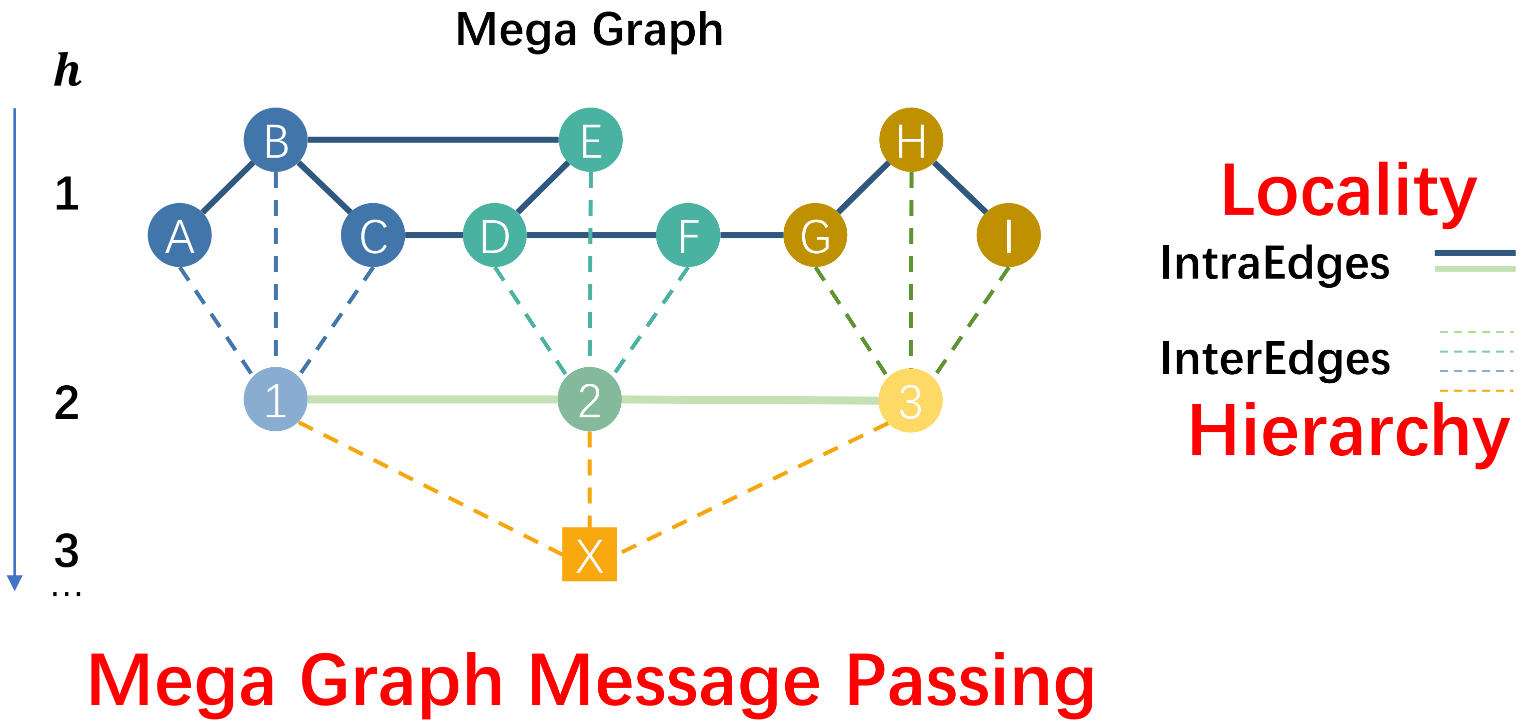
Connect the Graph Hierarchy

… edges in graphs of each level as intra-graph edges (IntraEdges)



The Mega Graph

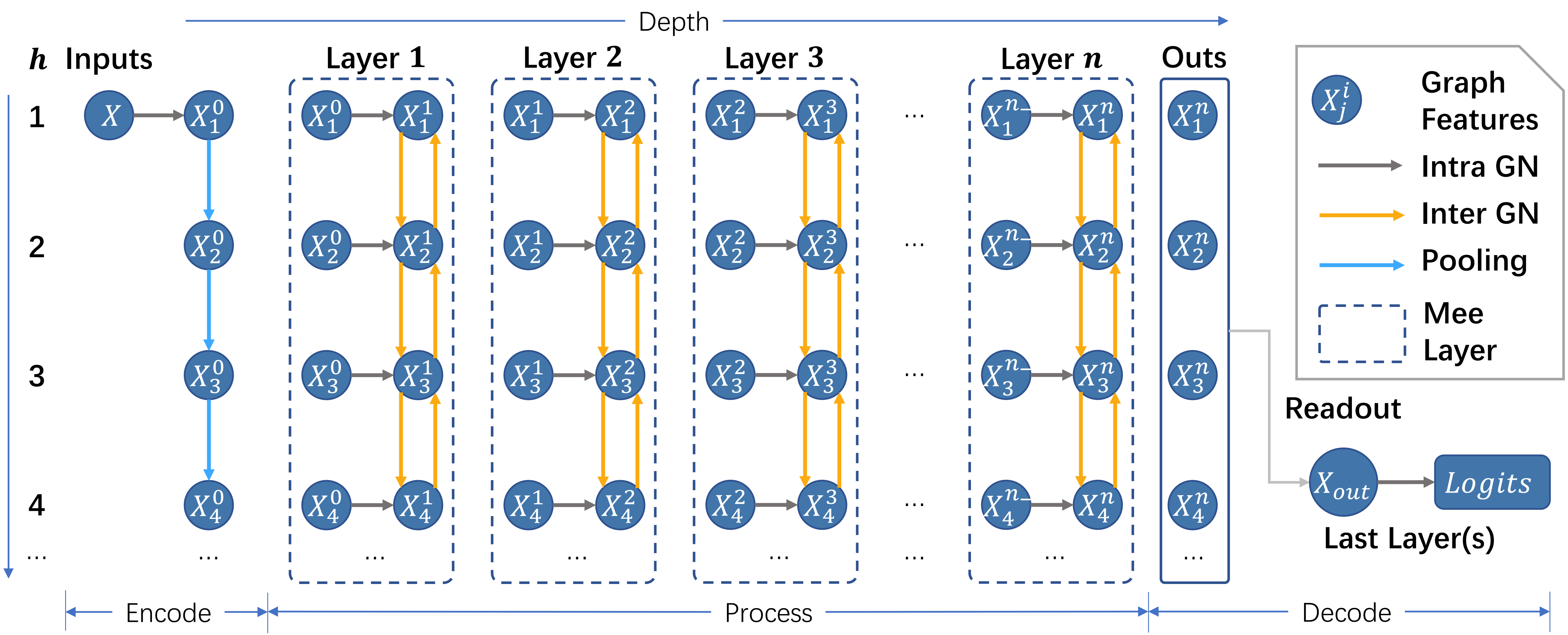
IntraEdges and InterEdges connects nodes into a mega graph.



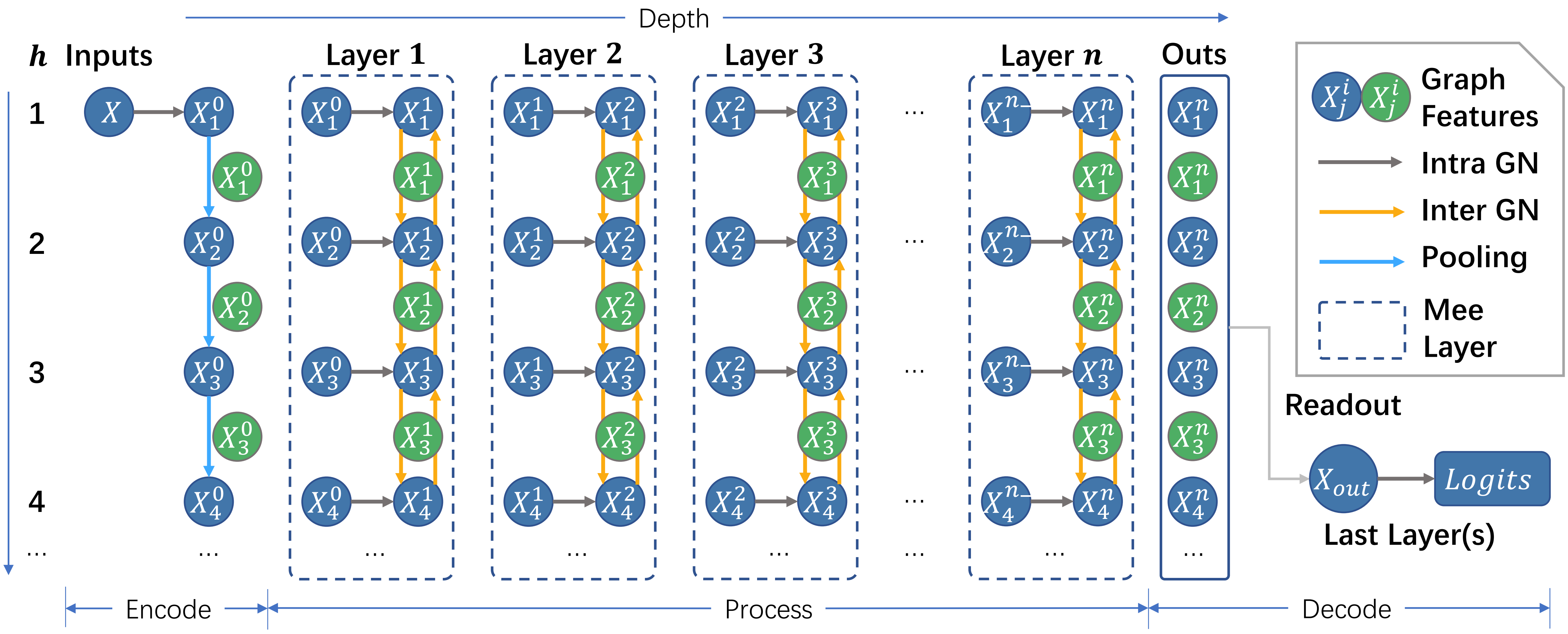
Basic Notations

- **Graph** $\mathcal{G} = (\mathcal{V}, \mathcal{E})$: A graph contains nodes $\mathcal{V}$ and edges $\mathcal{E}$
- **Graph features** $X = (u, V, E)$
  - u : global (graph-level) features
  - V : features of nodes $\mathcal{V}$
  - E : features of edges $\mathcal{E}$
- **GNN Blocks (GN)**: (i.e. GNN layer)
  - **Input:** graph $\mathcal{G}$ and its features X, **output:** new features X'
- **Graph pooling**:
  - **input:** graph $\mathcal{G}$ and its features X
  - **output:** The pooled graph $\tilde{\mathcal{G}}$ and its features $\tilde{X}$
  - **output:** The inter-graph $\hat{\mathcal{G}}$ and its features $\hat{X}$

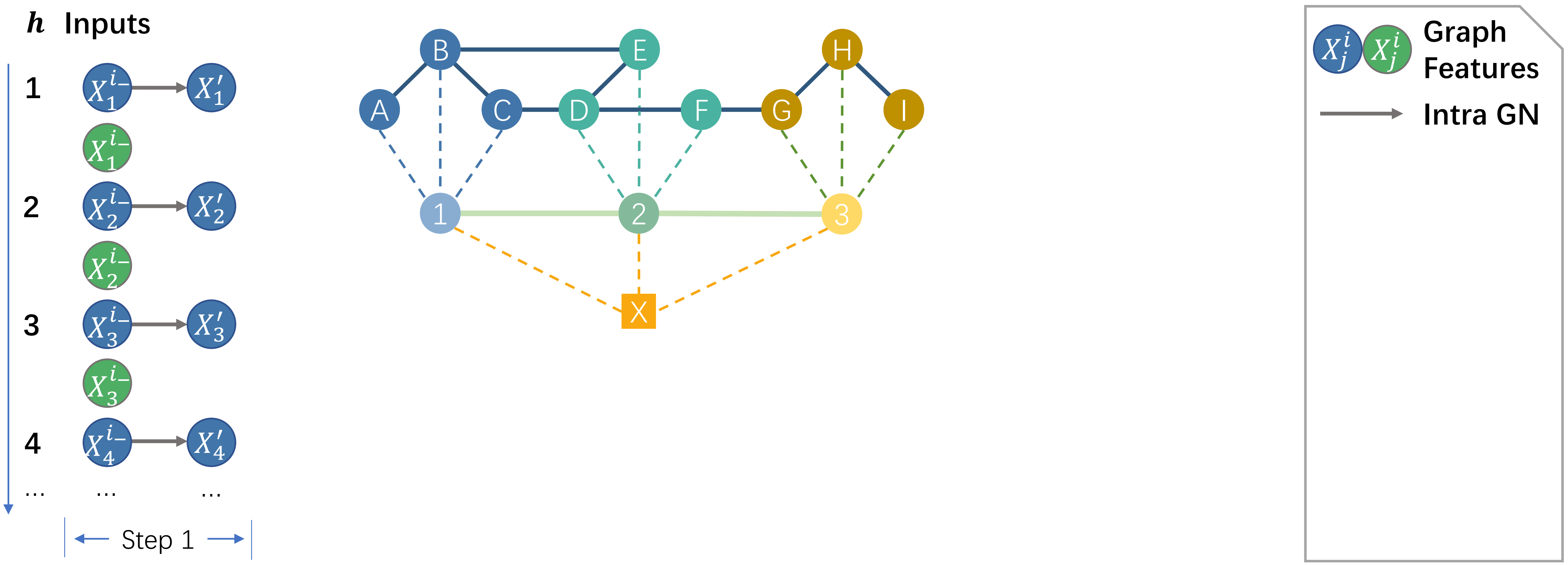
MeGraph Architecture



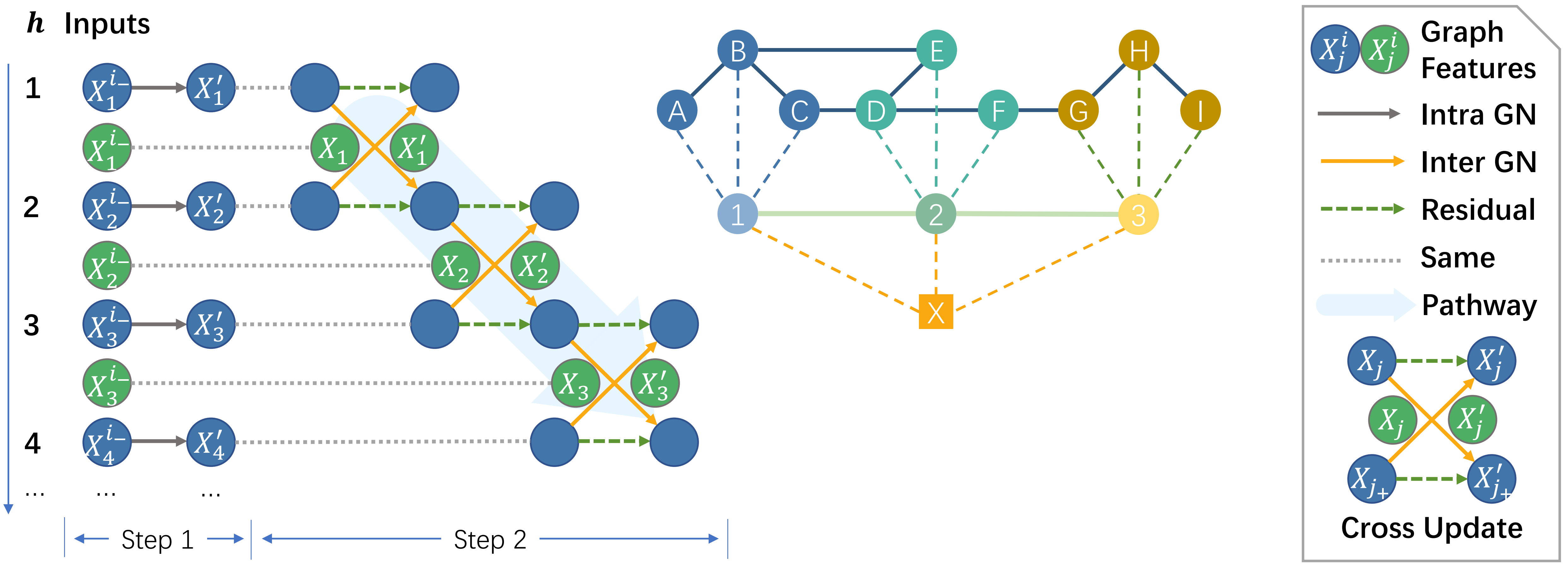
MeGraph Architecture



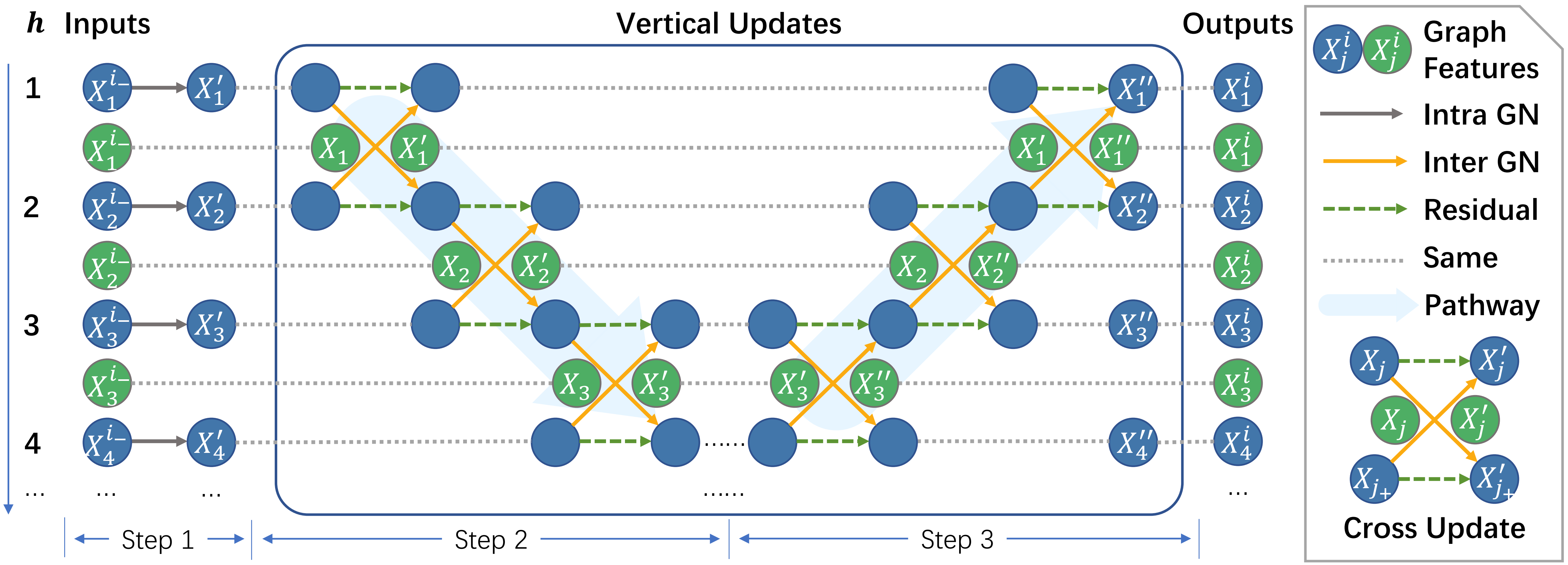
Mee Layer: Bidirectional Pathway



Mee Layer: Bidirectional Pathway



Mee Layer: Bidirectional Pathway



Experiments: Graph Theory Benchmark

- We establish a Graph Theory Benchmark for LRIs
  - 5 tasks related to graph theory (e.g. shortest path distance)
  - 14 random graph generation methods (e.g. tree graph)
  - Medium size: 300 graph, $|\mathcal{V}| \in [20, 50]$
  - Large size: 500 graphs, $|\mathcal{V}| \in [100, 200]$

Model	Param	SP_{sssd}	MCC	Diameter	SP_{ss}	ECC	Average
Baseline(h=1)	h=1,n=5	3.898	1.229	5.750	12.354	18.971	8.440
MeGraph	h=5,n=5	**0.615**	**0.702**	**0.651**	**0.434**	**0.975**	**0.675**
Graph-UNets	h=5	1.118	1.008	2.031	1.166	2.584	1.581

Experiments: Long Range Graph Benchmark

Methods	Use PE	Peptide-func ↑	Peptide-struct ↓
GCN [17]		59.30 ±0.23	0.3496 ±0.0013
GINE [17]		55.43 ±0.78	0.3547 ±0.0045
GatedGCN [17]		58.64 ±0.77	0.3420 ±0.0013
GatedGCN+RWSE [17]	✓	60.69 ±0.35	0.3357 ±0.0006
GatedGCN+RWSE+VN [10]	✓	66.85 ±0.62	0.2529 ±0.0009
Transformer+LapPE [17]	✓	63.26 ±1.26	0.2529 ±0.0016
SAN+LapPE [17]	✓	63.84 ±1.21	0.2683 ±0.0043
SAN+RWSE [17]	✓	64.39 ±0.75	0.2545 ±0.0012
GPS [46]	✓	65.35 ±0.41	0.2500 ±0.0005
MGT+WavePE [43]	✓	68.17 ±0.64	**0.2453** ±0.0025
GNN-AK+ [27]		64.80 ±0.89	0.2736 ±0.0007
SUN [27]		67.30 ±0.78	0.2498 ±0.0008
GraphTrans+PE [27]	✓	63.13 ±0.39	0.2777 ±0.0025
GINE+PE [27]	✓	64.05 ±0.77	0.2780 ±0.0021
GINE-MLP-Mixer+PE [27]	✓	69.21 ±0.54	0.2485 ±0.0004
MeGraph (h=9,n=1)		67.52 ±0.78	0.2557 ±0.0011
MeGraph (h=9,n=4)		**69.45** ±0.77	0.2507 ±0.0009

Conclusion and Takeaway Message

- We introduced the **MeGraph** architecture
  - Naturally alternates local and hierarchical updates
- Locality and hierarchy are both important for capturing LRIs

- Please visit our project page and read our paper for more details
  - https://sites.google.com/view/megraph