

# Weisfeiler and Lehman Go Cellular: CW Networks

Cristian Bodnar<sup>1,\*</sup> Fabrizio Frasca<sup>2,3,\*</sup> Nina Otter<sup>7</sup>  
Yu Guang Wang<sup>4,5,6\*</sup> Pietro Liò<sup>1</sup> Guido Montúfar<sup>4,7,\*</sup> Michael Bronstein<sup>2,3</sup>



Imperial College  
London



MAX-PLANCK-  
GESELLSCHAFT



上海交通大学  
SHANGHAI JIAO TONG UNIVERSITY



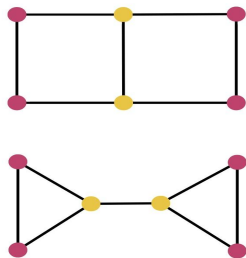
UNSW  
SYDNEY

UCLA

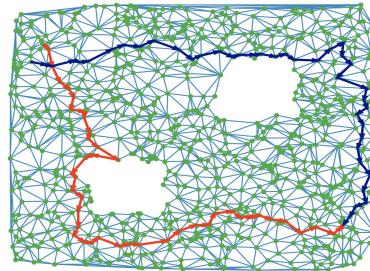


# Motivation: Limitations of Graph Neural Networks

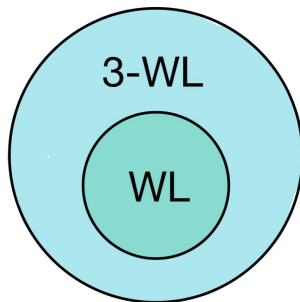
**Substructures**



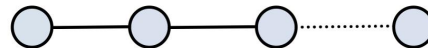
**Higher-order signals**



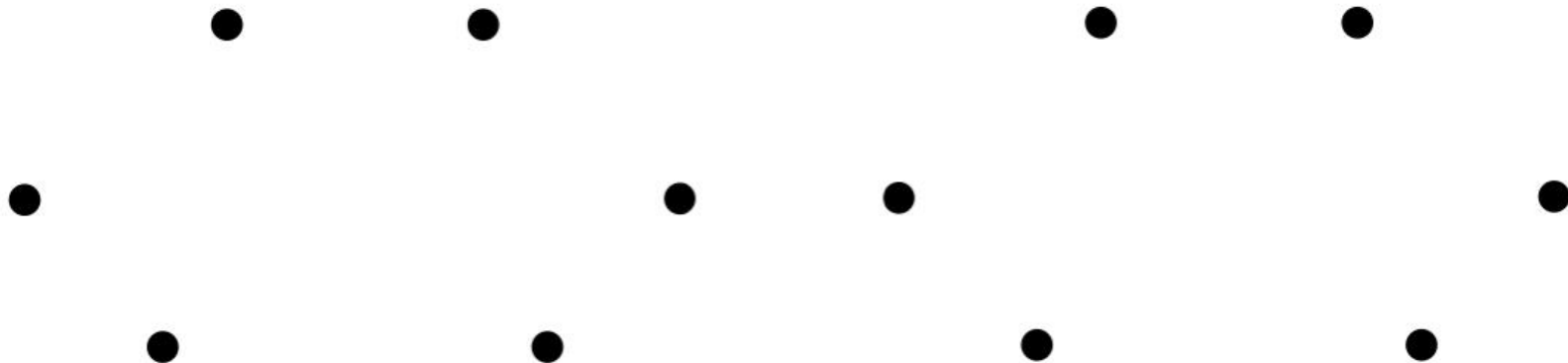
**Expressive power**



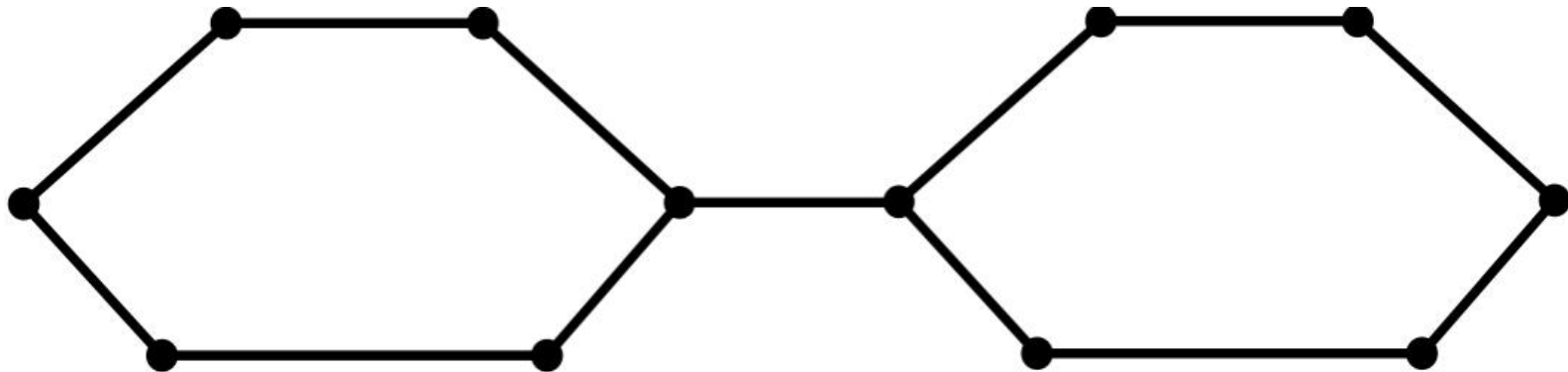
**Long-range interactions**



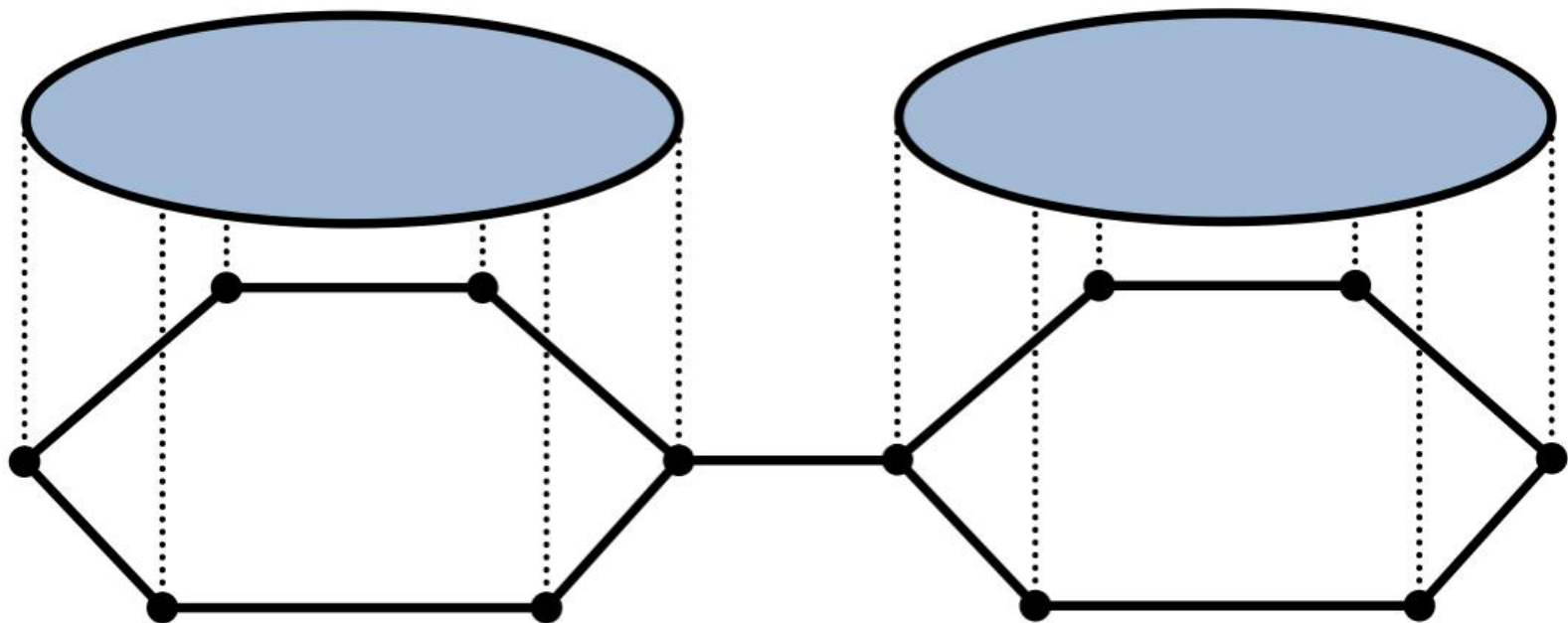
# CW Complexes



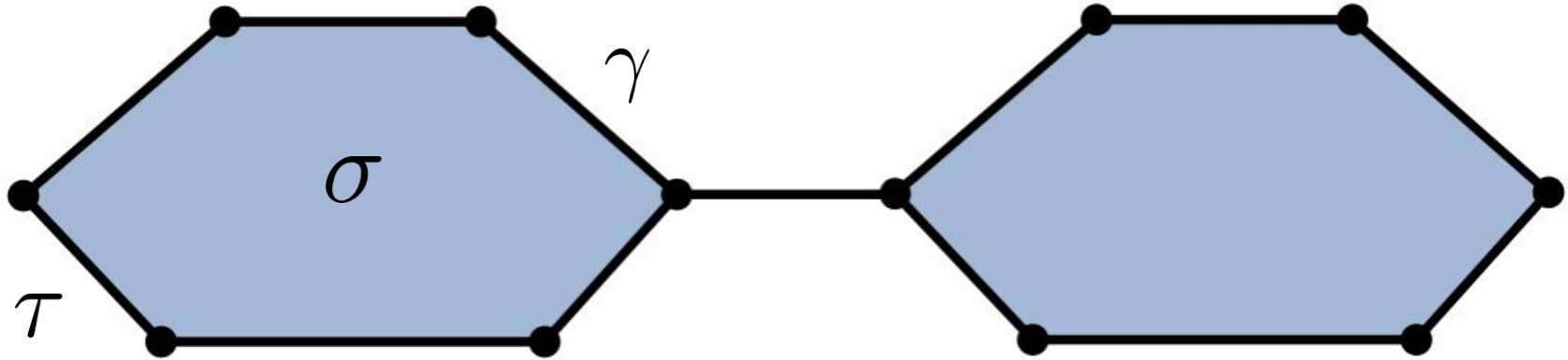
# CW Complexes



# CW Complexes



# CW Complexes

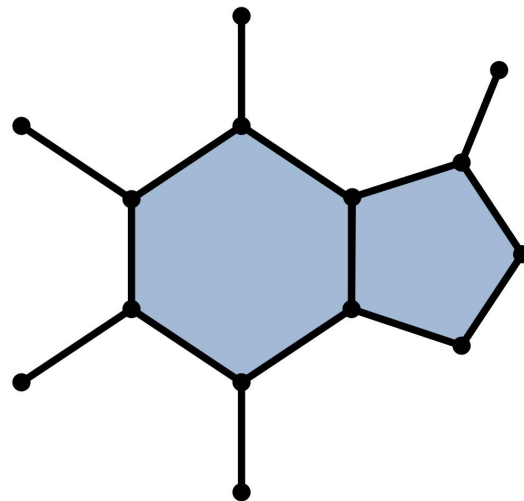
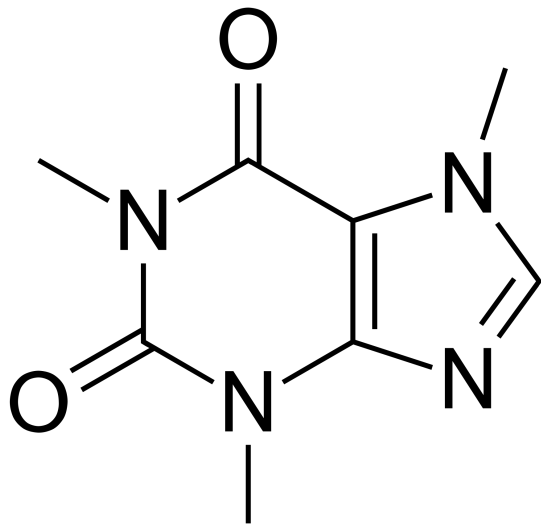


**Boundary adjacency:** Cell  $\tau$  is on the boundary of cell  $\sigma$

**Upper adjacency:**  $\tau, \gamma$  are on the boundary of cell  $\sigma$

# Molecules as Cell Complexes

Graph representations of molecules date back to the nineteenth century [1]. However, it is not necessarily the best representation. We propose modelling **molecules as cell complexes**.



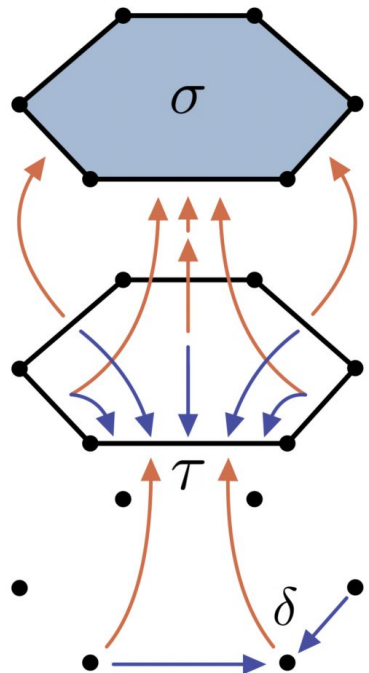
# CWNs and Molecular Message Passing

$$\underbrace{h_\sigma^{t+1}} = U\left(h_\sigma^t, \underbrace{m_{\mathcal{B}}^t(\sigma)}_{\text{orange}}, \underbrace{m_{\uparrow}^{t+1}(\sigma)}_{\text{blue}}\right)$$

$$\underbrace{m_{\mathcal{B}}^{t+1}(\sigma)}_{\text{orange}} = \text{AGG}_{\tau \in \mathcal{B}(\sigma)} \left( M_{\mathcal{B}}(h_\sigma^t, h_\tau^t) \right)$$

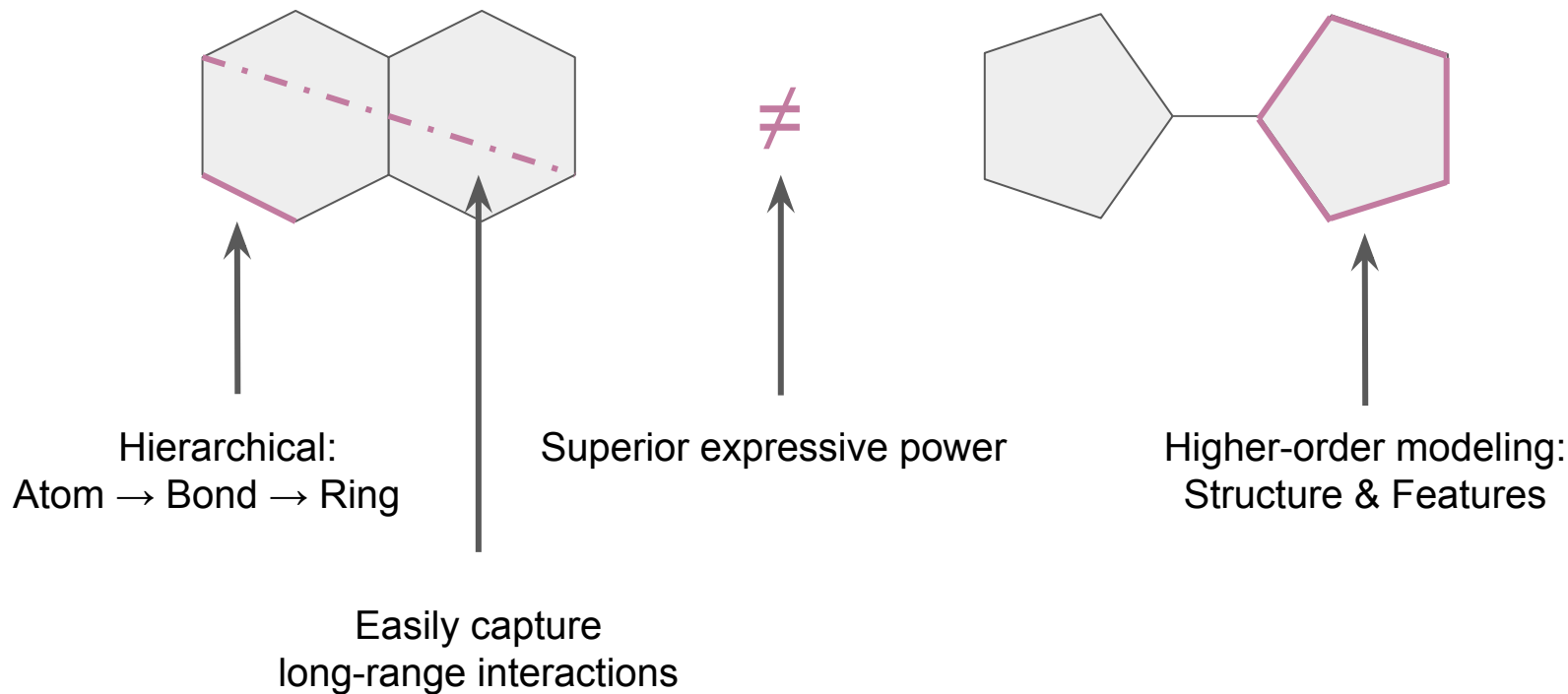
$$\underbrace{m_{\uparrow}^{t+1}(\sigma)}_{\text{blue}} = \text{AGG}_{\tau \in \mathcal{N}_{\uparrow}(\sigma), \delta \in \mathcal{C}(\sigma, \tau)} \left( M_{\uparrow}(h_\sigma^t, h_\tau^t, h_\delta^t) \right)$$

$$h_X = \text{READOUT}\left(\underbrace{\{\{h_\sigma^L\}\}_{\dim(\sigma)=0}}_{\text{purple}}, \underbrace{\{\{h_\sigma^L\}\}_{\dim(\sigma)=1}}_{\text{purple}}, \underbrace{\{\{h_\sigma^L\}\}_{\dim(\sigma)=2}}_{\text{purple}}\right)$$





# Practical Benefits of CWNs



# Real-world Experiments: ZINC + MolHIV

Table 3: ZINC (MAE), ZINC-FULL (MAE) and Mol-HIV (ROC-AUC).

Method	ZINC ↓		ZINC-FULL ↓	MOLHIV ↑
	No Edge Feat.	With Edge Feat.	All methods	All methods
GCN [45]	0.469±0.002	N/A	N/A	76.06±0.97
GAT [67]	0.463±0.002	N/A	N/A	N/A
GatedGCN [10]	0.422±0.006	0.363±0.009	N/A	N/A
GIN [72]	0.408±0.008	0.252±0.014	0.088±0.002	77.07±1.49
PNA [19]	0.320±0.032	0.188±0.004	N/A	79.05±1.32
DGN [5]	0.219±0.010	0.168±0.003	N/A	79.70±0.97
HIMP [26]	N/A	0.151±0.006	0.036±0.002	78.80±0.82
GSN [9]	0.139±0.007	0.108±0.018	N/A	77.99±1.00
<b>CIN-small (Ours)</b>	0.139±0.008	0.094±0.004	0.044±0.003	80.55±1.04
<b>CIN (Ours)</b>	<b>0.115±0.003</b>	<b>0.079±0.006</b>	<b>0.022±0.002</b>	<b>80.94±0.57</b>

# Real-world Experiments: ZINC (ablation study)

Method	MAE	
● GatedGCN [10]	$0.363 \pm 0.009$	
● GIN [72]	$0.252 \pm 0.014$	
● PNA [19]	$0.188 \pm 0.004$	
● DGN [5]	$0.168 \pm 0.003$	
● ● HIMP [26]	$0.151 \pm 0.006$	Importance of updating edges
● GSN [9]	$0.108 \pm 0.018$	
● GIN-E Custom	$0.196 \pm 0.007$	
● CIN No-Rings small	$0.174 \pm 0.006$	Importance of capturing rings
● CIN No-Rings	$0.159 \pm 0.007$	
● ● CIN-small	$0.094 \pm 0.004$	
● ● CIN	<b><math>0.079 \pm 0.006</math></b>	

● “Standard”

● Updates edges

● Captures rings

# Thanks!

Cristian Bodnar

Twitter: @crisbodnar

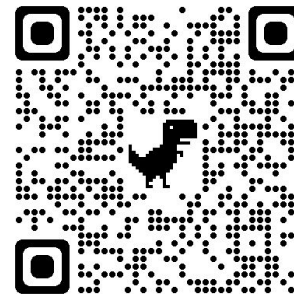
Email: cb2015@cam.ac.uk

Fabrizio Frasca

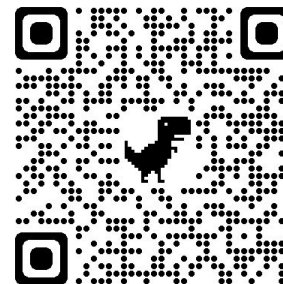
Twitter: @ffabffrasca

Email: ffrasca@twitter.com

*ArXiv*



*Github*



# Backup: Computational Complexity

(Message passing)

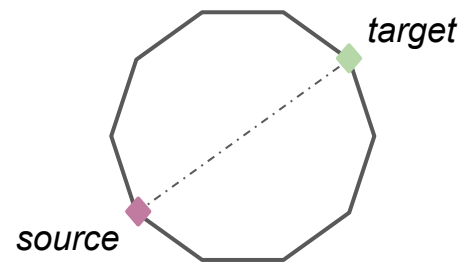
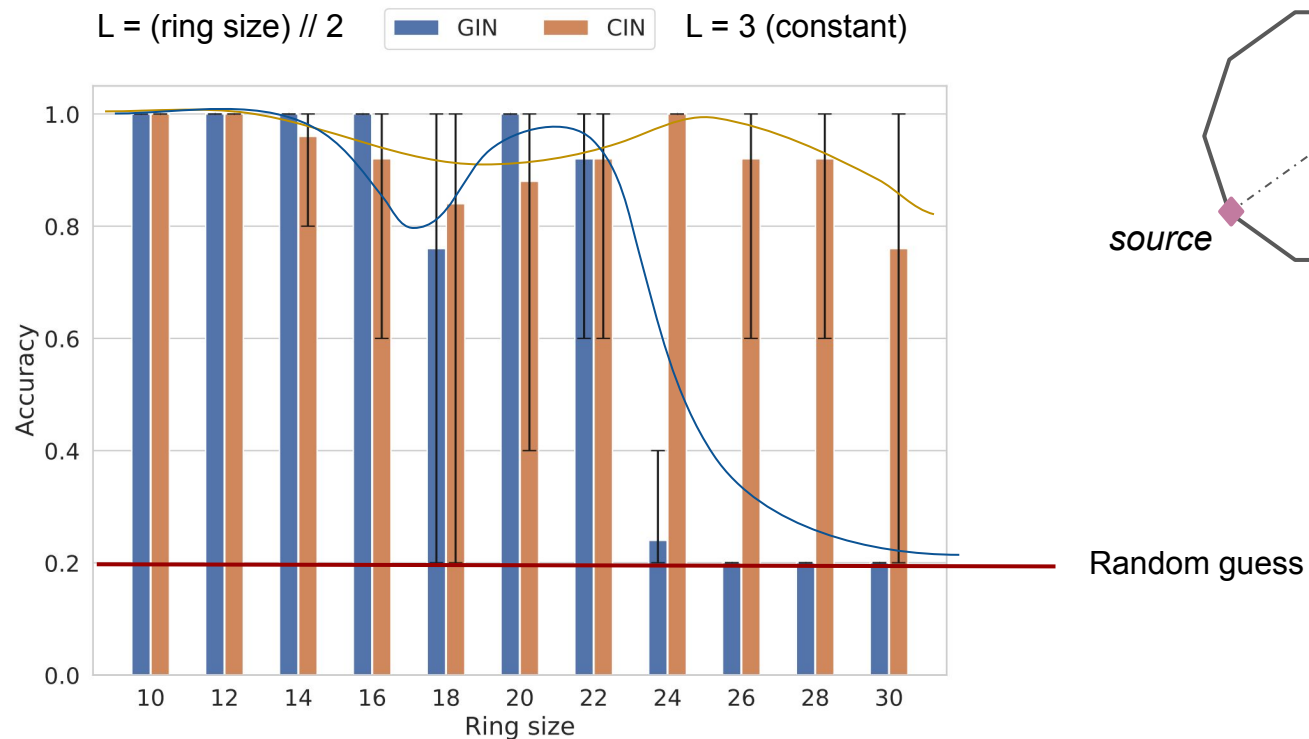
$$\mathcal{O}\left(\underbrace{\sum_{p=1}^d \underbrace{B_p S_p}_{\text{Boundary msgs}}}_{\text{Max boundary size}} + \underbrace{2 * \left(\underbrace{B_p}_2\right) S_p}_{\text{Upper msgs}}\right) \xrightarrow{\text{In practice: } B_p \text{ fixed (small) constant}} \Theta\left(\sum_{p=1}^d S_p\right)$$

Empirically, GIN is only 19% – 35% faster at inference time (ZINC)

(Lifting)

Dataset ↓ / Processes →	linear trend →					
	Seq.	2	4	8	16	32
ZINC (12k)	320.27 ± 0.54	169.95 ± 0.32	84.90 ± 0.21	43.38 ± 0.07	23.17 ± 0.68	18.59 ± 0.68
Mol-HIV (41k)	1178.98 ± 3.90	635.58 ± 0.83	319.01 ± 0.40	164.26 ± 0.52	86.92 ± 0.77	60.62 ± 2.05
ZINC-FULL (250k)	6805.35 ± 16.50	3549.16 ± 7.73	1782.41 ± 3.84	918.38 ± 3.46	492.77 ± 6.13	383.92 ± 3.30

# Backup: RingTransfer



# Backup: TUDatasets

Dataset	MUTAG	PTC	PROTEINS	NCI1	NCI109	IMDB-B	IMDB-M	RDT-B
RWK [28]	79.2 $\pm$ 2.1	55.9 $\pm$ 0.3	59.6 $\pm$ 0.1	>3 days	N/A	N/A	N/A	N/A
GK ( $k = 3$ ) [62]	81.4 $\pm$ 1.7	55.7 $\pm$ 0.5	71.4 $\pm$ 0.31	62.5 $\pm$ 0.3	62.4 $\pm$ 0.3	N/A	N/A	N/A
PK [56]	76.0 $\pm$ 2.7	59.5 $\pm$ 2.4	73.7 $\pm$ 0.7	82.5 $\pm$ 0.5	N/A	N/A	N/A	N/A
WL kernel [63]	90.4 $\pm$ 5.7	59.9 $\pm$ 4.3	75.0 $\pm$ 3.1	<b>86.0</b> $\pm$ 1.8	N/A	73.8 $\pm$ 3.9	50.9 $\pm$ 3.8	81.0 $\pm$ 3.1
DCNN [3]	N/A	N/A	61.3 $\pm$ 1.6	56.6 $\pm$ 1.0	N/A	49.1 $\pm$ 1.4	33.5 $\pm$ 1.4	N/A
DGCNN [74]	85.8 $\pm$ 1.8	58.6 $\pm$ 2.5	75.5 $\pm$ 0.9	74.4 $\pm$ 0.5	N/A	70.0 $\pm$ 0.9	47.8 $\pm$ 0.9	N/A
IGN [50]	83.9 $\pm$ 13.0	58.5 $\pm$ 6.9	<b>76.6</b> $\pm$ 5.5	74.3 $\pm$ 2.7	<b>72.8</b> $\pm$ 1.5	72.0 $\pm$ 5.5	48.7 $\pm$ 3.4	N/A
GIN [72]	89.4 $\pm$ 5.6	64.6 $\pm$ 7.0	76.2 $\pm$ 2.8	82.7 $\pm$ 1.7	N/A	75.1 $\pm$ 5.1	52.3 $\pm$ 2.8	<b>92.4</b> $\pm$ 2.5
PPGNs [51]	<b>90.6</b> $\pm$ 8.7	66.2 $\pm$ 6.6	<b>77.2</b> $\pm$ 4.7	83.2 $\pm$ 1.1	<b>82.2</b> $\pm$ 1.4	73.0 $\pm$ 5.8	50.5 $\pm$ 3.6	N/A
Natural GN [20]	89.4 $\pm$ 1.6	<b>66.8</b> $\pm$ 1.7	71.7 $\pm$ 1.0	82.4 $\pm$ 1.3	N/A	73.5 $\pm$ 2.0	51.3 $\pm$ 1.5	N/A
GSN [9]	<b>92.2</b> $\pm$ 7.5	<b>68.2</b> $\pm$ 7.2	<b>76.6</b> $\pm$ 5.0	<b>83.5</b> $\pm$ 2.0	N/A	<b>77.8</b> $\pm$ 3.3	<b>54.3</b> $\pm$ 3.3	N/A
SIN [7]	N/A	N/A	76.4 $\pm$ 3.3	82.7 $\pm$ 2.1	N/A	<b>75.6</b> $\pm$ 3.2	<b>52.4</b> $\pm$ 2.9	<b>92.2</b> $\pm$ 1.0
<b>CIN (Ours)</b>	<b>92.7</b> $\pm$ 6.1	<b>68.2</b> $\pm$ 5.6	<b>77.0</b> $\pm$ 4.3	<b>83.6</b> $\pm$ 1.4	<b>84.0</b> $\pm$ 1.6	<b>75.6</b> $\pm$ 3.7	<b>52.7</b> $\pm$ 3.1	<b>92.4</b> $\pm$ 2.1