# INVARIANT GRAPH NETWORKS
# & EXPRESSIVITY

**02/04/25**

Khushee Kapoor
Master of Data Science and Artificial Intelligence

UNIVERSITY OF WATERLOO | FACULTY OF MATHEMATICS

Based on the blogs by *Haggai Maron and Yaron Lipman*

## Invariant and Equivariant Graph Networks

**Haggai Maron, Heli Ben-Hamu, Nadav Shamir & Yaron Lipman**
Department of Computer Science and Applied Mathematics
Weizmann Institute of Science
Rehovot, Israel

## On the Universality of Invariant Networks

**Haggai Maron [1] Ethan Fetaya [2 3] Nimrod Segol [1] Yaron Lipman [1]**

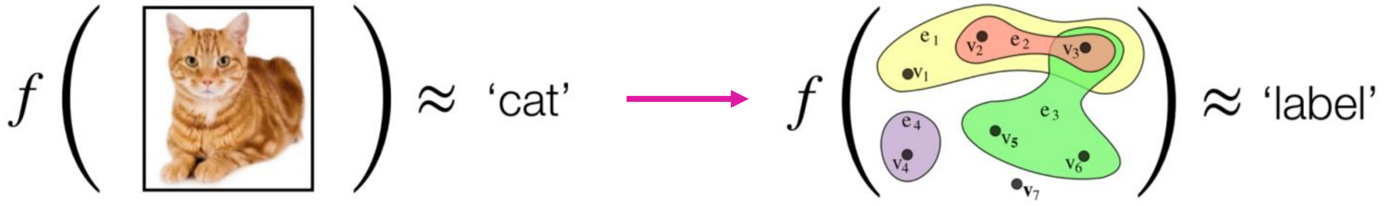
## Provably Powerful Graph Networks

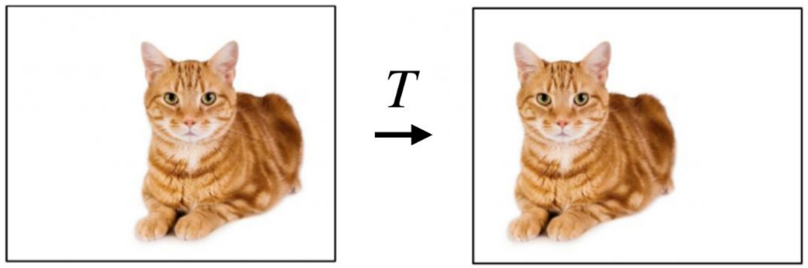
**Haggai Maron* Heli Ben-Hamu* Hadar Serviansky* Yaron Lipman**
Weizmann Institute of Science
Rehovot, Israel

# ALGEBRAIC VIEW OF GRAPH NEURAL NETWORKS

neural network architectures suitable for learning irregular data in the form of graphs - hypergraphs

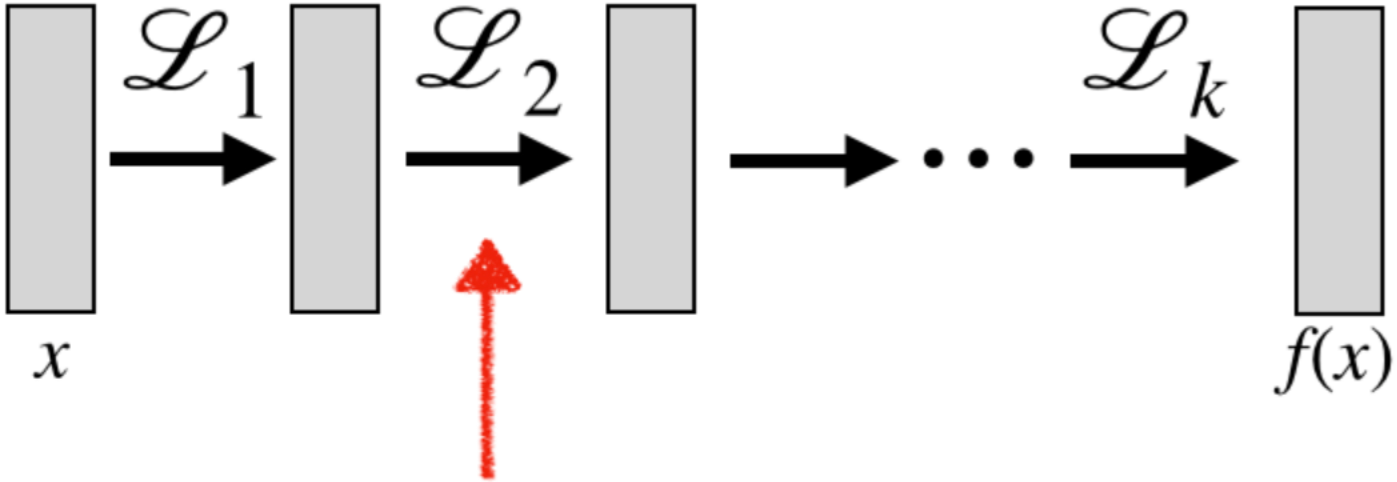
trade-off between expressivity and efficiency




$$f(x) \quad = \quad f(T.x)$$

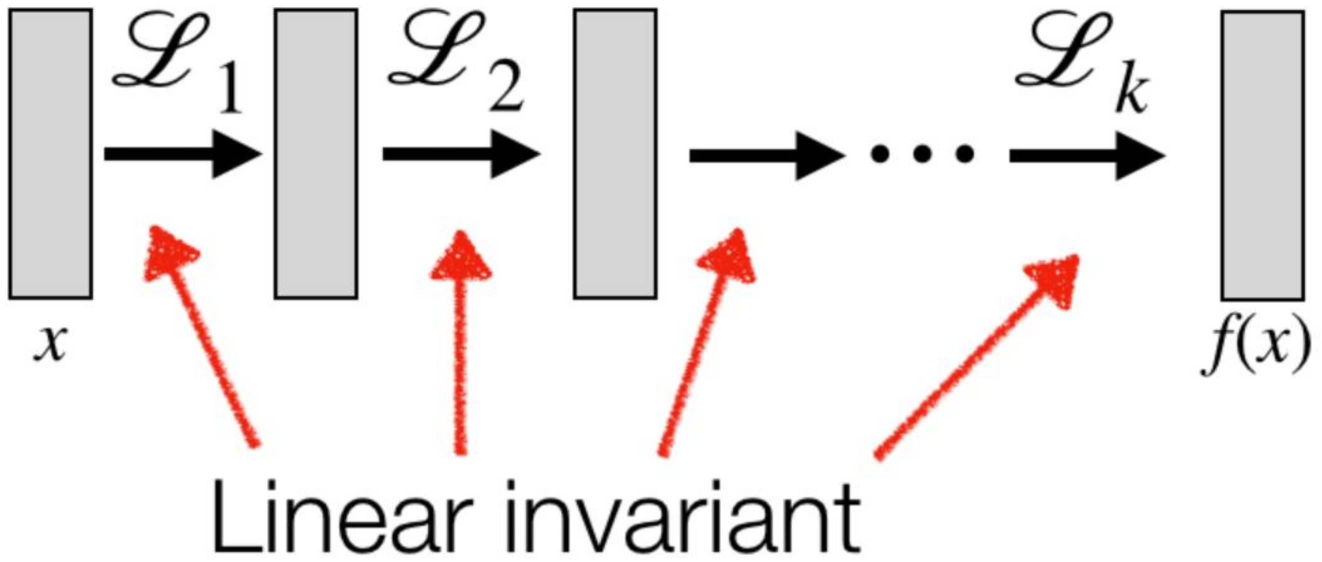
# MULTI LAYER PERCEPTRON




$$\mathscr{L}(x) = \sigma(L(x)) = \sigma(Ax + b)$$

$[100 \times 100 \times 3] \rightarrow [100 \times 100 \times 3] \sim 10^9$

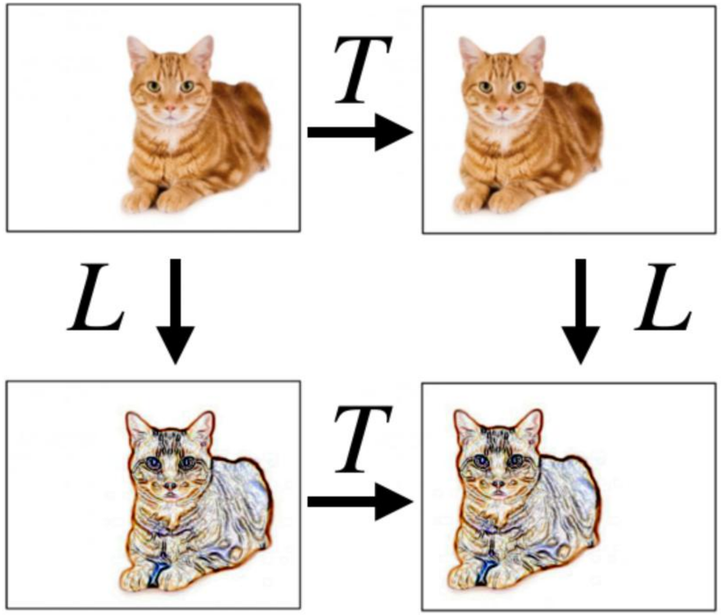
# LINEAR INVARIANT FUNCTIONS



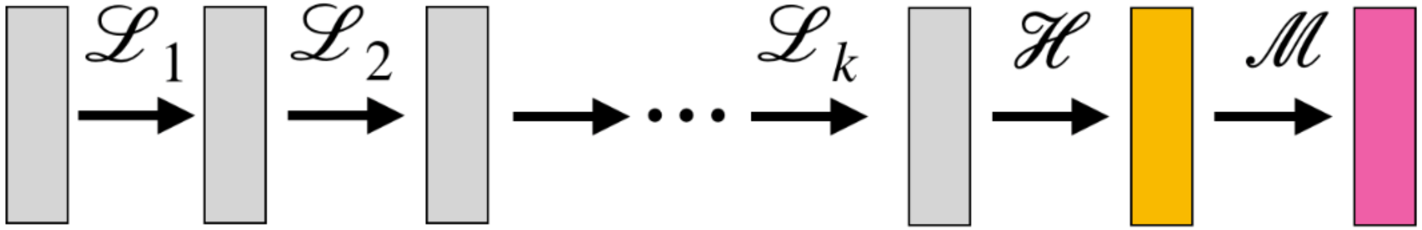

$$x \mapsto L(x) = Ax + b \qquad L(x) = L(T.x)$$

one dimensional vector space ~ sum operator

# EQUIVARIANT LINEAR OPERATORS




$$L(T.x) = T.L(x)$$



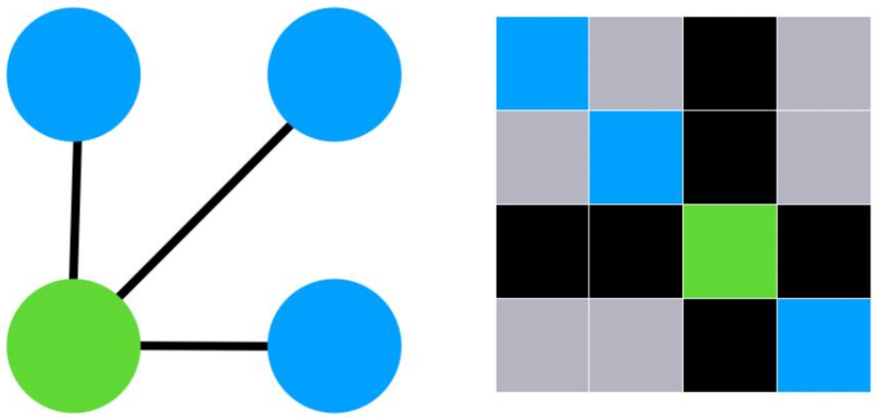



UNIVERSITY OF WATERLOO | FACULTY OF MATHEMATICS

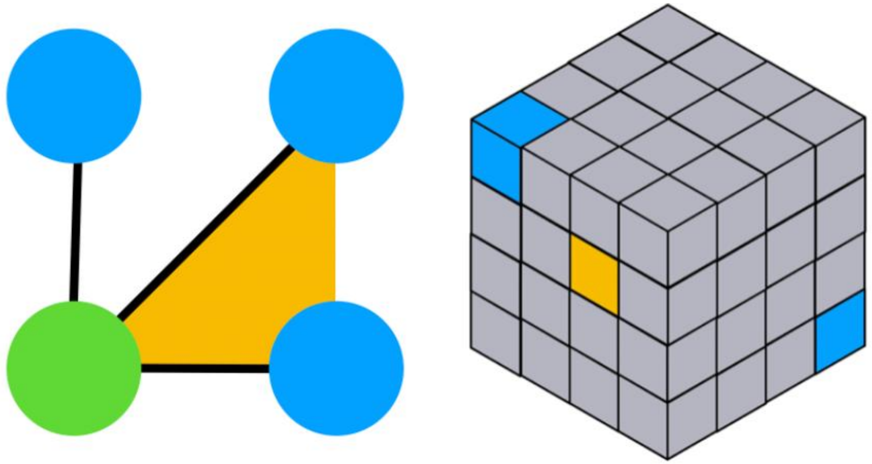
# REPRESENTING GRAPHS AS TENSORS

Definition of graph:
- **n** elements/ nodes
- $x_i$ information attached to i-th element
- some information attached to edges $x_{ij}$

Representation as tensor:
- $X \in \mathbb{R}^{nxn}$
- diagonal elements $X_{i,i} = x_i$ encode node data
- off-diagonal elements $X_{i,j} = x_{i,j}$ encode edge data



simple graph as tensor



hypergraph of order 3 as tensor

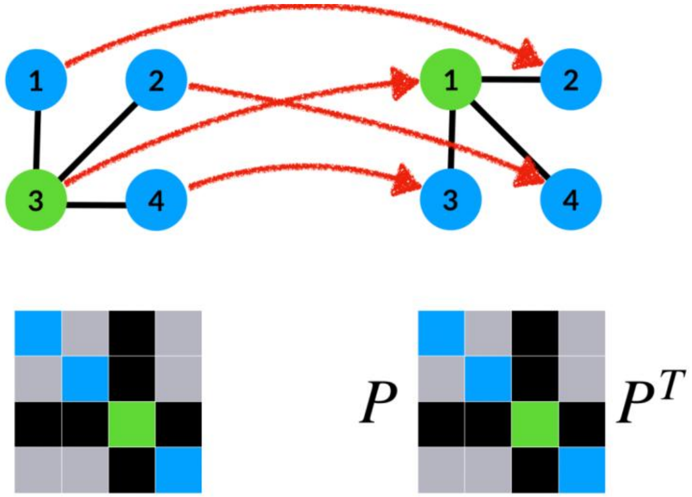
# SYMMETRIES OF GRAPHS

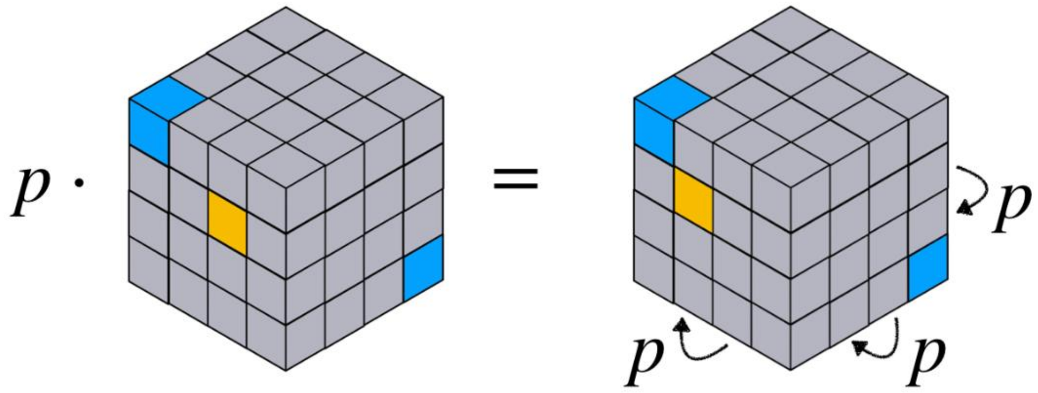
- Transformations that do not change the input data are called symmetries.
- Two graphs $X, Y$ are considered isomorphic if

$$Y = p.X \text{ and } (i\text{-}j)\text{th entry of } X \text{ is } X_{p^{-1}(i),p^{-1}(j)}$$

- Generalizing for hypergraphs

$$(i_1, \ldots, i_k)\text{th entry of } X \text{ is } X_{p^{-1}(i_1),\ldots,\, p^{-1}(i_k)}$$

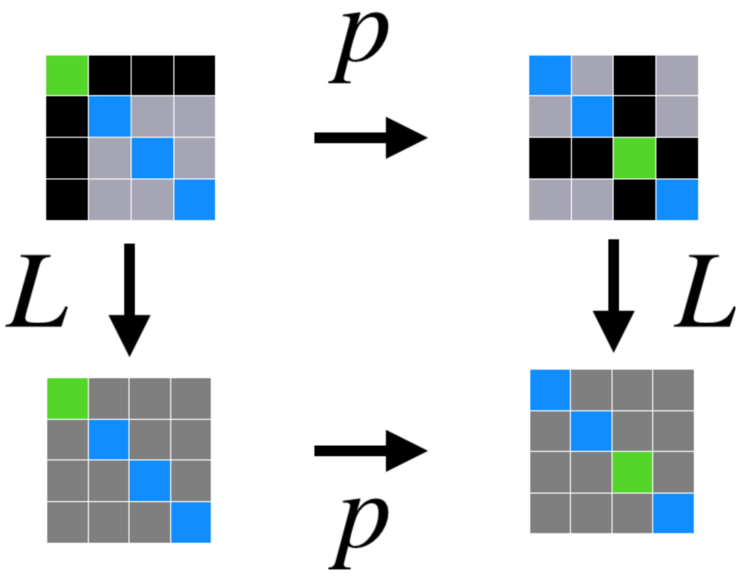


# INVARIANT GRAPH NETWORKS

- Trying to consider linear transformations of graph data that are invariant leads to a poor space of operators.
- Networks using only such invariant layers could not distinguish two graphs if they have the same number of nodes and edges.
- Remedy comes from considering equivariant operators.

$$L(p.x) = p.L(x)$$




$$f(X) = L_1(X) \circ \cdots \circ L_k \circ H \circ M$$

**Invariant Graph Network**

equivariant linear operators can map between different order tensors $\mathbb{R}^{n^k} \to \mathbb{R}^{n^l}$

# LINEAR OPERATORS AND THE FIXED-POINT EQUATIONS

Characterizing affine transformations $\mathbb{R}^{n^k} \to \mathbb{R}^{n^l}$ equivariant ($l \geq 1$) or invariant ($l = 0$) to $X \mapsto p.X$:

- Ignoring the bias, $L\colon \mathbb{R}^{n^k} \to \mathbb{R}^{n^l}$ can be treated as a tensor $L \in \mathbb{R}^{n^{k+l}}$
- Above is similar to $\mathbb{R}^k \to \mathbb{R}^l$ represented using matrix $\mathbb{R}^{k \times l}$
- $L(p.x) = p.L(x)$ can be expressed compactly as $p.L = L$

Fixed-point Equations

The space of equivariant/ invariant operators $L\colon \mathbb{R}^{n^k} \to \mathbb{R}^{n^l}$ is characterized by all tensors $L \in \mathbb{R}^{n^{k+l}}$ that are fixed by the action of the permutation group.

# SOLVING FIXED-POINT EQUATIONS

Any solution $L$ should be constant along the orbit of the permutation group.

- Considering $k = l = 2$, corresponding to equivariant operators $\mathbb{R}^{n\times n} \mapsto \mathbb{R}^{n\times n}$
- Taking $p = (12)$, then $L_{1,1,1,1} = p.L_{1,1,1,1} = L_{p^{-1}(1),\, p^{-1}(1),\, p^{-1}(1),\, p^{-1}(1)} = L_{2,2,2,2}$ implying $L_{i,i,i,i}$ is constant
- Similarly, $L_{1,1,2,3}$ equals all entries of the form $L_{i,i,j,s}$

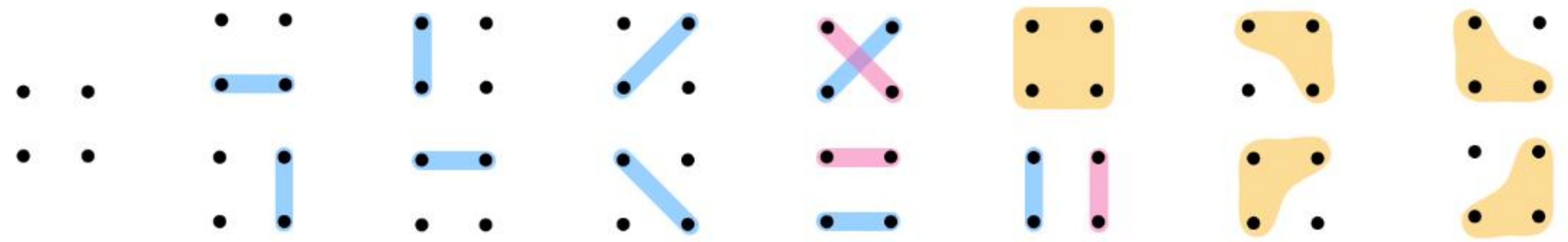
*$L$ is constant along all indices that preserve equality and inequality relations between pairs*

**Bell Number**

no. of different orbits

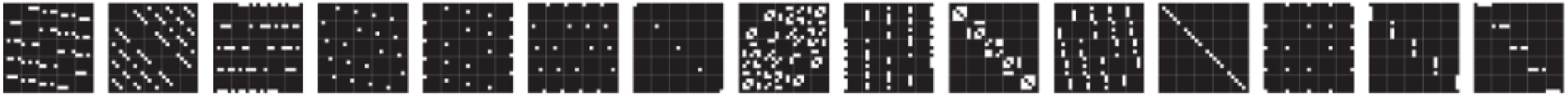
no. of different equality patterns of indices

no. of partitions of a set with elements = indices



# SOLVING FIXED-POINT EQUATIONS

- An orthogonal basis to the equivariant operators can be constructed using an equality pattern $\alpha$ and the indicator tensor $B^{\alpha}_{i,j,s,t} = 1$ if and only if $(i, j, s, t) \in \alpha$ and 0 otherwise.
- Applying equivariant operators can be done in $O(n^2)$ operations
- Thus, the linear equivariant operator has the form $L = \sum_{\alpha} w_{\alpha} B^{\alpha}$
- Solving fixed point equations for $l, k \in \mathbb{N}$ reduces to fixed point equation for $L \in \mathbb{R}^{n^{k+l}}$, in this case we will have $bell(k + l)$ basis elements the basis is given by the indicator tensors $B^{\alpha}$ of these equality patterns.



# HOW EXPRESSIVE ARE IGNs?

**Function Approximation POV**

- 2-IGNs can approximate MPNNs on compact sets.
- IGNs are universal for large tensor order k, approximating any continuous invariant function.

**Graph Discrimination POV**

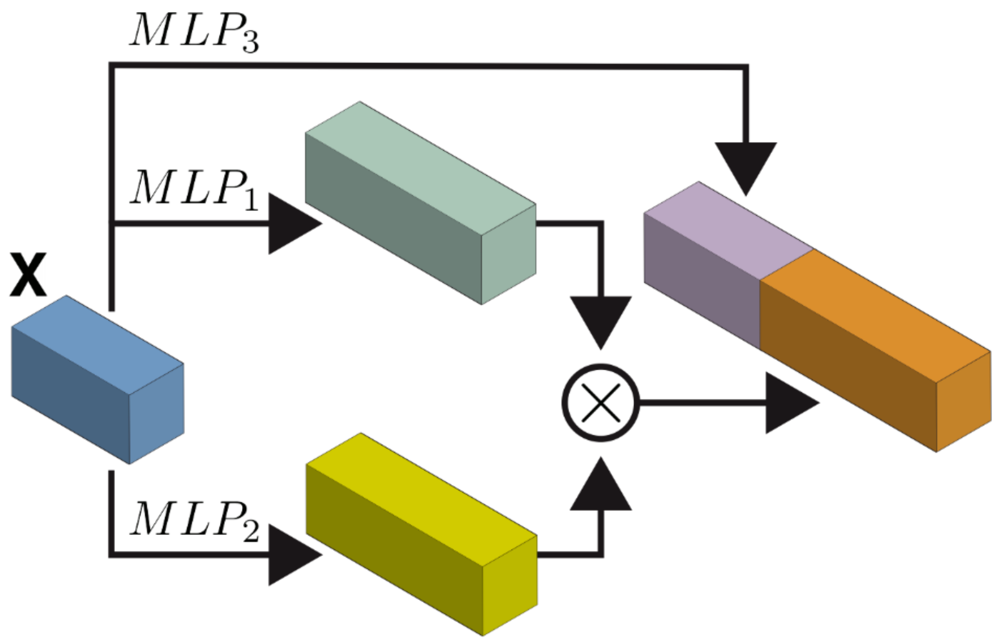
- k-IGNs match or exceed the graph discrimination power of k-WL, distinguishing graphs like 3-WL distinguishes regular graphs.

# EXPRESSIVE VARIANT OF IGNS

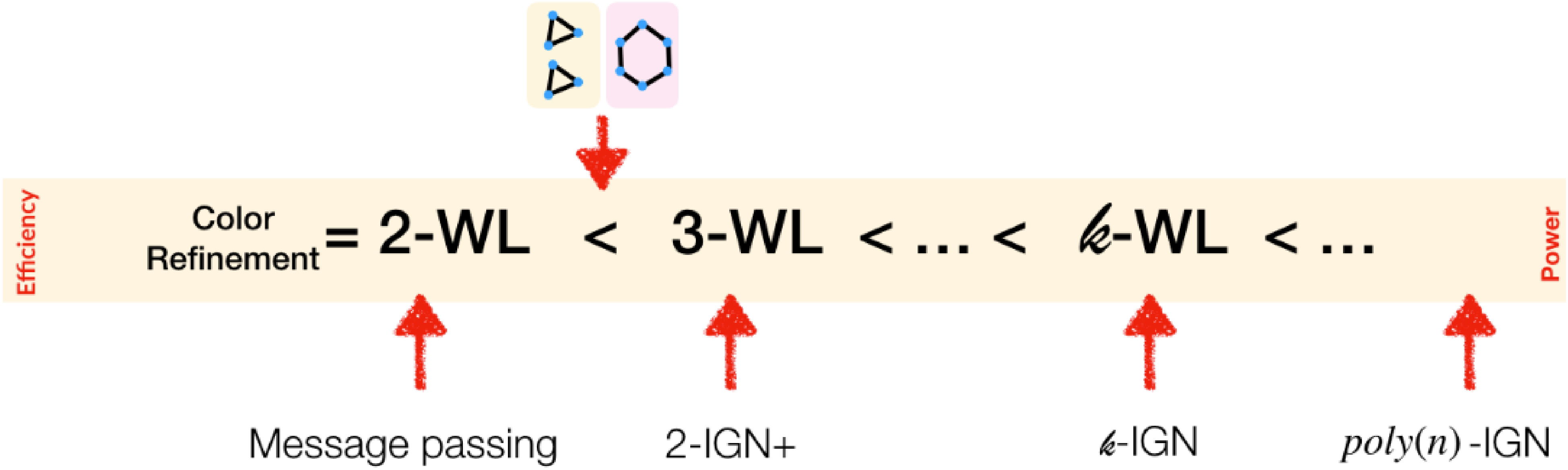
- Drawback of k-IGNs: storing and processing k-order tensors.

- **2-IGN+**: atleast as expressive as 3-WL test. Composed of:
  - 3 different MLPs
  - input 2-tensor
  - output of two multiplied feature-wise
  - output of third concatenated to product

- **Strengths:**
  - MLPs on feature dimension can be implemented on 1x1 image convolution
  - matrix multiplication supported by all DL frameworks



# SUMMARY OF IGN EXPRESSIVENESS





UNIVERSITY OF WATERLOO | FACULTY OF MATHEMATICS

UNIVERSITY OF WATERLOO

FACULTY OF MATHEMATICS