**Summary**

We prove universal approximation theorems for 1-layer invariant and equivariant Graph Neural Networks.

- The studied (theoretical) GNNs have unbounded width and tensorization order.
- Results are uniformly valid for (hyper)graphs of varying number of nodes, for a single set of parameters.
- The equivariant case is much more involved and requires a new Stone-Weierstrass theorem.

**Notations**

- **Graph:** $W \in \mathbb{R}^{m \times n}$, $d$-**Hypergraph:** $W \in \mathbb{R}^{d \times n}$
- **Permutation:** bijection $\sigma : [n] \to [n]$
- **Permuted (hyper)graph:** $\sigma \ast W \in \mathbb{R}^{m \times n}$
- **Invariant function:** $f(\sigma \ast W) = f(W)$
- **Equivariant function:** $f(\sigma \ast W) = \sigma \ast f(W)$

**Invariant and equivariant linear layers**

**Theorem (Maron et al. [1])**

There is a basis of $b(k + p)$ equivariant linear operators $\mathbb{R}^d \to \mathbb{R}^d$, where $b(k)$ is the $k$th Bell number.

*(invariant case: just take $p = 0$)*

- Does not depend on $n$. Ex: there are exactly 15 equivariant linear operators $\mathbb{R}^2 \to \mathbb{R}^2$.
- The number of trainable parameters of $\ell$-layer GNNs is \( \sum_{k=1}^{\ell} \binom{(d+n)}{n} (k+1) + 1 \)
- A GNN (1) with a fixed set of parameters can be applied to graphs of any size.

**Sketch of proof**

**Invariant case**

Apply Stone-Weierstrass theorem (like in Hornik et al. [4]), quotienting $\mathbb{G}$ by graph isomorphisms.

**Theorem (Stone-Weierstrass)**

An algebra of continuous functions that separates points is dense in the set of continuous functions.

**Algebra of GNNs (aka “the cos trick”)**

1. Authorize product of GNNs to obtain an algebra
2. Prove universality for $\rho = \cos$
3. A product of cos is also a sum!
4. Approximate cos with any $\rho$ using MLP universality theorem

**Separation of points**

- For any two distinct points, there is a function that distinguishes them.
- Here: "For two non-isomorphic graphs, there is a GNN that distinguishes them."
- We prove: “Two graphs that coincide for every GNNs are isomorphic.”

**Equivariant case**


**Theorem (Stone-Weierstrass for equivariant functions; Keriven and Peyré [3])**

An algebra of equivariant continuous functions that separates points and separates coordinates is dense in the set of continuous functions.

**Separation of coordinates:** “for a given graph $W$, and any two coordinates $1 \leq i, j \leq n$ that are not related by an automorphism of $W$ (i.e. $\sigma \ast W = W$), there is an equivariant GNN that distinguishes them.”

**Proof:** non-trivial adaptation of [5]

**Main results: universality of GNNs**

Compact set of graphs: $\mathbb{G} = \{ W \in \mathbb{R}^{d \times n} ; n \leq n_{\text{max}}, \|W\| \leq R \}$

**Theorem (Maron et al. [2]; Keriven and Peyrè [3])**

The set $\mathcal{F}_0$ of invariant GNNs is dense in the set of invariant continuous functions on $\mathbb{G}$ (for the sup norm).

**Theorem (Keriven and Peyré [3])**

The set $\mathcal{F}_0$ of equivariant GNNs is dense in the set of equivariant continuous functions on $\mathbb{G}$ (for the sup norm).

- A single set of parameters approximate functions on graphs of varying size uniformly well
- Equivariant case: much more difficult to prove (see below). Valid only for full group of permutations, and order-1 output $y \in \mathbb{R}^d$.

**Outlooks**

- **Convolutional GNNs**
- Approximation power/stability with respect to weaker metrics on graphs (e.g. cut-metric)
- Behavior in the large-graph limit (see [6])

**Numerics (toy)**

Approximation results on synthetic data for invariant (left) or equivariant (right) GNNs. The tensorization order $k$ plays a greater role than the width $s$.

**References**