On Representing Linear Programs by Graph Neural Networks

ZIANG CHEN, JIALIN LIU, XINSHANG WANG, JIANFENG LU, AND WOTAO YIN

Linear Programming (LP)

$$\min_{x\in\mathbb{R}^n} \ c^\top x, \quad \text{s.t. } Ax \circ b, \ l \leq x \leq u,$$

$$A \in \mathbb{R}^{m\times n},\ c \in \mathbb{R}^n,\ b \in \mathbb{R}^m,\ l \in (\mathbb{R} \cup \{-\infty\})^n,\ u \in (\mathbb{R} \cup \{+\infty\})^n,\ \circ \in \{\leq, =, \geq\}^m$$

Three cases for LP:

- **Infeasible:** the feasible set $\mathcal{X}_F := \{x \in \mathbb{R}^n : Ax \circ b,\ l \leq x \leq u\}$ is empty.
- **Unbounded:** The feasible set is non-empty, but the objective value can be arbitrarily good.
- **Feasible and bounded:** There exists $x^* \in \mathcal{X}_F$ such that $c^\top x^* \leq c^\top x$ for all $x \in \mathcal{X}_F$, x^* is an optimal solution, $c^\top x^*$ is the optimal objective value.

Problem Setup

$(A, b, c, l, u, \circ) \longrightarrow$ $\boxed{\textbf{ML models}}$ $\longrightarrow$

Feasibility {0, 1}

Boundedness {R, -∞, ∞ }

One of its optimal solutions (if exist){ $\mathbb{R}^n$ }

LP represented as weighted bipartite graph

$$
\begin{aligned}
\min_{x\in\mathbb{R}^3} \ & x_1 + x_2 + 2x_3, \\
\text{s.t. } & x_1 + 2x_2 + x_3 \geq 1, \\
& 2x_1 + x_2 + x_3 \geq 1, \\
& x_1 \geq -2, \\
& x_2 \geq -2, \\
& -3 \leq x_3 \leq 3.
\end{aligned}
$$

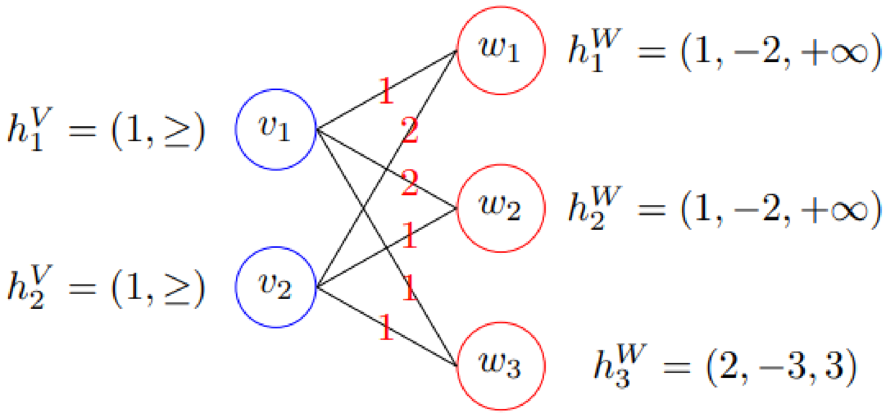


FIGURE 1. An example of LP-graph

- An LP is represented as a graph $(G, H) \in \mathcal{G}_{m,n} \times \mathcal{H}_m^V \times \mathcal{H}_n^W$.

Problem Setup

$(A, b, c, l, u, \circ) \longrightarrow$ **GNN models** $\longrightarrow$

Feasibility $\{0, 1\}$

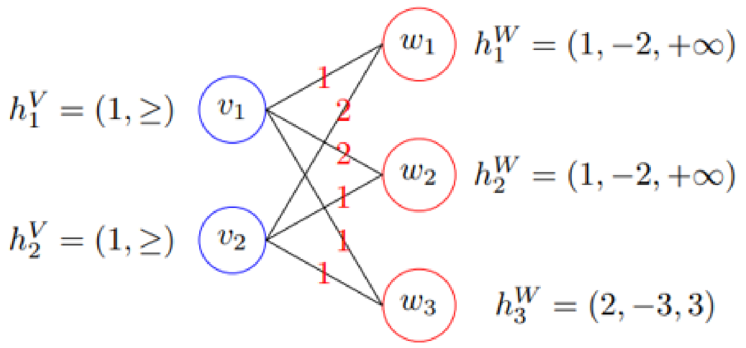
Boundedness $\{\mathrm{R}, \ -\infty, \ \infty\}$

One of its optimal solutions (if exist) $\{\mathbb{R}^n\}$

$$\mathcal{G}_{m,n} \times \mathcal{H}_m^V \times \mathcal{H}_n^W \to \mathbb{R}$$

$$\mathcal{G}_{m,n} \times \mathcal{H}_m^V \times \mathcal{H}_n^W \to \mathbb{R}^n$$

GNNs-LP



$$h_i^{0,V} = f_{\text{in}}^V(h_i^V),\ h_j^{0,W} = f_{\text{in}}^W(h_j^W),\ i = 1, 2, \dots, m,\ j = 1, 2, \dots, n, \tag{2.1}$$

$$h_i^{l,V} = g_l^V\left(h_i^{l-1,V}, \sum_{j=1}^{n} E_{i,j} f_l^W(h_j^{l-1,W})\right), \quad i = 1, 2, \dots, m, \tag{2.2}$$

$$h_j^{l,W} = g_l^W\left(h_j^{l-1,W}, \sum_{i=1}^{m} E_{i,j} f_l^V(h_i^{l-1,V})\right), \quad j = 1, 2, \dots, n, \tag{2.3}$$

$$y_{\text{out}} = f_{\text{out}}\left(\sum_{i=1}^{m} h_i^{L,V}, \sum_{j=1}^{n} h_j^{L,W}\right). \tag{2.4}$$

$$y_{\text{out}}(w_j) = f_{\text{out}}^W\left(\sum_{i=1}^{m} h_i^{L,V}, \sum_{j=1}^{n} h_j^{L,W}, h_j^{L,W}\right), \quad j = 1, 2, \cdots, n. \tag{2.5}$$

$$\begin{aligned} \mathcal{F}_{\text{GNN}} &= \{F : \mathcal{G}_{m,n} \times \mathcal{H}_m^V \times \mathcal{H}_n^W \to \mathbb{R} \mid F \text{ yields } (2.1), (2.2), (2.3), (2.4)\}, \\ \mathcal{F}_{\text{GNN}}^W &= \{F : \mathcal{G}_{m,n} \times \mathcal{H}_m^V \times \mathcal{H}_n^W \to \mathbb{R}^n \mid F \text{ yields } (2.1), (2.2), (2.3), (2.5)\}. \end{aligned} \tag{2.6}$$

$f_{\text{in}}^V, f_{\text{in}}^W, f_{\text{out}}, f_{\text{out}}^W, \{f_l^V, f_l^W, g_l^V, g_l^W\}_{l=0}^L$ are usually parameterized with multi-linear perceptrons (MLP)

GNNs-LP

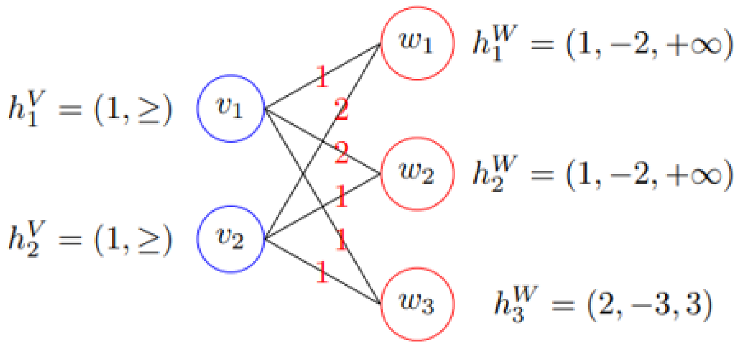
$f_{\text{in}}^V, f_{\text{in}}^W, f_{\text{out}}, f_{\text{out}}^W, \{f_l^V, f_l^W, g_l^V, g_l^W\}_{l=0}^L$ are usually parameterized with multi-linear perceptrons (**MLP**)

Advantages:

- **Invariance and Equivariance**

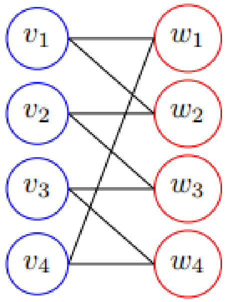
$$F(G,H) = F\left((\sigma_V, \sigma_W) * (G,H)\right), \ \forall \sigma_V \in S_m, \sigma_W \in S_n,$$

$$\sigma_W(F_W(G,H)) = F_W\left((\sigma_V, \sigma_W) * (G,H)\right), \ \forall \sigma_V \in S_m, \sigma_W \in S_n,$$



GNNs-LP **Beyond WL-Test**

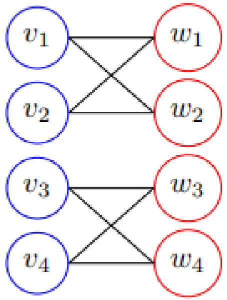
infeasible **unbounded**



$$\begin{aligned} \min\ & x_1 + x_2 + x_3 + x_4, \\ \text{s.t. }\ & x_1 + x_2 = 1, \\ & x_2 + x_3 = 1, \\ & x_3 + x_4 = 1, \\ & x_4 + x_1 = 1, \\ & x_j \geq 1,\ 1 \leq j \leq 4. \end{aligned}$$

$$\begin{aligned} \min\ & x_1 + x_2 + x_3 + x_4, \\ \text{s.t. }\ & x_1 + x_2 \leq 1, \\ & x_2 + x_3 \leq 1, \\ & x_3 + x_4 \leq 1, \\ & x_4 + x_1 \leq 1, \\ & x_j \leq 1,\ 1 \leq j \leq 4. \end{aligned}$$

$$\begin{aligned} \min\ & x_1 + x_2 + x_3 + x_4, \\ \text{s.t. }\ & x_1 + x_2 = 1, \\ & x_2 + x_3 = 1, \\ & x_3 + x_4 = 1, \\ & x_4 + x_1 = 1, \\ & x_j \leq 1,\ 1 \leq j \leq 4. \end{aligned}$$



$$\begin{aligned} \min\ & x_1 + x_2 + x_3 + x_4, \\ \text{s.t. }\ & x_1 + x_2 = 1, \\ & x_2 + x_1 = 1, \\ & x_3 + x_4 = 1, \\ & x_4 + x_3 = 1, \\ & x_j \geq 1,\ 1 \leq j \leq 4. \end{aligned}$$

$$\begin{aligned} \min\ & x_1 + x_2 + x_3 + x_4, \\ \text{s.t. }\ & x_1 + x_2 \leq 1, \\ & x_2 + x_1 \leq 1, \\ & x_3 + x_4 \leq 1, \\ & x_4 + x_3 \leq 1, \\ & x_j \leq 1,\ 1 \leq j \leq 4. \end{aligned}$$

$$\begin{aligned} \min\ & x_1 + x_2 + x_3 + x_4, \\ \text{s.t. }\ & x_1 + x_2 = 1, \\ & x_2 + x_1 = 1, \\ & x_3 + x_4 = 1, \\ & x_4 + x_3 = 1, \\ & x_j \leq 1,\ 1 \leq j \leq 4. \end{aligned}$$

Feasible bounded with the same optimal solution

$(1/2, 1/2, 1/2, 1/2)$ with the smallest ℓ_2-norm

GNNs approximate Lp

$(A, b, c, l, u, \circ) \longrightarrow$ **ML models** $\longrightarrow$

Feasibility $\{0, 1\}$

Boundedness $\{R, -\infty, \infty\}$

One of its optimal solutions (if exist) $\{\mathbb{R}^n\}$

Three mapping:

- **Feasibility mapping**

$$\Phi_{\text{feas}} : \mathcal{G}_{m,n} \times \mathcal{H}_m^V \times \mathcal{H}_n^W \to \{0, 1\}$$

- **Optimal objective value mapping**

$$\Phi_{\text{obj}} : \mathcal{G}_{m,n} \times \mathcal{H}_m^V \times \mathcal{H}_n^W \to \mathbb{R} \cup \{\infty, -\infty\}$$

- **Optimal solution mapping**

$$\Phi_{\text{solu}} : \Phi_{\text{obj}}^{-1}(\mathbb{R}) \to \mathbb{R}^n \qquad \text{with the smallest } \ell_2\text{-norm}$$

GNNs approximate Lp

Are there GNNs that can accurately approximate Φ_{feas}, Φ_{obj} and Φ_{solu}?

GNNs approximate Lp

Three mapping:

- **Feasibility mapping**

$$\Phi_{\text{feas}} : \mathcal{G}_{m,n} \times \mathcal{H}_m^V \times \mathcal{H}_n^W \to \{0,1\}$$

$\mathcal{F}_{\text{GNN}}$

- **Optimal objective value mapping**

$$\Phi_{\text{obj}} : \mathcal{G}_{m,n} \times \mathcal{H}_m^V \times \mathcal{H}_n^W \to \mathbb{R} \cup \{\infty, -\infty\}$$

- **Optimal solution mapping**

$$\Phi_{\text{solu}} : \Phi_{\text{obj}}^{-1}(\mathbb{R}) \to \mathbb{R}^n \text{ with the smallest } \ell_2\text{-norm}$$

$\mathcal{F}_{\text{GNN}}^W$

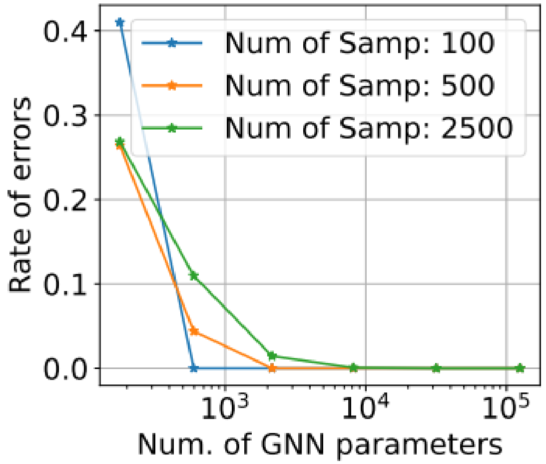
GNNs approximate Lp

- **Feasibility mapping**

Theorem 3.2. *Let $X \subset \mathcal{G}_{m,n} \times \mathcal{H}_m^V \times \mathcal{H}_n^W$ be measurable with finite measure. For any $\epsilon > 0$, there exists some $F \in \mathcal{F}_{GNN}$, such that*

$$Meas\left(\left\{(G,H) \in X : \mathbb{I}_{F(G,H)>1/2} \neq \Phi_{feas}(G,H)\right\}\right) < \epsilon,$$

where $\mathbb{I}_{\cdot}$ is the indicator function, i.e., $\mathbb{I}_{F(G,H)>1/2} = 1$ if $F(G,H) > 1/2$ and $\mathbb{I}_{F(G,H)>1/2} = 0$ otherwise.



(A) Feasibility

GNNs approximate Lp

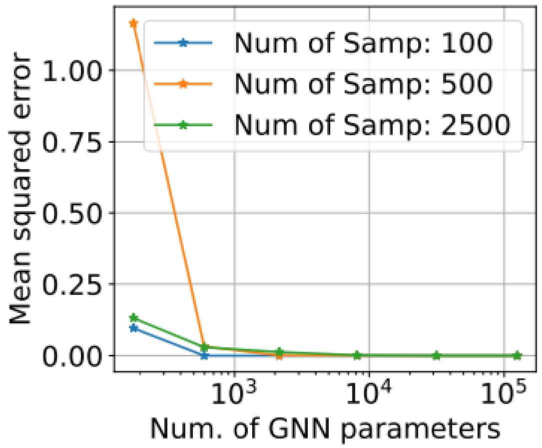
- **Optimal objective value mapping**

Theorem 3.4. *Let $X \subset \mathcal{G}_{m,n} \times \mathcal{H}_m^V \times \mathcal{H}_n^W$ be measurable with finite measure. For any $\epsilon > 0$, there exists $F_1 \in \mathcal{F}_{GNN}$ such that*

$$\text{(3.1)} \qquad Meas\left(\{(G,H) \in X : \mathbb{I}_{F_1(G,H)>1/2} \neq \mathbb{I}_{\Phi_{obj}(G,H)\in\mathbb{R}}\}\right) < \epsilon.$$

For any $\epsilon, \delta > 0$, there exists $F_2 \in \mathcal{F}_{GNN}$ such that

$$\text{(3.2)} \qquad Meas\left(\left\{(G,H) \in X \cap \Phi_{obj}^{-1}(\mathbb{R}) : |F_2(G,H) - \Phi_{obj}(G,H)| > \delta\right\}\right) < \epsilon.$$



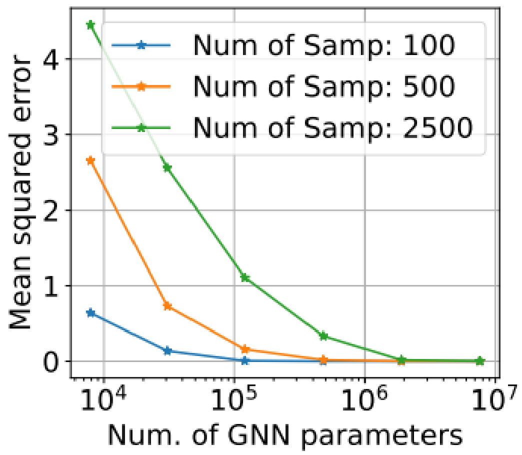
(B) Optimal objective value

GNNs approximate Lp

- **Optimal solution mapping**

Theorem 3.6. *Let $X \subset \Phi_{obj}^{-1}(\mathbb{R}) \subset \mathcal{G}_{m,n} \times \mathcal{H}_m^V \times \mathcal{H}_n^W$ be measurable with finite measure and be closed under actions in $S_m \times S_n$. For any $\epsilon, \delta > 0$, there exists some $F_W \in \mathcal{F}_{GNN}^W$, such that*

$$Meas\left(\{(G,H) \in X : \|F(G,H) - \Phi_{solu}(G,H)\| > \delta\}\right) < \epsilon.$$



(c) Optimal solution

Conclusion

- GNN can approximate Φ_{feas}, Φ_{obj}, and Φ_{solu}
- The idea of using neural network to approximate discrete values is provided