Lagrangian Propagation
Graph Neural Networks

Matteo Tiezzi\textsuperscript{1}, Giuseppe Marra\textsuperscript{2,1}, Stefano Melacci\textsuperscript{1}, Marco Maggini\textsuperscript{1}, Marco Gori\textsuperscript{1}

\textsuperscript{1}DIISM, University of Siena, Italy
\textsuperscript{2}DINFO, University of Florence, Italy

e-mail: mtiezzi@diism.unisi.it

\begin{align*}
g(\bar{x}) &= 0 \\
g(\bar{x}) &< 0 & \lambda \to - \\
g(\bar{x}) &> 0 & \lambda \to +
\end{align*}
Propose a novel constrained formulation approach to learning w.r.t. GNNs [Scarselli 2009]
Avoid epoch-wise fixed-point convergence, still take advantage of the multi-hop diffusion.

\[
\min \sum_{v \in S} L(g_w(x_v, k-1), y_v)
\]

s.t. \( G(x_v, k - f^k_{w,v}) = 0, \quad \forall v \in V, \forall k \in [0, K - 1] \)

Cast it in the Lagrangian framework.

\[
\min_{\theta_{f_w}, \theta_{g_w}, X} \max_{\Lambda} \mathcal{L}(\theta_{f_w}, \theta_{g_w}, X, \Lambda)
\]

- Jointly optimize transition function and node state representation
- Diffusion as a differential optimization process, aimed at fulfilling the constraints
- Mixed strategy:
  - Still rely on BackPropagation to learn the transition and output functions
  - Exploit constraints to define the diffusion mechanism