How Training Label Noise Affects the Learning Curves of Graph Neural Networks

Ivan Markov - imarkov@tudelft.nl - CSE3000 - Research Project

Supervisors: Elvin Isufi, Mohamed Jebali, Chengen Liu

TU Delft

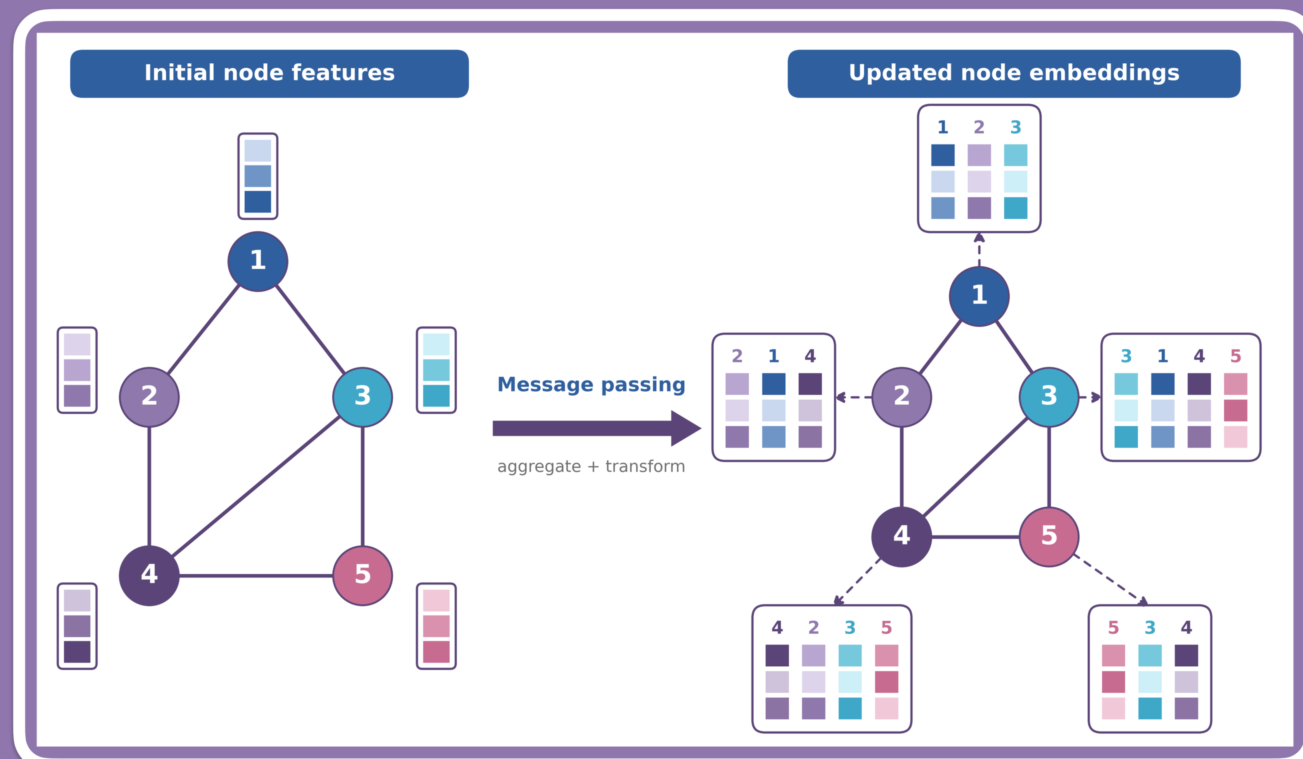
Introduction

GNNs - a deep learning model for graph-structured data

- Used in: drug discovery, social
- networks, fraud detection, and more
- A key open question: How does GNN performance scale with labeled training data?

How do GNNs work?

- Aggregate and transform features of neighbouring nodes
- Repeat multiple times - each node contains information from multiple neighbours



Research Question

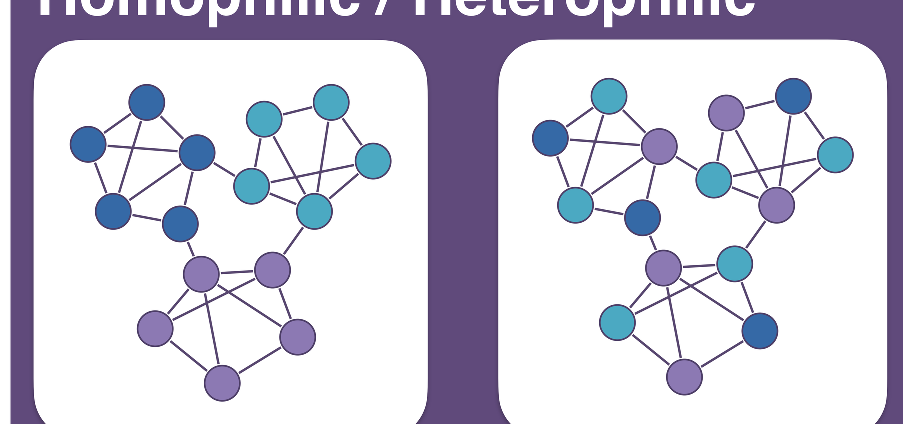
How does training label noise affect the shape of the learning curve of a GNN on node classification tasks?

- Shape
- Plateau
- Homophilic vs heterophilic
- Noise injection protocols

Key Concepts

Learning curve: Test-set metric as a function of number of labeled training nodes

Homophilic / Heterophilic



ChebNet: Spectral GNN using Chebyshev polynomial filters

Methodology and Experiment

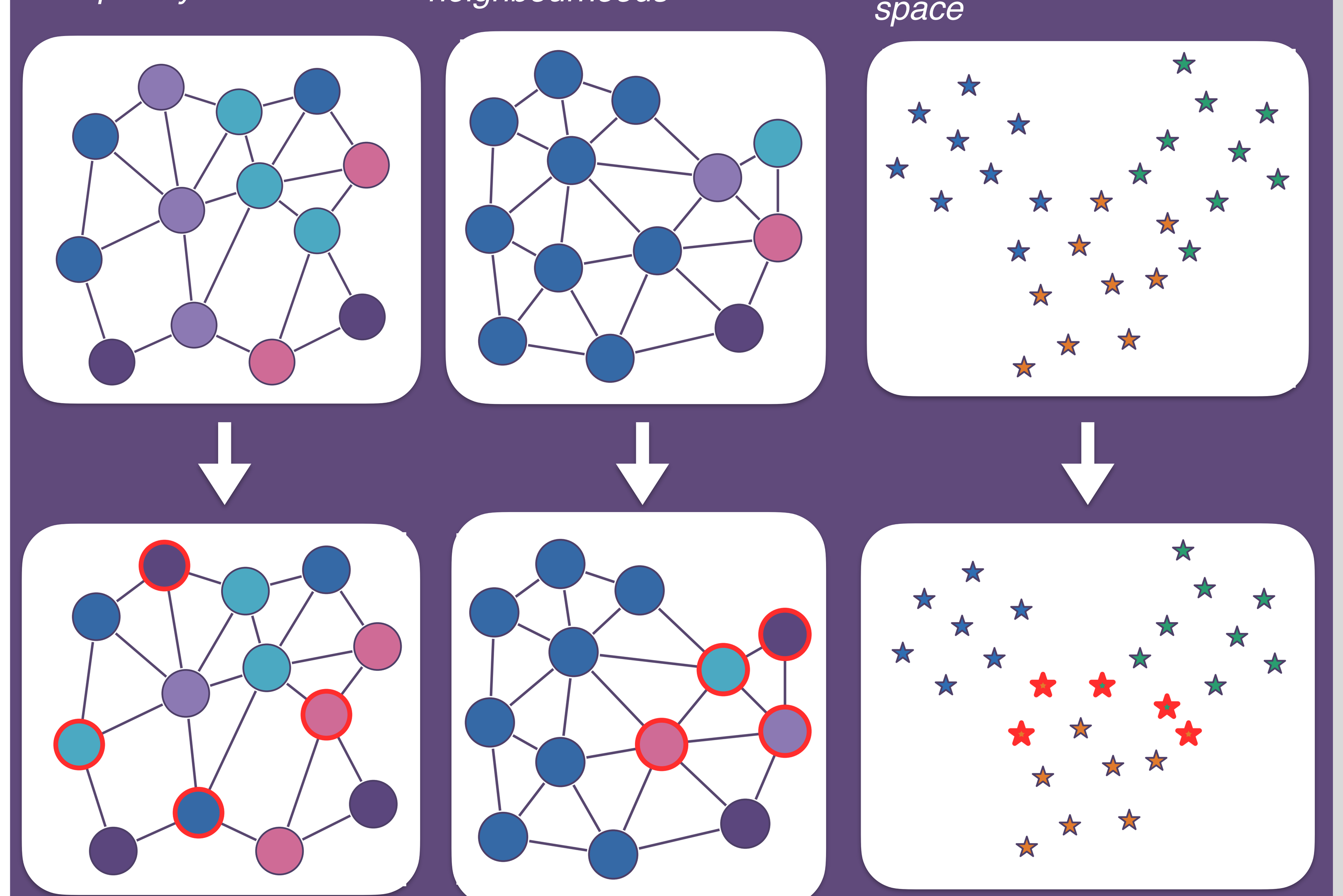
1.Inject Noise in Clean Datasets

Datasets: Cora, Pubmed, Chameleon, Squirrel

Symmetric (Sym) *corrupts nodes completely at random*

Structure-symmetric (Struct) *targets heterophilic neighbourhoods*

Feature-asymmetric (Feat) *targets nodes far from class centroids in feature space*



2. Node Classification

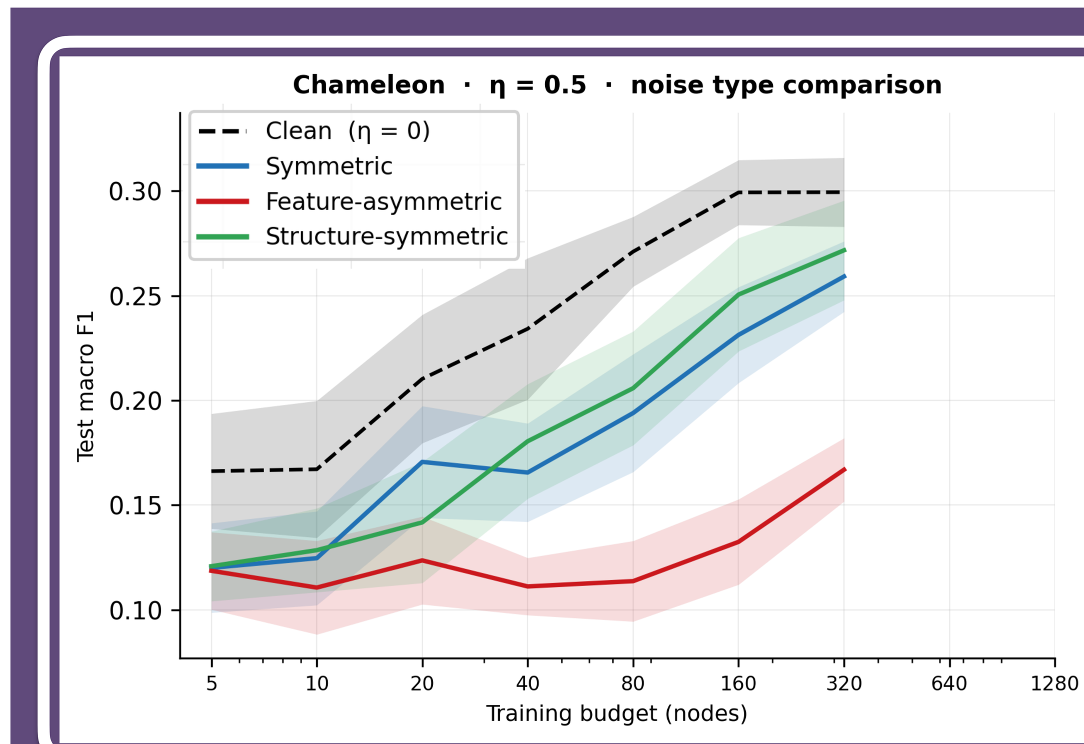
Model

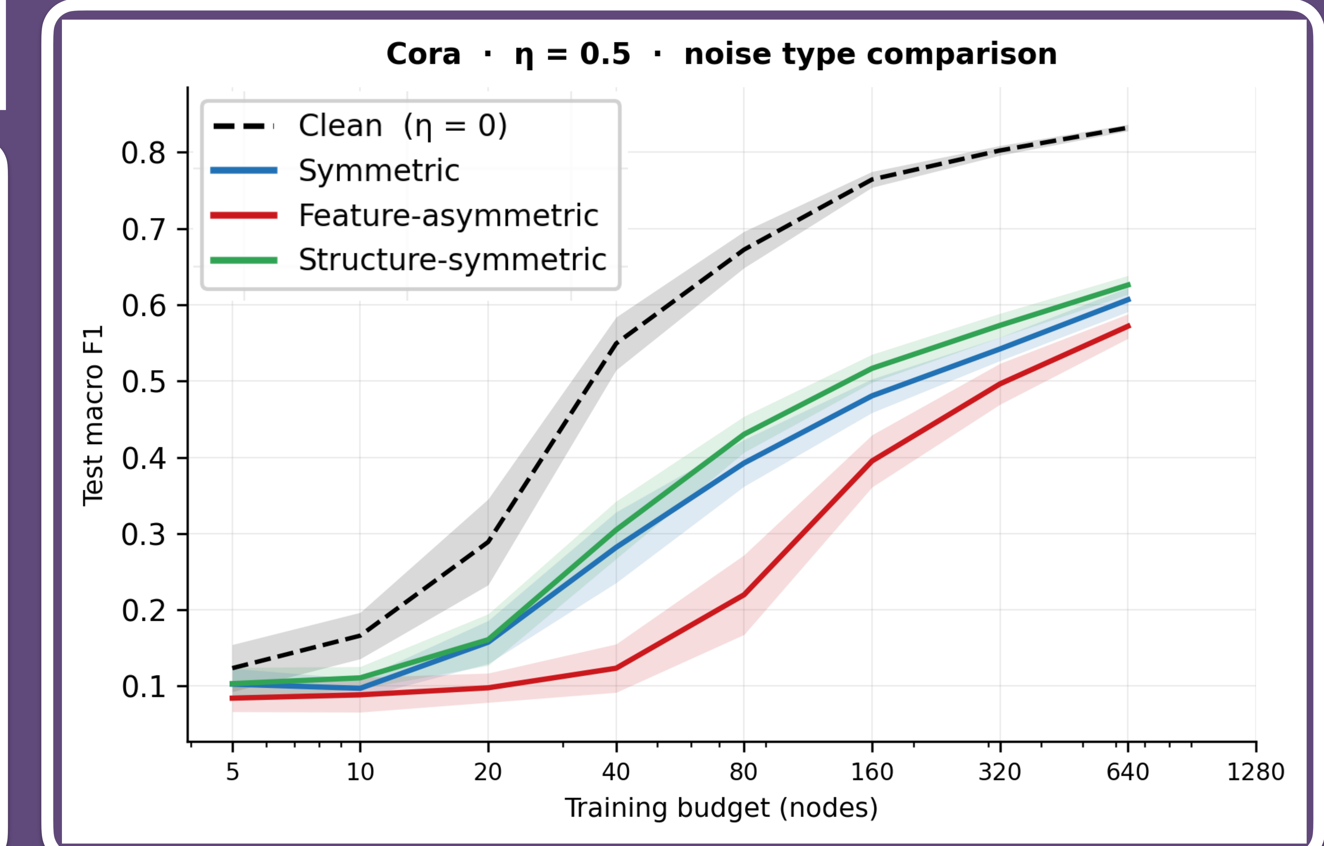
- ChebNet: spectral convolution via K=3 Chebyshev polynomials
- 2 layers · dropout 0.5 · Adam

Protocol

- Only training labels noised - validation, test sets remain clean
- Training budget $n_n \in \{10, 20, 40, 80, 160, 320, 640, 1280\}$
- 20 independent runs per: dataset, noise type, noise level, budget

Results





Dataset	Clean	$\eta = 0.1$ Sym	Feat	Struct	$\eta = 0.3$ Sym	Feat	Struct	$\eta = 0.5$ Sym	Feat	Struct
Cora	0.832	0.807	**0.802**	0.815	0.738	**0.720**	0.736	0.607	**0.572**	0.626
Pubmed	0.844	0.825	**0.815**	0.828	0.786	**0.722**	0.782	0.662	**0.528**	0.651
Chameleon	0.299	0.291	**0.283**	0.298	0.286	**0.218**	0.281	0.259	**0.167**	0.272
Squirrel	0.228	**0.220**	0.224	0.226	0.222	**0.206**	0.216	0.204	**0.189**	0.220

***Macro F1** for max budget for each dataset Bold: Lowest Macro F1 (higher is better)*

Dataset	Clean	$\eta = 0.1$ Sym	Feat	Struct	$\eta = 0.3$ Sym	Feat	Struct	$\eta = 0.5$ Sym	Feat	Struct
Cora	0.549	0.516	**0.463**	0.510	0.442	**0.304**	0.457	0.282	**0.123**	0.305
Pubmed	0.622	0.584	**0.575**	0.595	0.474	**0.442**	0.535	0.372	**0.297**	0.398
Chameleon	0.234	**0.229**	0.240	0.232	0.203	**0.157**	0.208	0.166	**0.111**	0.181
Squirrel	0.149	0.160	**0.144**	0.146	0.155	**0.135**	0.148	0.148	**0.123**	0.138

***Macro F1** for budget 40 for each dataset Bold: Lowest Macro F1 (higher is better)*

Analysis

Noise changes shape

- Gap between noisy and clean performance widens - invest in label cleaning

Feature-asymmetric noise - most harmful

- Corrupts highest value training signal at class boundary

Structure-symmetric noise - least harmful

- Targets already degraded representations on homophilic graphs

Architecture bottleneck - noise is irrelevant

- Heterophilic graphs already challenge GNNs - invest in different architectures

Conclusion

Summary:

- Investigated how noise affects GNN learning curves - ChebNet
- Studied 3 types of label noise

Key finding - Noise matters:

- Flattening curve - diminishing benefit of adding more noisy labels
- Noise type is important

Practical implications:

- For homophilic graphs: invest in label cleaning
- For heterophilic graphs: invest in different architectural choices

Future work:

- More noise types
- Different architectures