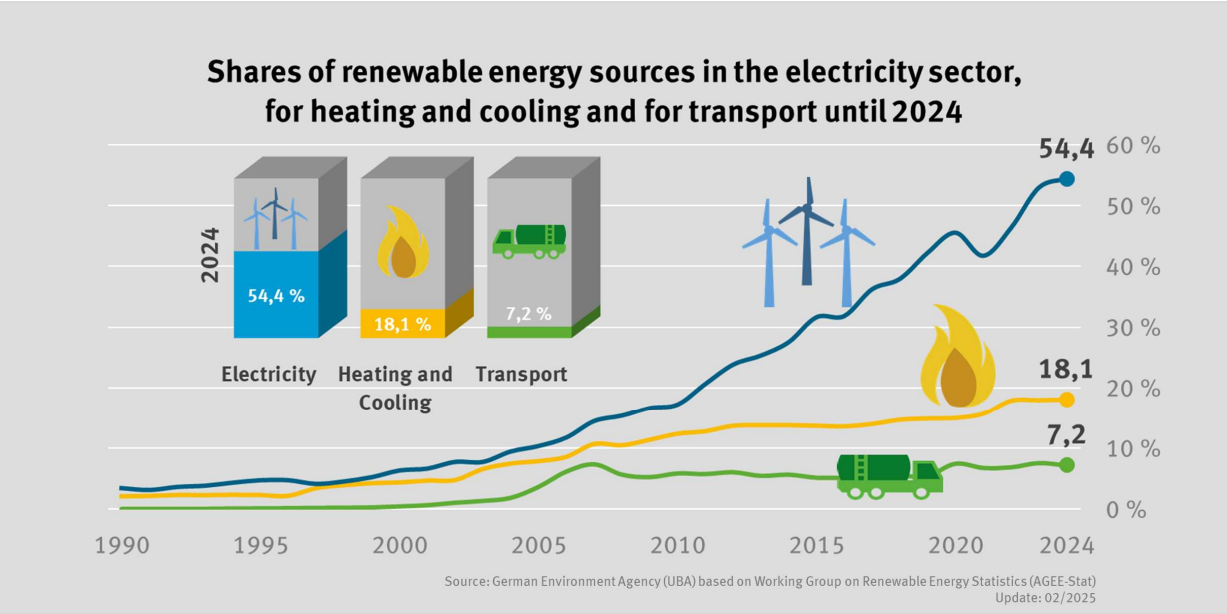
Exploring Variational Graph Autoencoders for Distribution Grid Data Generation

Syed Zain Abbas
Ehimare Okoyomon

Tackling Climate Change with Machine Learning workshop
at NeurIPS 2025

Problem Statement

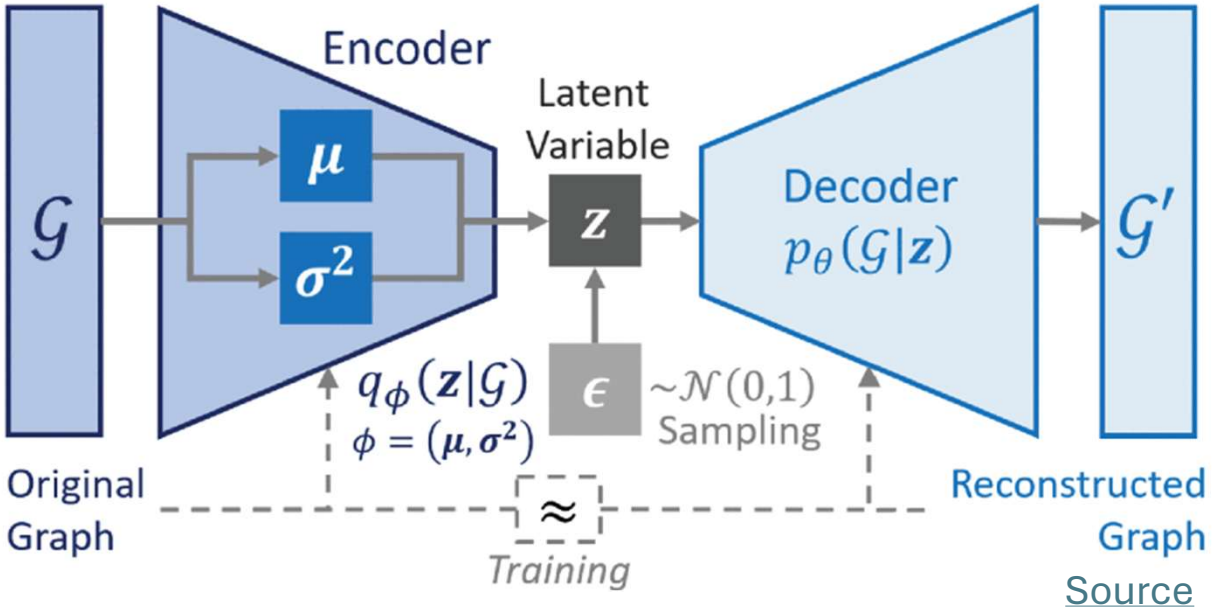
- Increase in Renewable Energy Resources
- Dangerous for Grid Stability
- Lack of Public Power System Data
- Machine Learning can be used to solve this problem



Methodology

VGAE

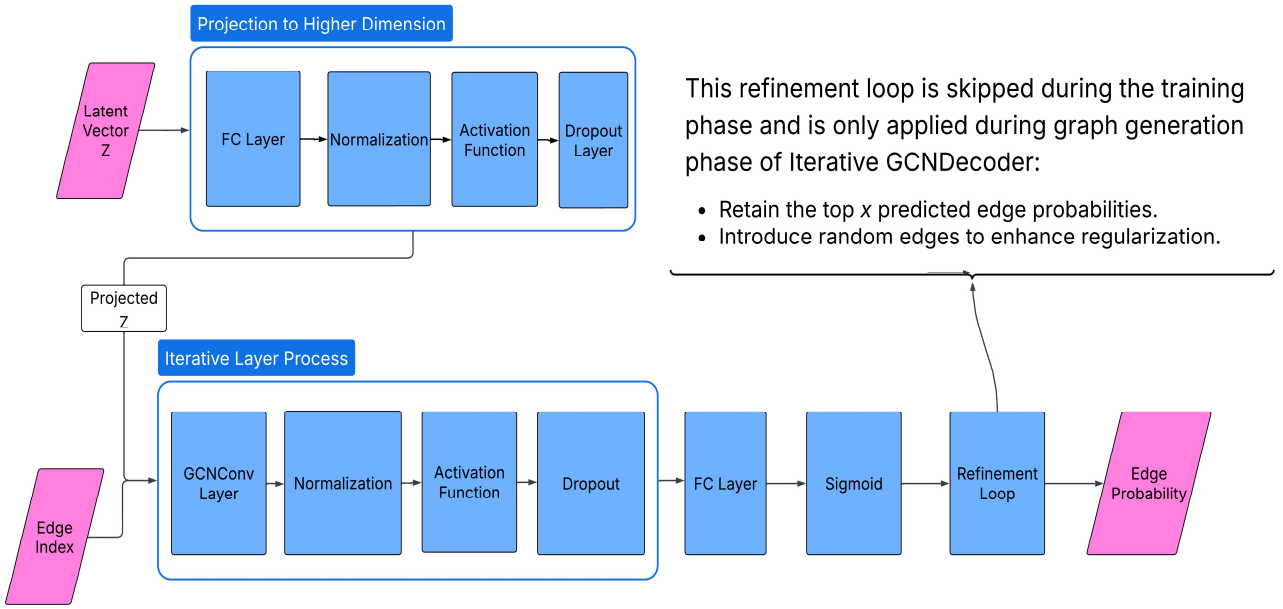
- Encoder generates a latent representation
- Decoder tries to reconstruct the graph from the latent representation



Source

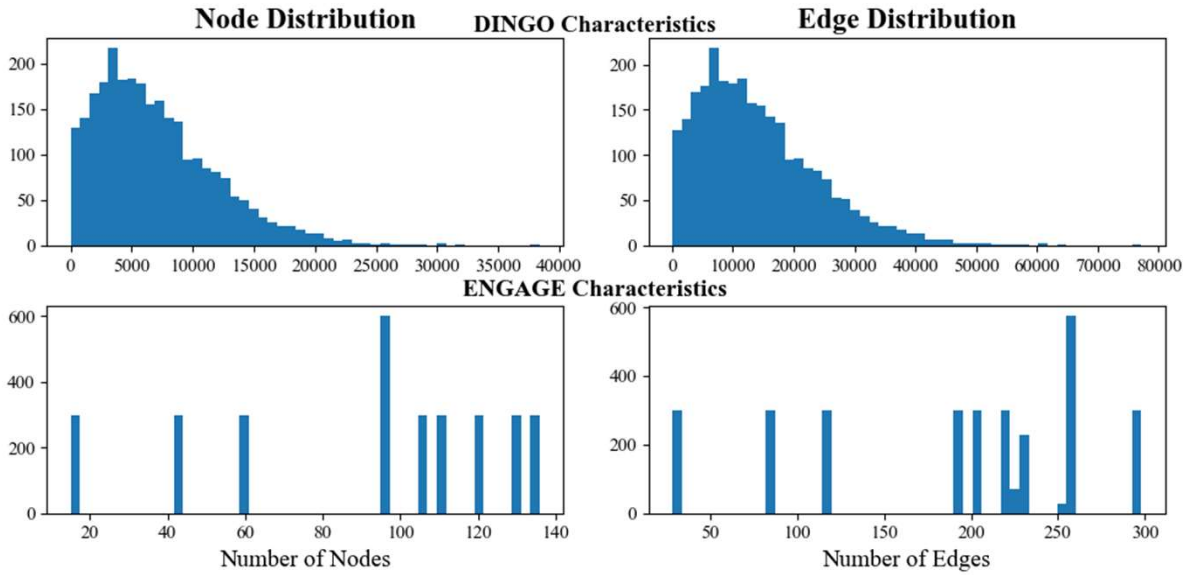
Model Architecture

- GCN Encoder
- Inner Product Decoder
- MLP Decoder
- GCN Decoder
- Iterative GCN Decoder



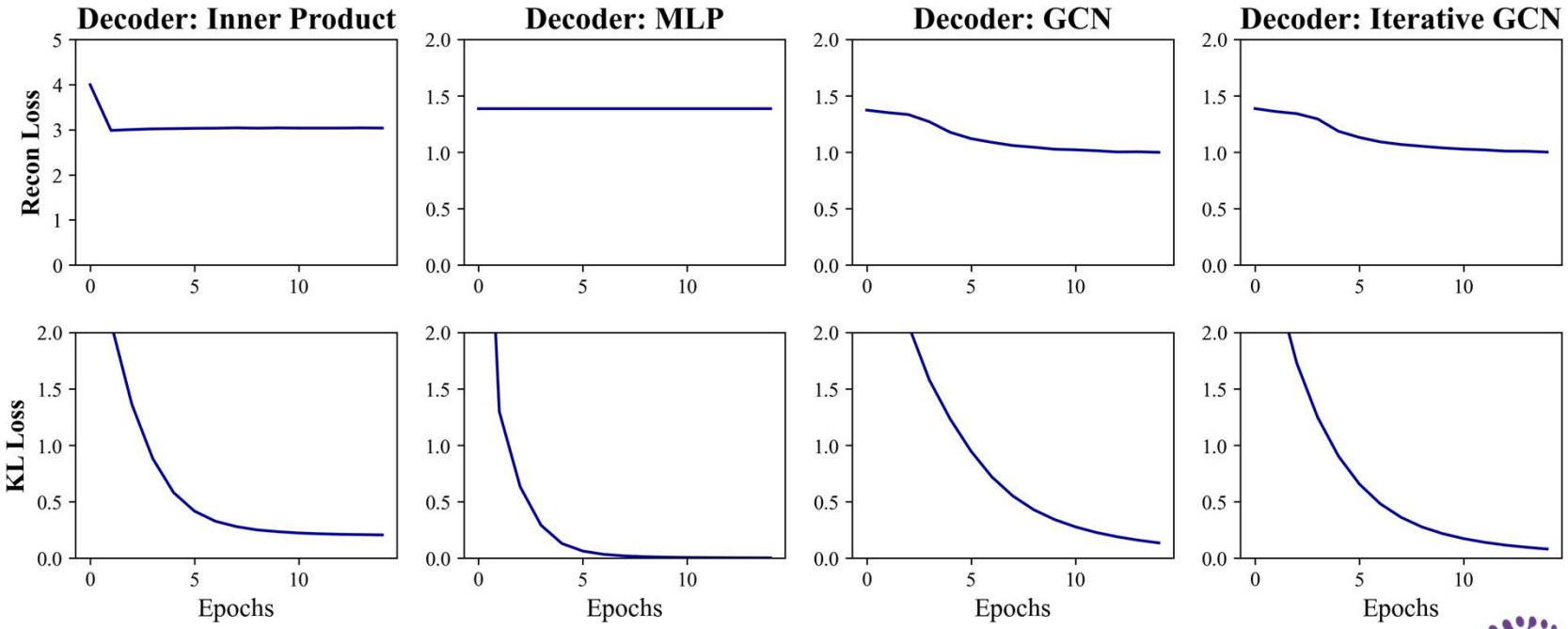
Data Characteristic

- ENGAGE is based on Simbench and contains small and less diverse grids
- DINGO, on the other hand, is more diverse

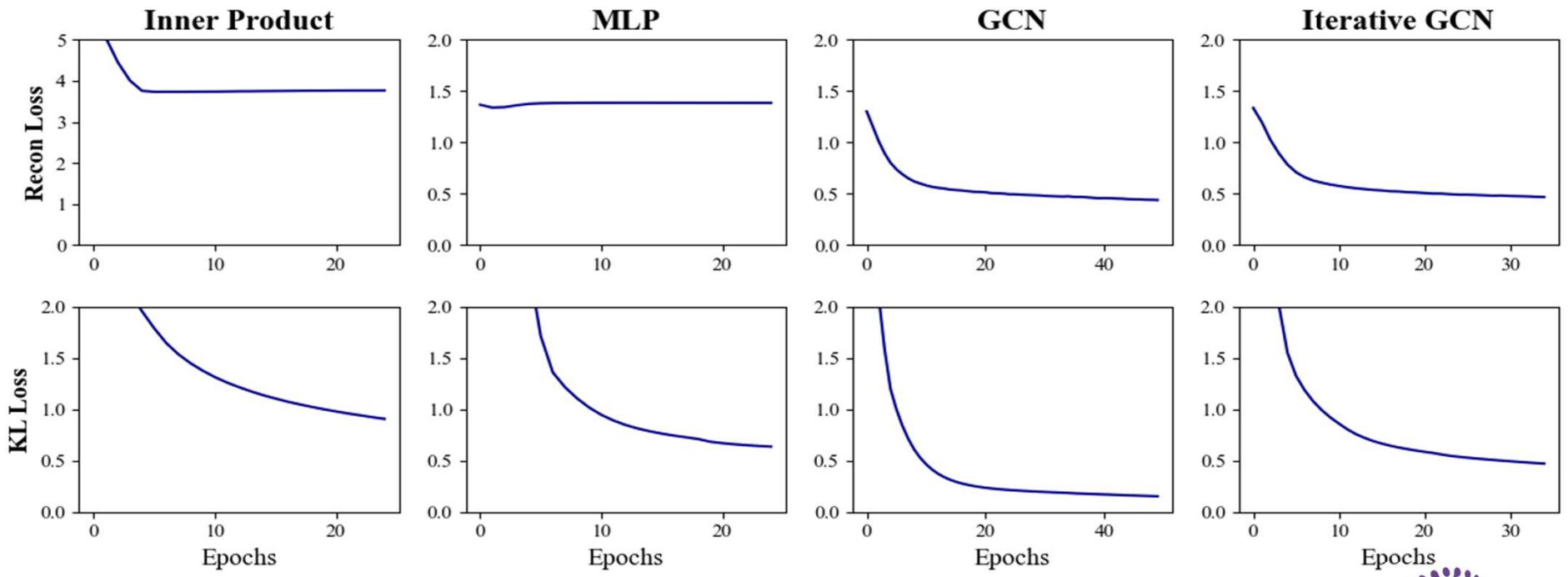


Results

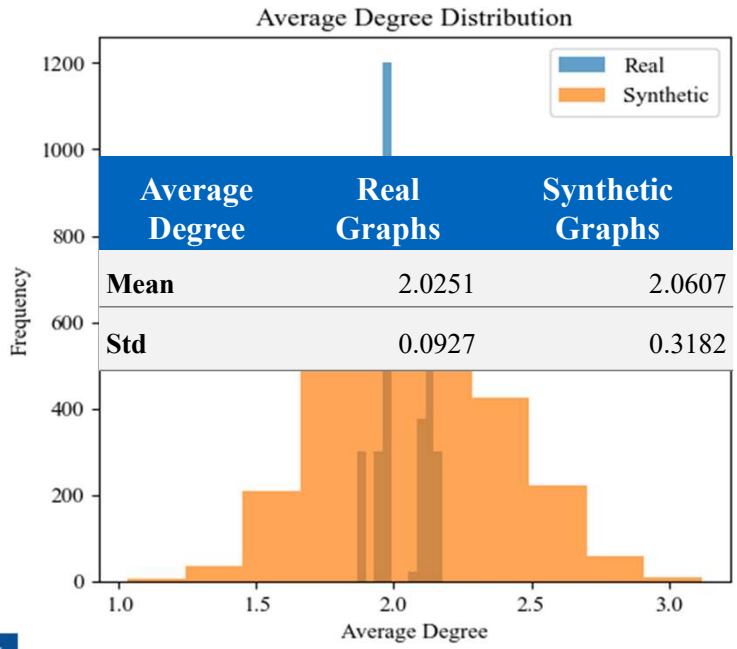
Training Losses – ENGAGE



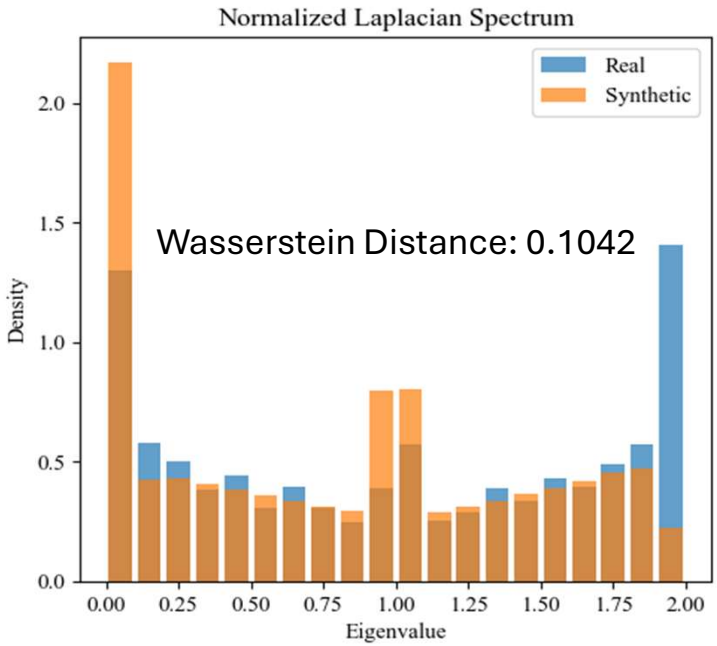
Training Losses – DINGO



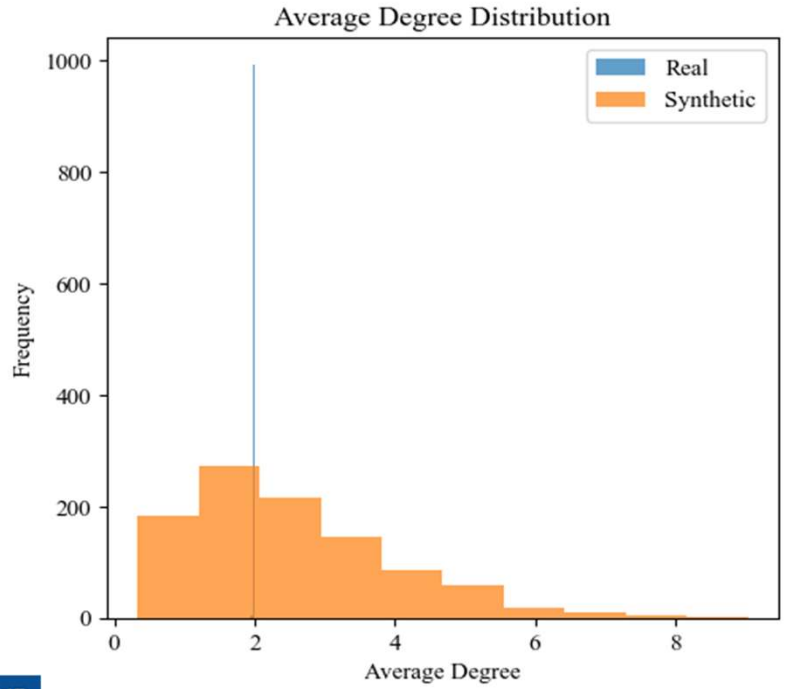
Performance Evaluation – ENGAGE

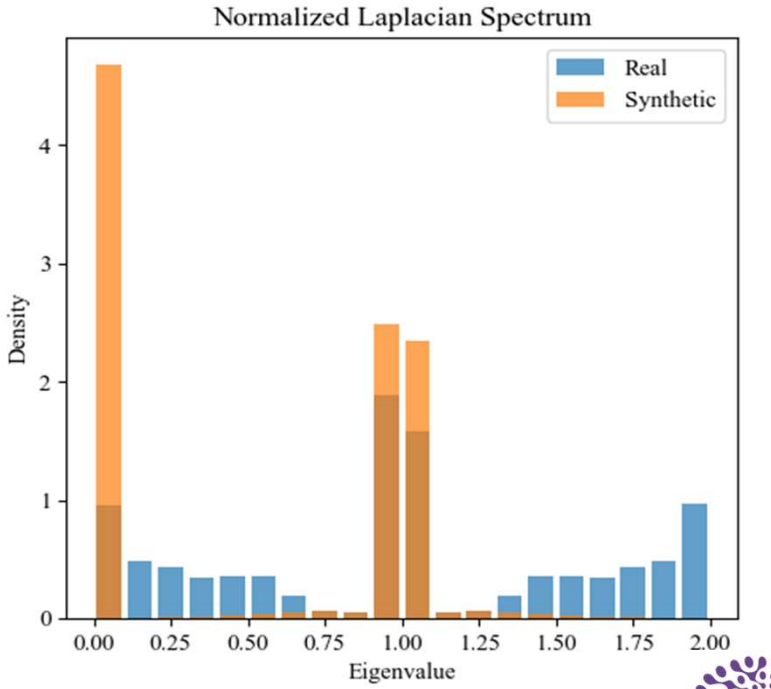


Average Degree	Real Graphs	Synthetic Graphs
Mean	2.0251	2.0607
Std	0.0927	0.3182



Performance Evaluation – DINGO





Limitations

Training Instability:

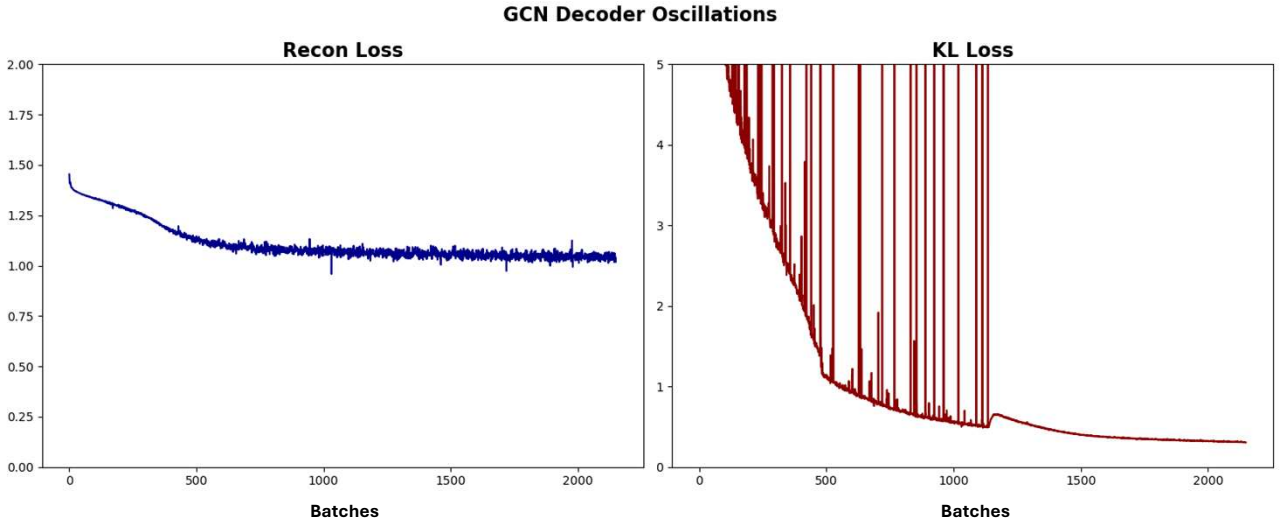
- Exhibits sensitivity to hyperparameter selection
- Prone to loss oscillation during training phases

Computational Challenges:

- Performance degradation with large-scale graphs
- Deeper architectures required for complex graphs increase computational demands
- Susceptibility to over-smoothing in deeper configurations

Data Availability Constraints:

- Limited by the availability of suitable training datasets
- Can be addressed through hybrid approaches combining simpler models with domain-specific heuristics to generate supplementary training data



Average Degree	Real Graphs	Synthetic Graphs
Mean	1.9988	4,0036
Std	0.01	5,0084

Conclusion & Outlook

Key Research Findings

- Successfully implemented PyTorch Geometric's VGAE framework for synthetic graph generation
- Generated graphs demonstrate similarity to real-world network structures with appropriate hyperparameter optimization

Future Research Directions

Exploration of advanced generative frameworks:

- Evaluation of moderate-complexity approaches such as GANs
- State-of-the-art diffusion models for enhanced graph generation
- Investigation of hybrid architectures combining VGAE with diffusion models to improve link prediction capabilities

Thank You!