

Is Graph Mixup Beneficial?

Investigating Interpolation And Empirical Performance of Graph Mixup Methods

Simon Forbat, Rainer Gemulla

Mannheim, Germany

Is graph mixup beneficial?

Our study:

We **analyzed** graph mixup methods across popular datasets and models.

Key contributions:

- Independent evaluation of graph mixup
- **Interpolation analysis** of mixup graphs

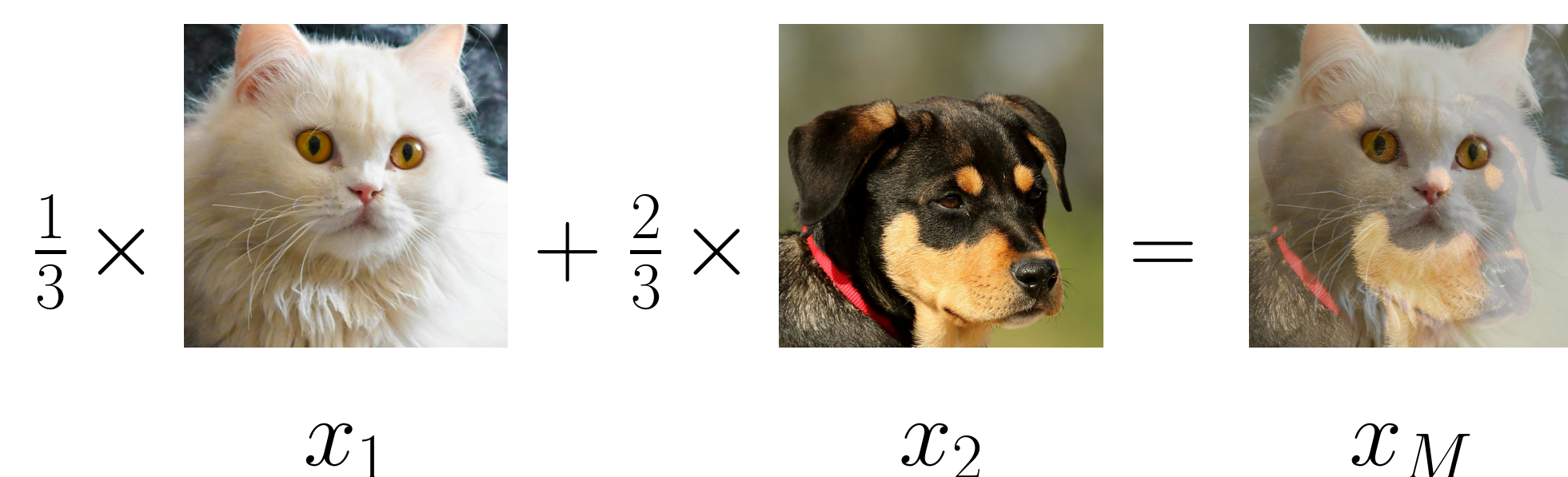
Key findings:

- **No significant gains** over the baseline
- Most methods produced **bad interpolations**
- Even **optimal interpolation** did not significantly improve performance

Mixup: why and what?

Benefits:

- Encourages smoother decision boundaries
- Reduces overfitting
- Improves generalization
- Highly effective in CV & NLP



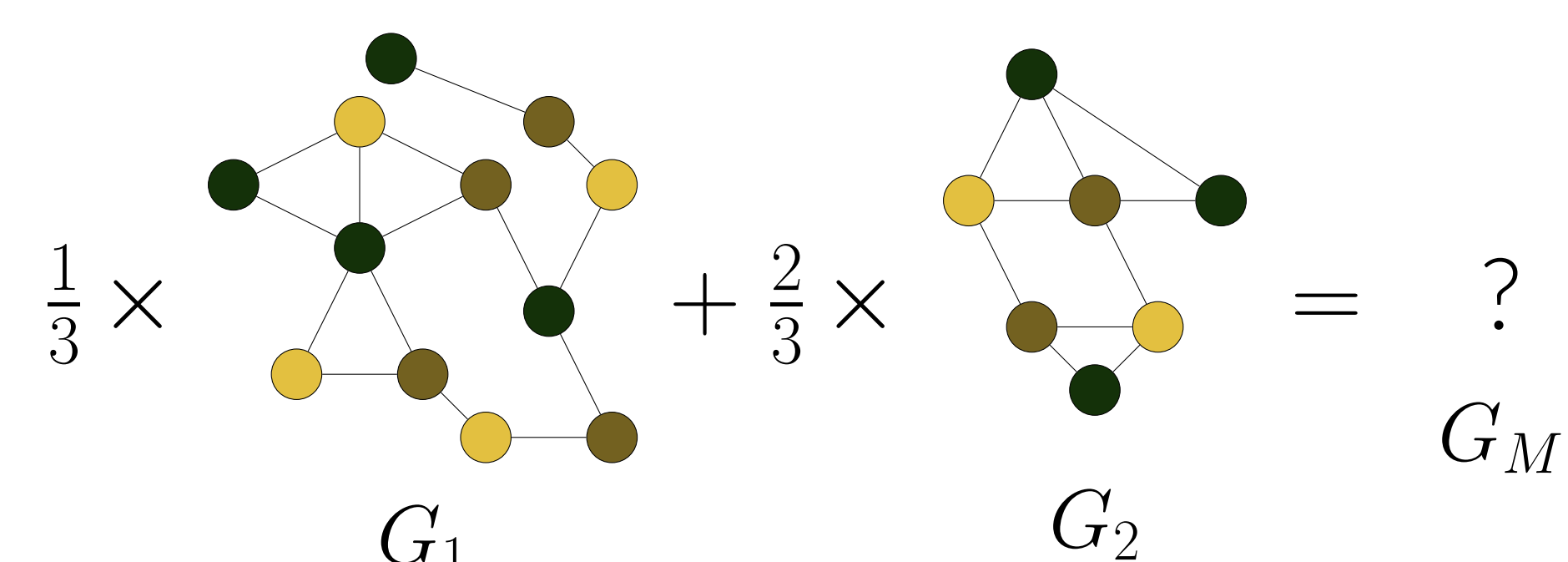
Definition: Construct a synthetic item (x_M, y_M)

$$x_M = (1-\lambda) x_1 + \lambda x_2,$$

$$y_M = (1-\lambda) y_1 + \lambda y_2,$$

where λ is sampled from $[0, 1]$.

Mixup for graphs

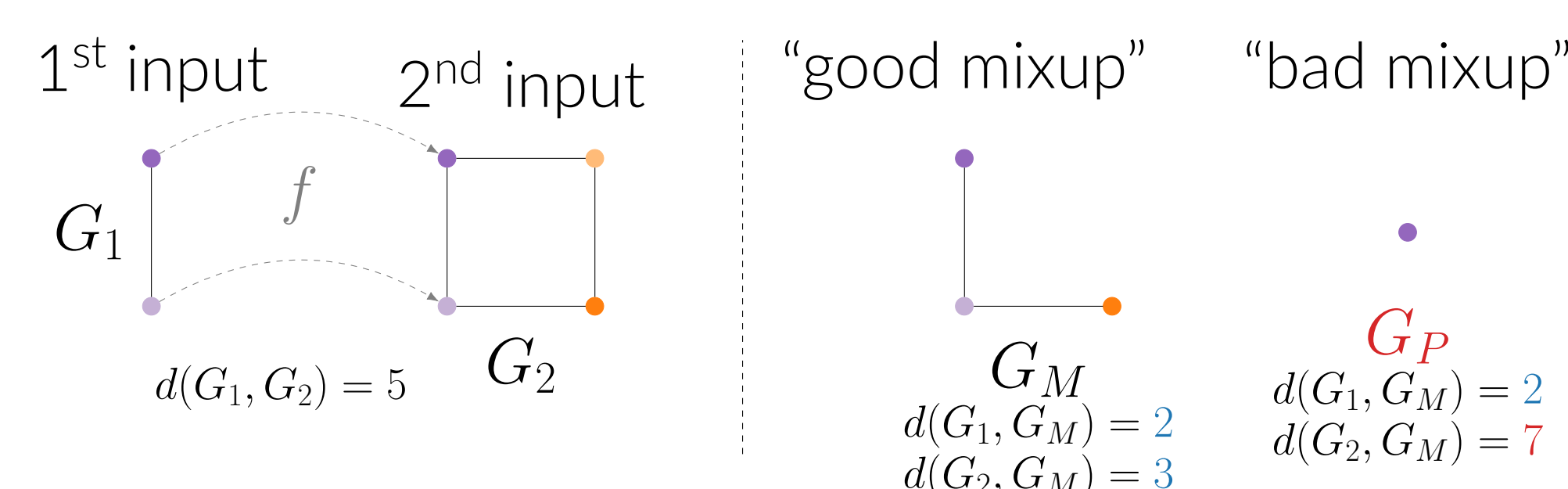
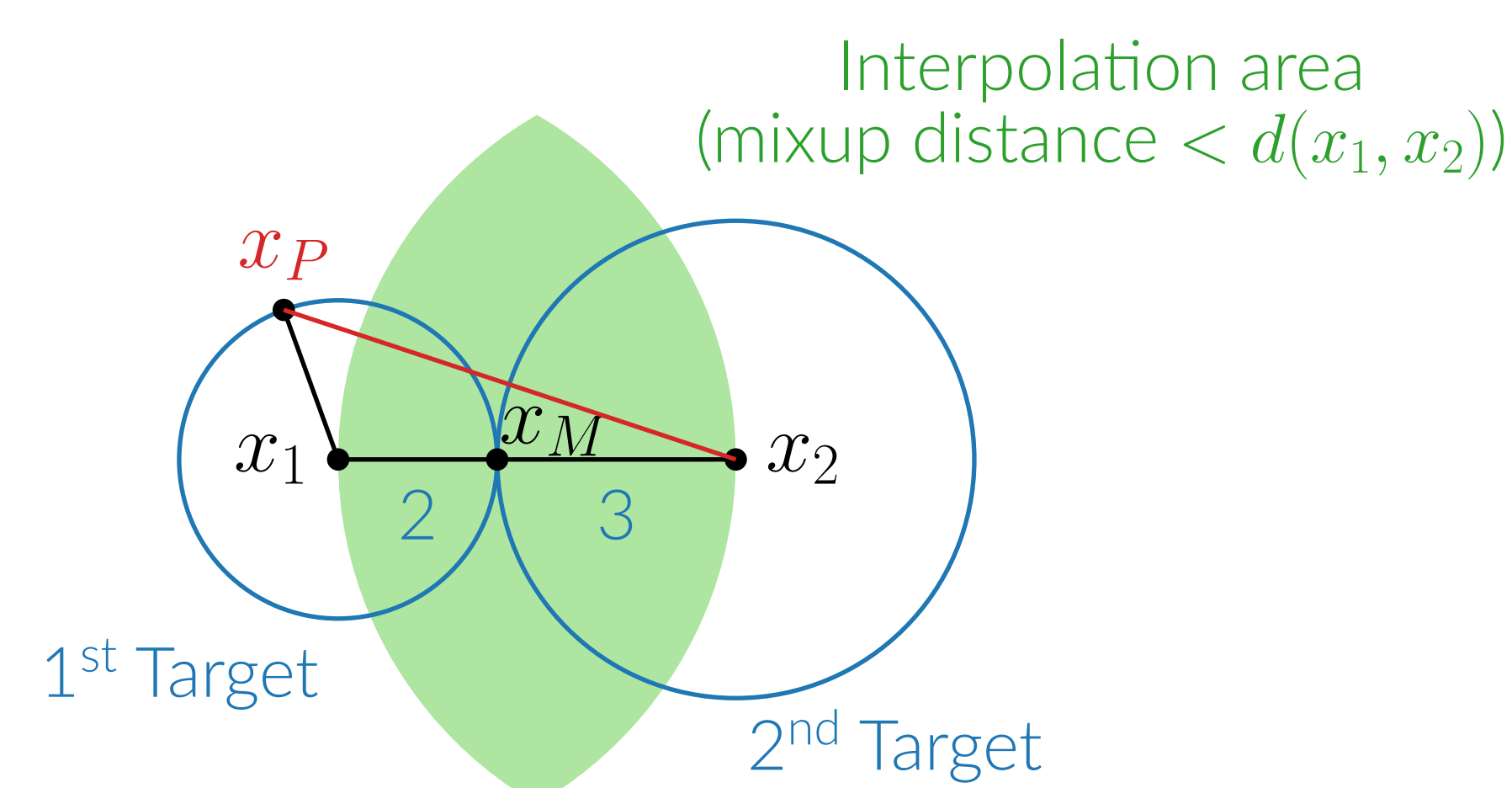


- Graph datasets are **small** and GNNs **overfit**
- Many methods proposed: Embedding Mixup, FGW-Mixup, G-Mixup, GeoMix, If-Mixup, S-Mixup, SubMix, and others
- **No independent evaluation** exists
- Reproducibility & evaluation issues raised

Mixup aims to interpolate

In theory:

- (Graph) mixup methods are **designed to interpolate** between their inputs
- Distances in the graph domain are based on the graph edit distance (GED)



Many methods interpolated poorly

In practice:

- Many graph mixup methods interpolated poorly as they **did not meet their targets**
- Achieving optimal interpolation is possible

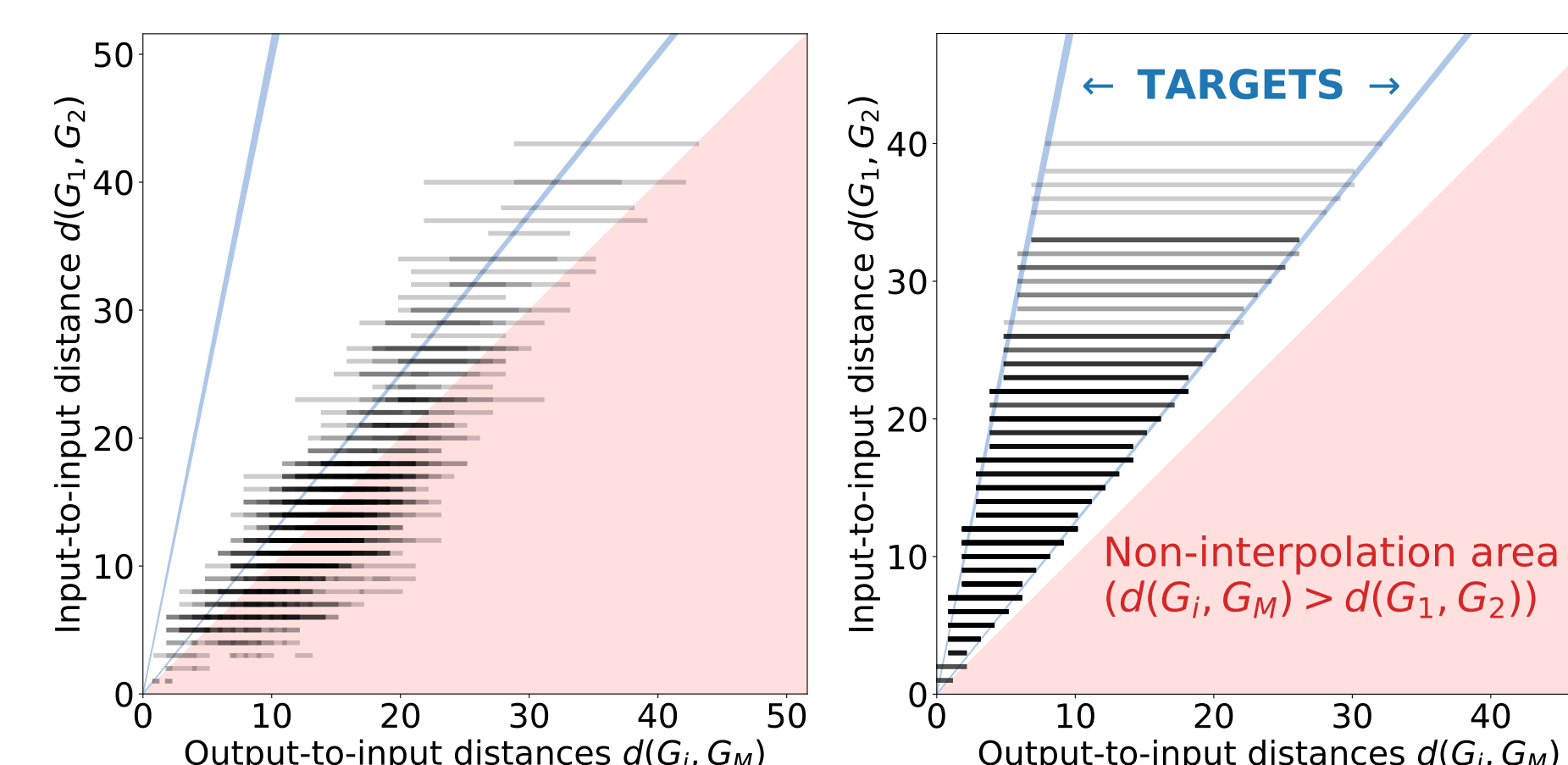


Figure: Distance from mixup graph to its inputs vs. distance between the inputs

Good interpolation helped but was insufficient

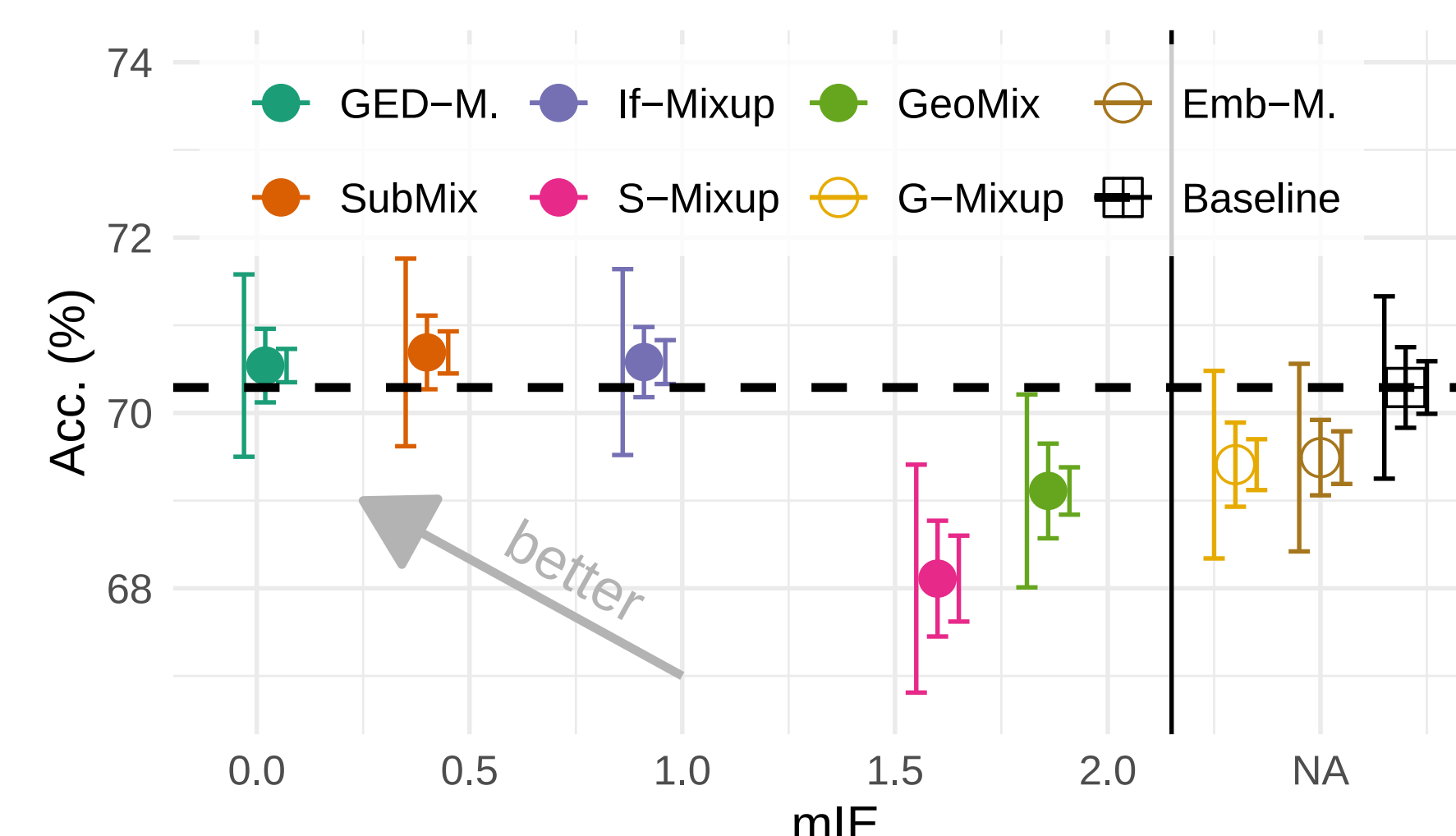


Figure: Accuracy vs. interpolation error

- No method significantly outperformed the no-mixup baseline
- Methods which interpolated better also performed better

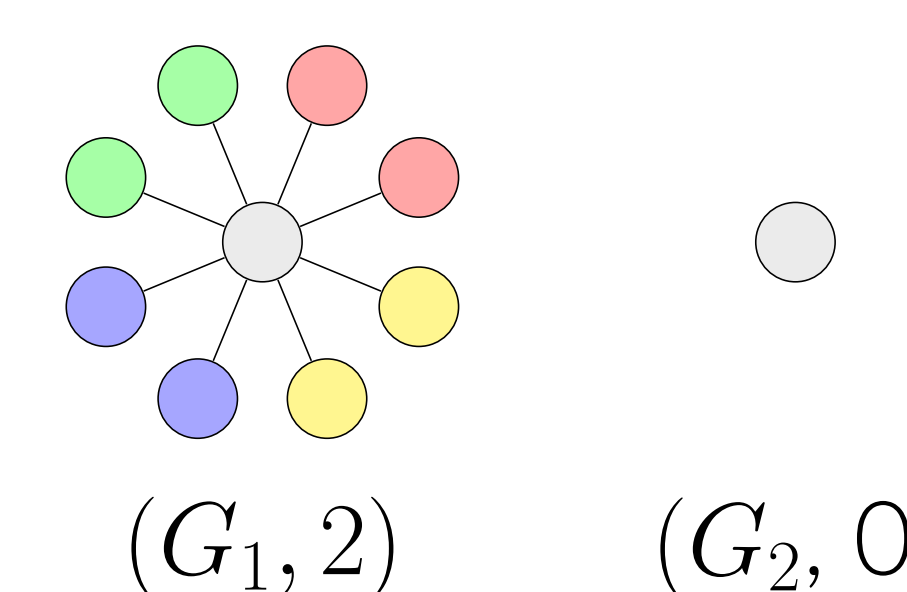
The road ahead

Problem:

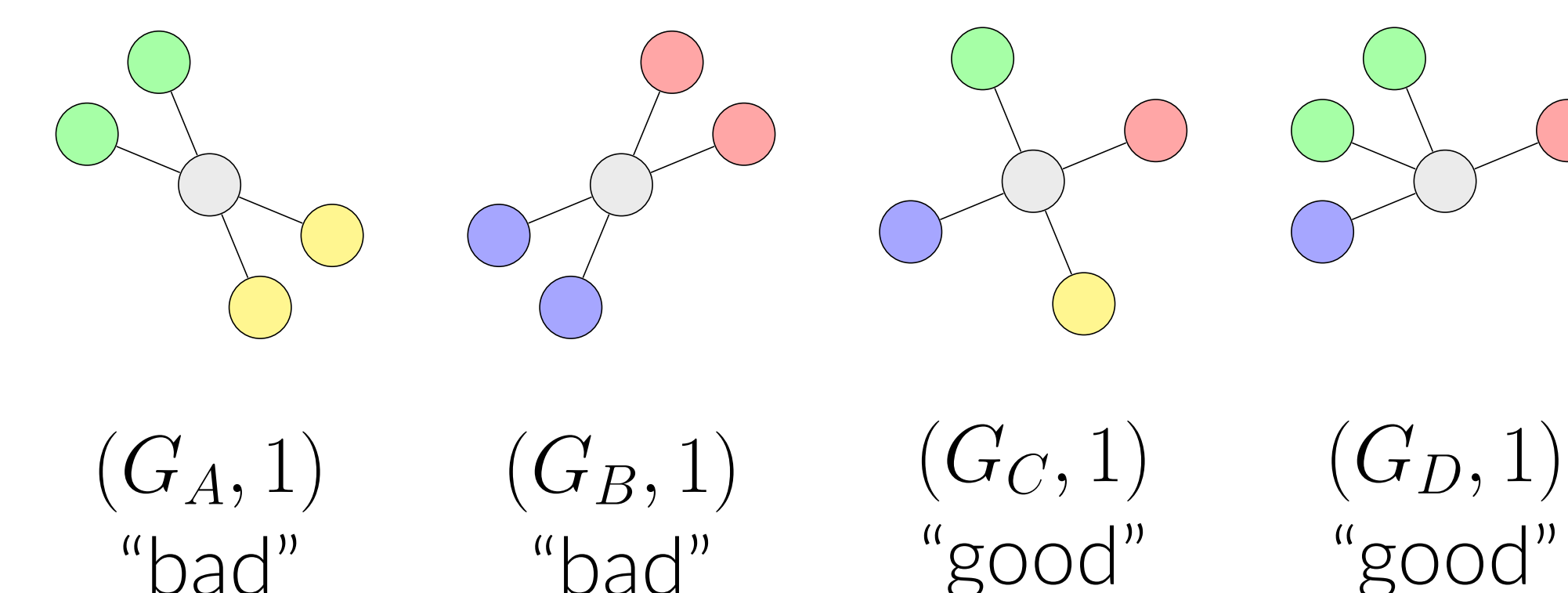
Only a subset of optimally interpolated mixup graphs is beneficial → how can they be **selected**?

Example: Predict the number of blue nodes.

Inputs:



Optimally interpolated mixup graphs:



How do we fix this? — Promising direction for future work!

Check out the full paper



<https://icml.cc/virtual/2026/poster/64758>