Quantile-Free Uncertainty Quantification in Graph Neural Networks

Soyoung Park[1], Hwanjun Song[2], Sungsu Lim[1]
[1]Chungnam National University, [2]KAIST

ICML 2026

Contents

Graphs and GNNs

- Graphs model connected entities, and GNNs learn from their relationships

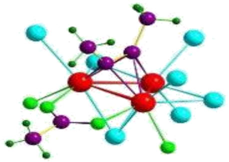
$$G = (V, E), \qquad V = nodes/entities, \qquad E = edges/relationships$$

- GNNs learn node representations through message passing

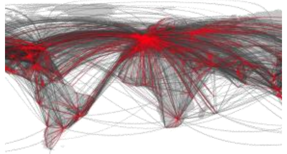
$$h_v^{(k)} = \mathrm{UPDATE}(h_v^{(k-1)}, \mathrm{AGG}\{h_u^{(k-1)} : u \in \mathcal{N}(v)\})$$

Social Graphs

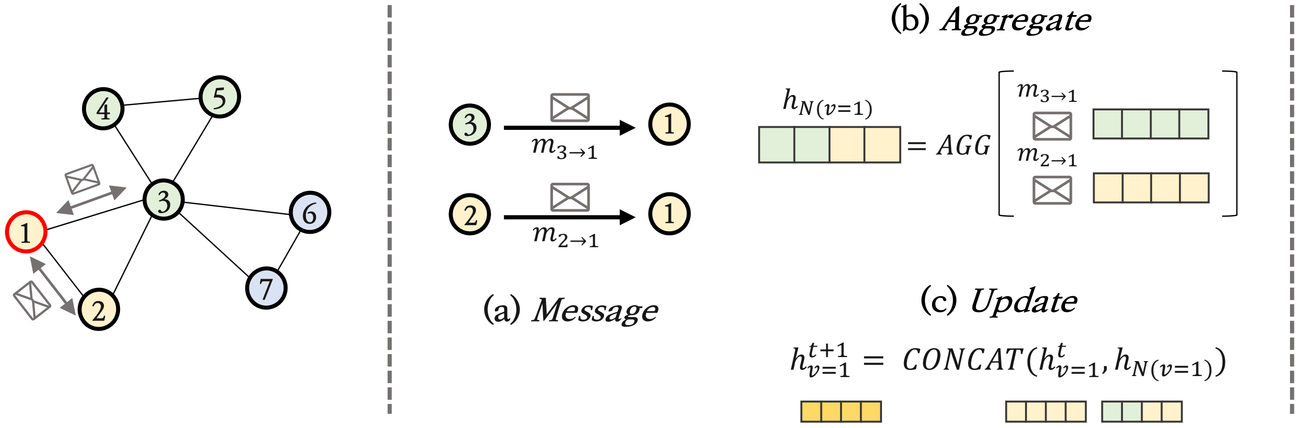
Web Graphs



Molecular Graphs



Transportation Graphs



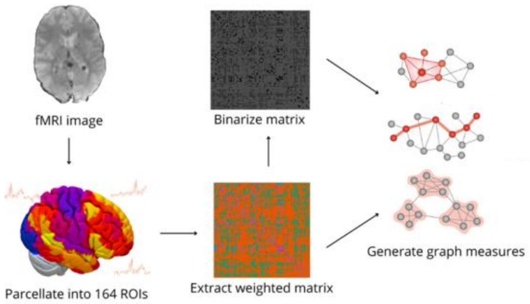
GNNs in High-Stakes Application

- Graph-structured data appears in decision-critical domains

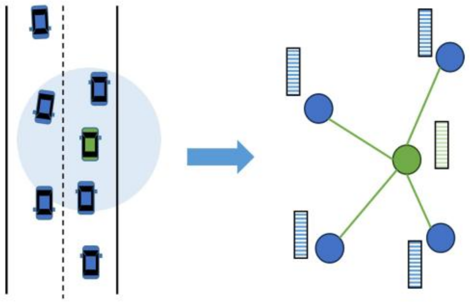
$$features + topology \rightarrow context - aware\ prediction$$

- GNNs exploit relational structure to improve predictions
- But in high-stakes settings, accuracy alone is not enough

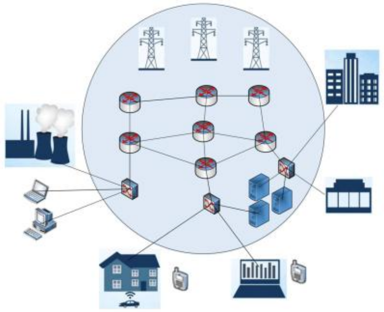
Financial Fraud Detection



Medical Diagnosis



Autonomous Driving



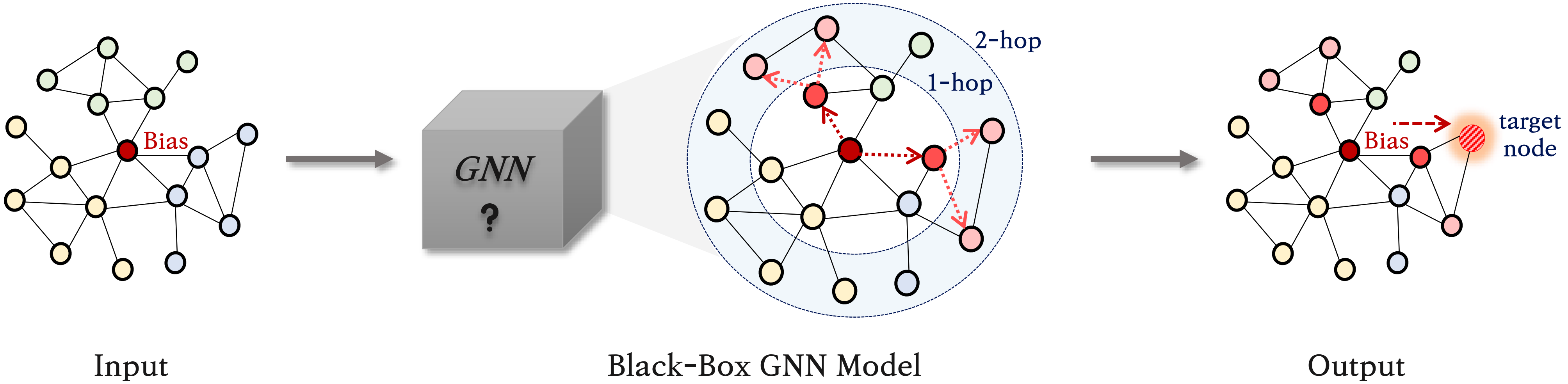
Power Grid Management

Problem of Message Passing

- Message passing can propagate errors across connected nodes

noisy node → neighbor aggregation → uncertain prediction

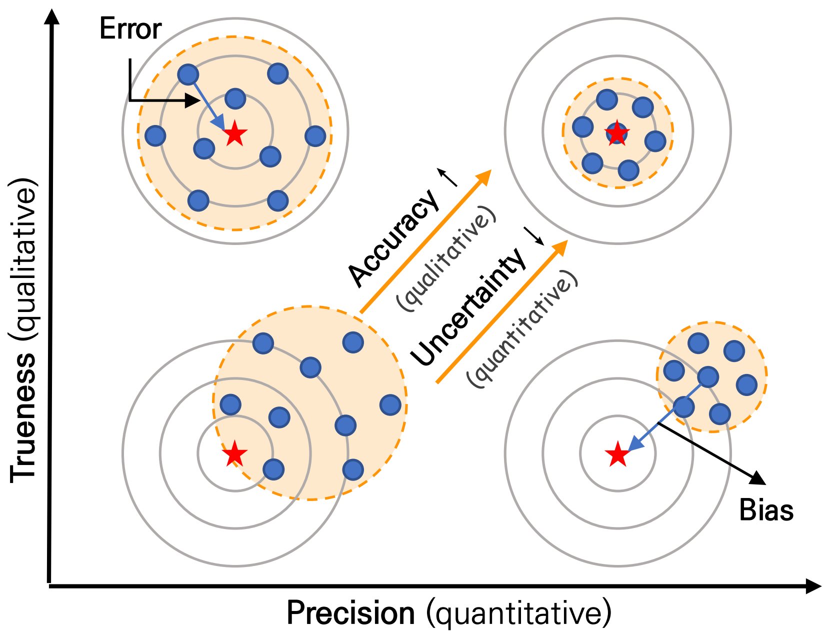
- Noisy features can spread to neighbors
- Incorrect edges can affect connected nodes

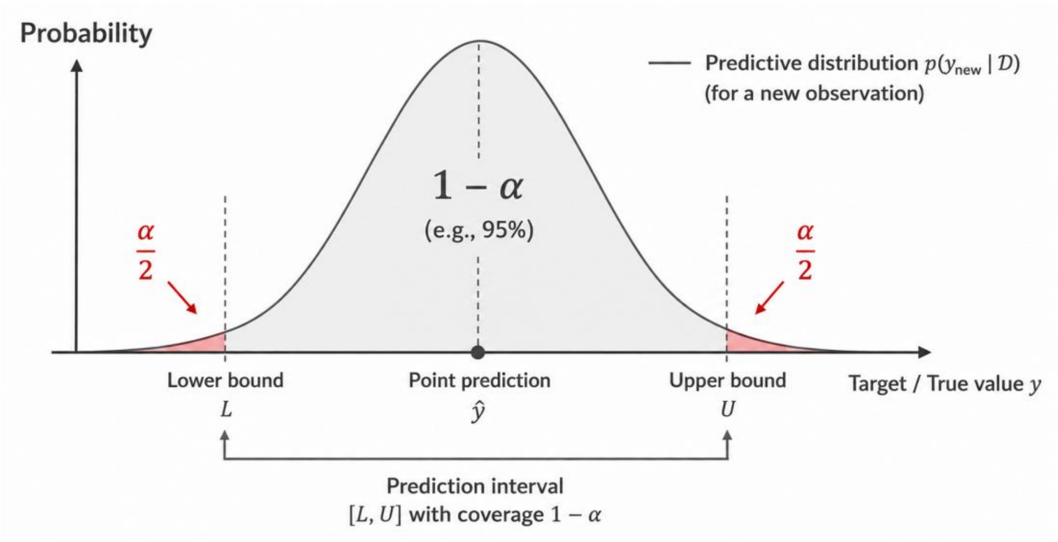


Uncertainty Quantification (UQ)

- Point prediction alone does not show confidence
- A prediction interval provides a plausible range where the true value is expected to lie

$$\hat{y}_v \rightarrow [\hat{y}_v^{low}, \hat{y}_v^{up}], \qquad Good\ prediction\ inteval \rightarrow calibrated + compact$$





Contents

Node-wise Prediction Intervals

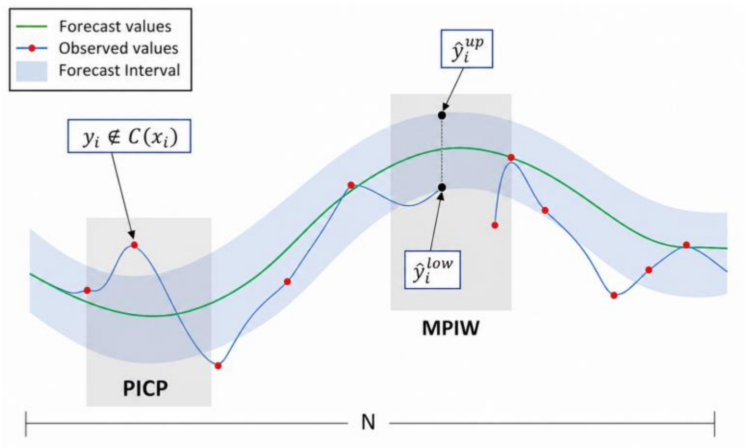
- Goal: estimate reliable uncertainty for each node-level regreesion target
  - ➢ Instead of predicting only a point value $\hat{y}_v$, we predict a node-wise interval $[\hat{y}_v^{low}, \hat{y}_v^{up}]$
  - ➢ A useful interval should satisfy $P(\hat{y}_v^{low} \leq y_v \leq \hat{y}_v^{up}) \geq 1 - \alpha \quad with \quad \mathbb{E}[\hat{y}_v^{up} - \hat{y}_v^{low}] \; minimized$

- In experiments, these are measure by
  - ➢ Prediction Interval Coverage Probability (PICP)

$$\frac{1}{N}\sum_{i=1}^{N} \mathbb{I}(y_i \in C(x_i))$$

  - ➢ Mean Prediction Interval Width (MPIW)

$$\frac{1}{N}\sum_{i=1}^{N} |\hat{y}_i^{up} - \hat{y}_i^{low}|$$



Assumption Burden in UQ

- UQ methods differ in how much extra assumption or procedure they require
  - Bayesian/Ensemble → priors, posterior approximation, or sampling
  - Conformal Prediction (CP) → exchangeability and post-hoc calibration

- Quantile Regression (QR) learns lower and upper quantiles and provides direct interval

$$C(x) = [\hat{q}_{\alpha/2}(x), \hat{q}_{1-\alpha/2}(x)]$$

- Compared to Bayesian and CP methods, QR turns interval estimation into a direct learning problem, rather than relying on model sampling or post-hoc calibration

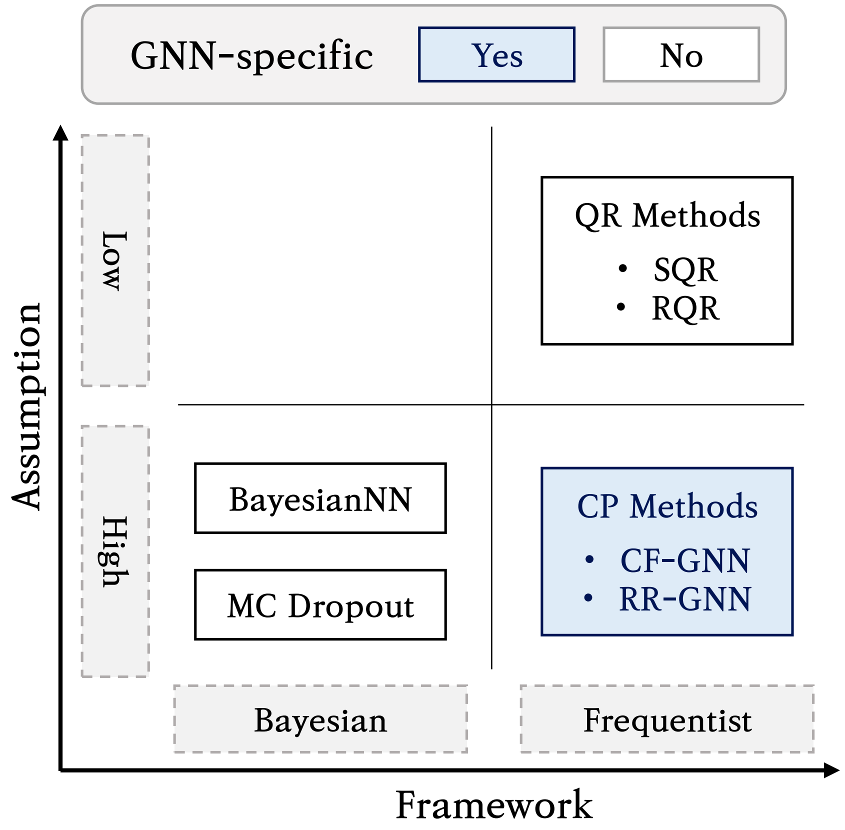
QR for GNNs

- QR is direct and assumption-light, but not graph-native
- Existing QR methods can be combined with GNN embeddings

$$h_v = GNN\big(x_v, \mathcal{N}(v)\big)$$

$$C_v = [\hat{q}_{\alpha/2}(h_v), \hat{q}_{1-\alpha/2}(h_v)]$$

- But message passing makes node intevals dependent
  - ➢ Quantile Input Dependency
  - ➢ Quantile Crossing
  - ➢ Over-smoothing



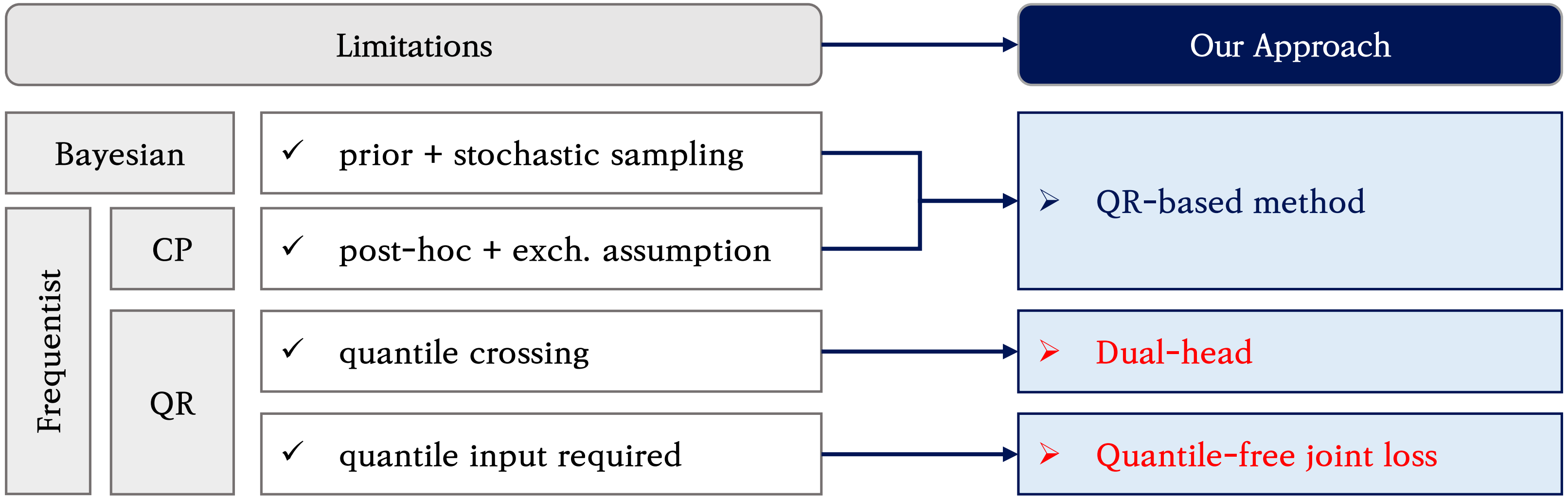
Contents

QpiGNN Framework

- QpiGNN learns node-wise prediction intervals directly

$$H_v \rightarrow (\hat{y}_v, \hat{d}_v) \rightarrow C_v$$

- Instead of predicting multiple quantiles, QpiGNN predicts a center $\hat{y}_v$ and half width $\hat{d}_v$
- This enables quantile-free interval learning directly from graph-awre representations

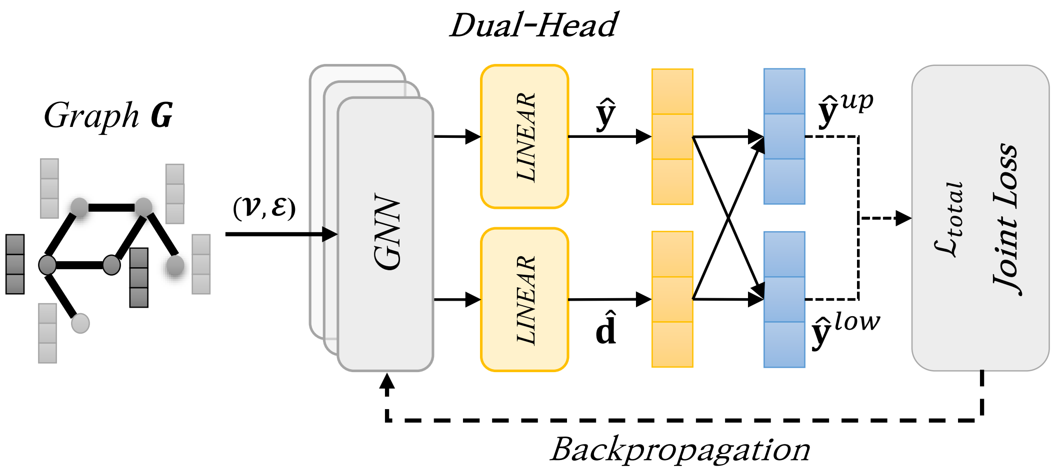


Dual Head Architecture

- The dual head separates prediction and uncertainty after message passing
- A GNN encoder first learns graph-aware node representations $H_v = GNN(X, A)$
- Then two heads predict the center and uncertainty width

$$\hat{y}_v = W_{pred} H_v + b_{pred}, \qquad \hat{d}_v = Softplus\left(W_{diff} H_v + b_{diff}\right)$$

- The interval is constructed as $C_v = [\hat{y}_v - \hat{d}_v,\ \hat{y}_v + \hat{d}_v]$

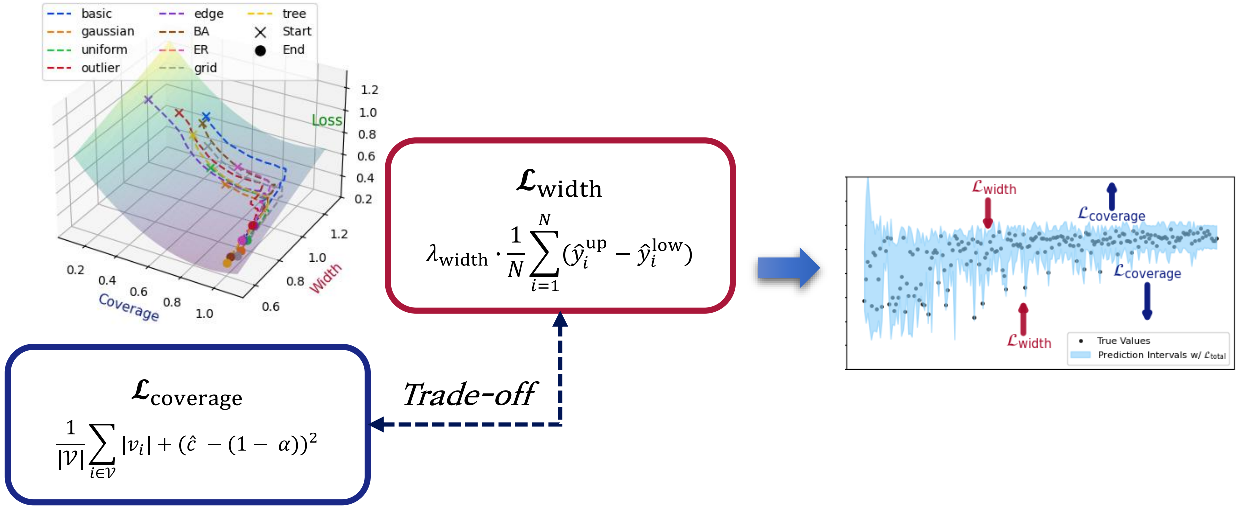


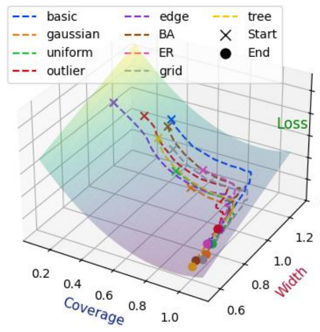
Quantile-Free Joint Loss

- QpiGNN learns calibrated and compact intervals without quantile levels
- Quantile-free joint objective with empirical coverage $\hat{c}$

$$\mathcal{L}_{total} = \underbrace{(\hat{c} - (1-\alpha))^2 + \hat{\ell}_{viol}}_{\mathcal{L}_{\text{coverage}}} + \lambda_{width} \cdot \underbrace{\mathbb{E}_{v\sim V}\left[\hat{y}_v^{up} - \hat{y}_v^{low}\right]}_{\mathcal{L}_{\text{width}}}, \qquad \hat{c} = \mathbb{P}\left(\hat{y}^{low} \le y \le \hat{y}^{up}\right)$$

- The coverage term makes intervals valid, while the width term keeps them informative





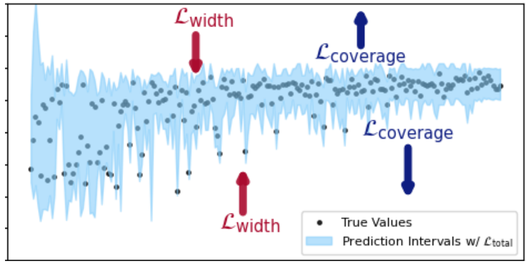
$\mathcal{L}_{\text{width}}$

$$\lambda_{\text{width}} \cdot \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i^{\text{up}} - \hat{y}_i^{\text{low}})$$

Trade-off

$\mathcal{L}_{\text{coverage}}$

$$\frac{1}{|\mathcal{V}|}\sum_{i\in\mathcal{V}}|v_i| + (\hat{c} - (1-\alpha))^2$$



Contents

Asymptotic Coverage Guarantee

- Empirical coverage approaches the target coverage $1-\alpha$
  - Z_v is bounded, so WLLN connects empirical coverage to expected coverage
  - Under mild assumptions, empirical coverage converges to the target level
    - bounded label noise and weak graph dependency
    - consistent prediction $\hat{y}_v$ and $\hat{d}_v$

$$Z_v = 1\left[\hat{y}_v^{low} \leq y_v \leq \hat{y}_v^{up}\right], \qquad \mathbb{E}[Z_v] \to 1-\alpha$$

$$\hat{c} = \frac{1}{N}\sum_{v=1}^{N} Z_v \xrightarrow{P} \mathbb{E}[Z_v] \quad by\, WLLN$$

$$\forall_\epsilon > 0, \lim_{N\to\infty} P(|\hat{c}-(1-\alpha)| > \epsilon) = 0$$

Finite-sample Coverage Stability

- Coverage fluctuation is controlled even for finite graphs
  - $Z_v \in \{0,1\}$, so coverage indicators are bounded
  - Concentration bounds control large deviation
  - δ_G captures extra graph dependency from message passing

$$|\hat{c} - \mathbb{E}[\hat{c}]| = O(\frac{1}{\sqrt{N}})$$

$$P(|\hat{c} - \mathbb{E}[\hat{c}]| > \epsilon) \leq 2\exp(-2N\epsilon^2)$$

$$\sup|\hat{c}(X_1, \ldots, X_v, \ldots, X_N) - \hat{c}(X_1, \ldots, X_v', \ldots, X_N)| \leq \frac{1}{N} + \delta_G$$

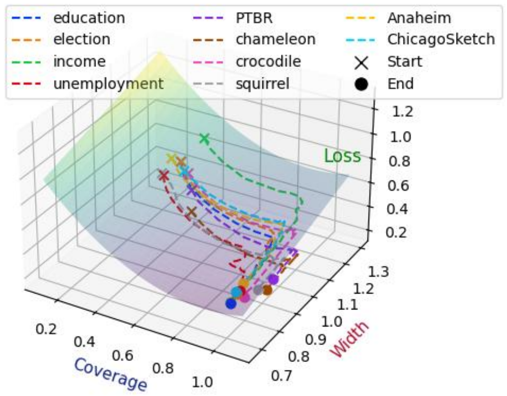
Near-optimal Width under Coverage Constraint

- QpiGNN reduces interval width while maintaining target coverage
  - ➢ Coverage-only intervals can become trivially wide
  - ➢ Under symmetry, optimal width corresponds to the required residual quantile
  - ➢ QpiGNN approximates the constrained objective with a coverage-width loss

$$\min_{\hat{d}_v \geq 0} {}^1/_N \sum_{v=1}^{N} 2\hat{d}_v \;\; s.t.\; P(\hat{y}_v^{low} \leq y_v \leq \hat{y}_v^{up}) \geq 1-\alpha$$

$$\hat{d}_v^* = F_v^{-1}\left(1-\frac{\alpha}{2}\right), \qquad [\hat{y}_v - \hat{d}_v^*, \hat{y}_v + \hat{d}_v^*]$$

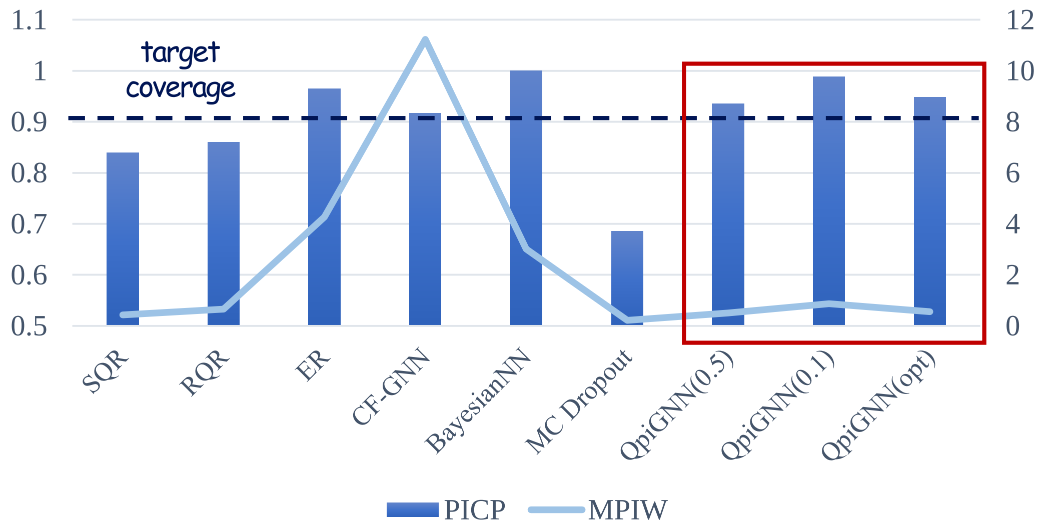
$$\mathcal{L}_{total} = \mathcal{L}_{coverage} + \lambda_{width} \cdot \mathcal{L}_{width}$$

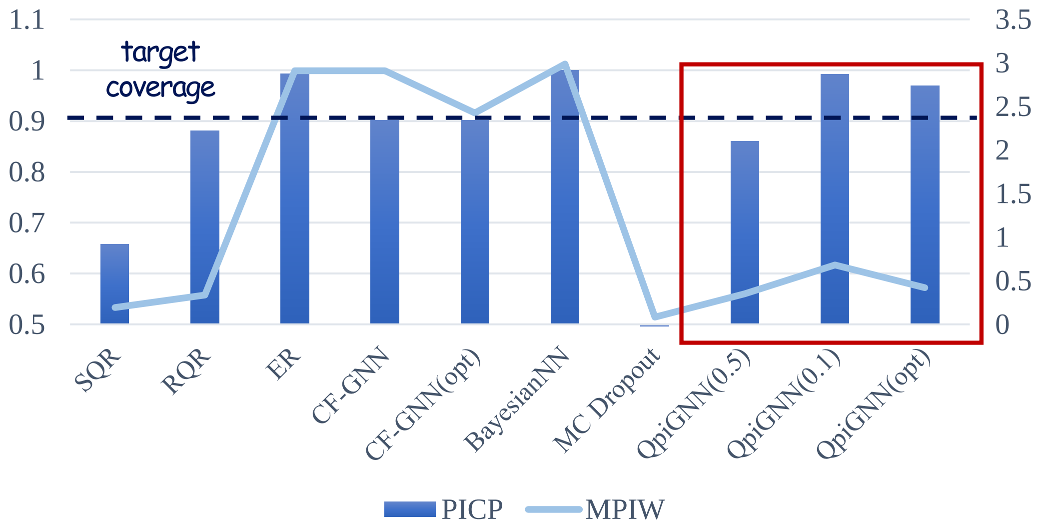


Contents

Quantitative Evaluation

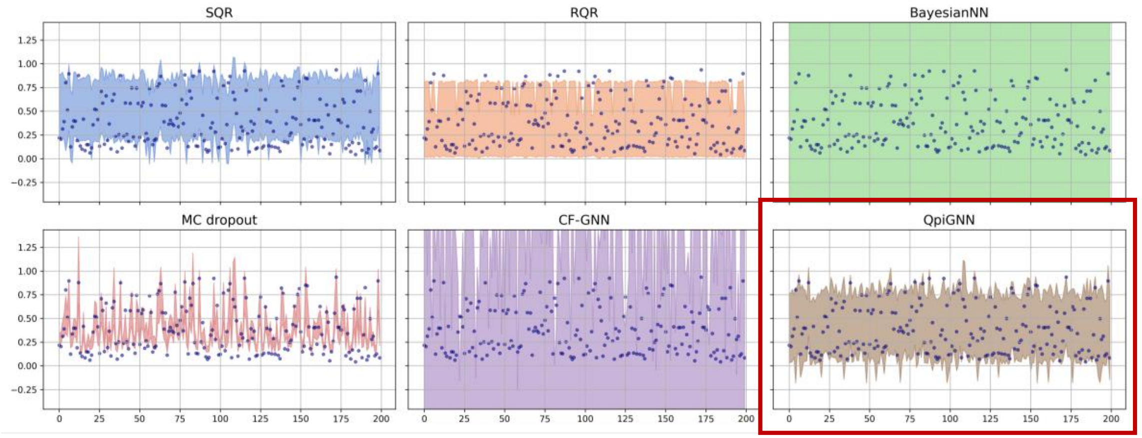
- QpiGNN achieves the best coverage–width trade-off
- +22% coverage, −50% width vs. baselines

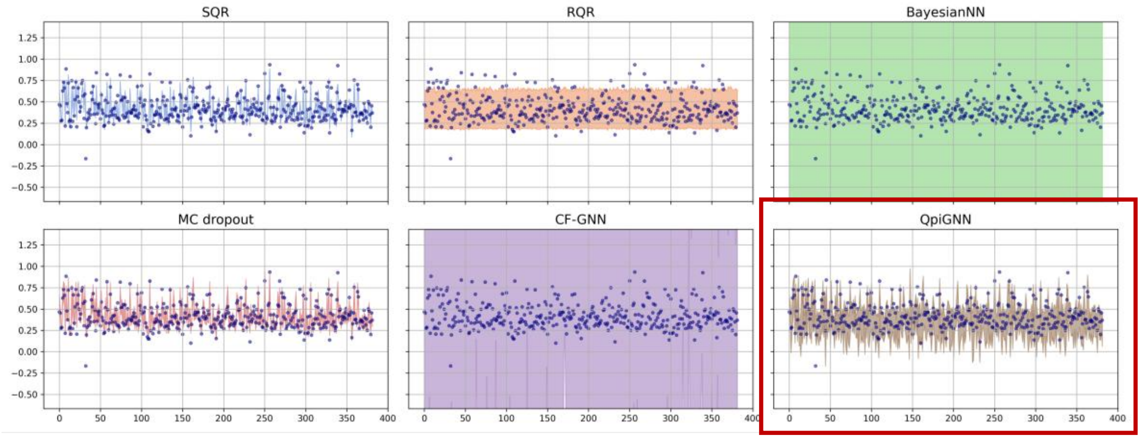




Qualitative Evaluation

- QpiGNN balanced coverage-width and robust across graph settings
- Baselines show a trade-off
  - SQR, MC Dropout: narrow intervals but under-coverage
  - BayesianNN, CF-GNN: adequate coverage but overly wide intervals

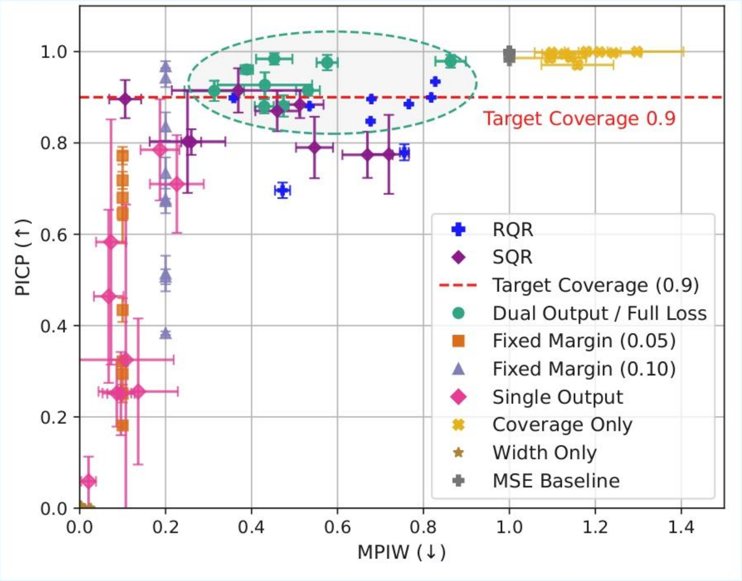




Ablation Study

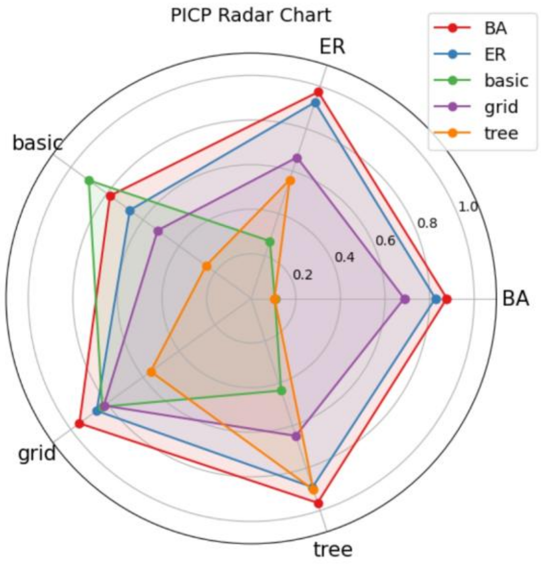
- Dual-head Architecture
  - ➢ Learnable-margin dual head performs best
  - ➢ Outperforms fixed-margin and single-output variants

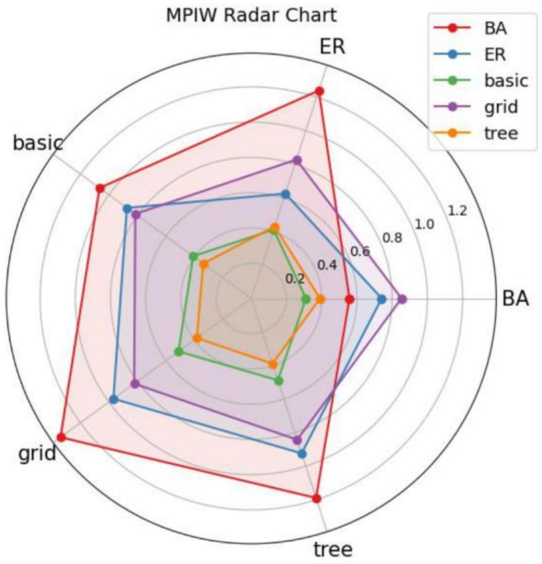
- Quantile-free Joint Loss
  - ➢ Full objective is essential for optimal performance
    - ▪ Coverage-only → overly wide intervals
    - ▪ Width-only → poor calibration
    - ▪ Baseline → trivial/collapsed outputs



Structural Shift Generalization

- Expressive graphs (e.g., BA, ER) → good generalization
- Simple graphs (e.g., Tree, Basic) → poor transfer





Robustness to Perturbations

- Maintains coverage under noise and edge dropout
- Intervals widen with noise; more stable than baselines

Perturbation	Level	$QpiGNN$		SQR		$RQR^{adj.}$	
		PICP≈0.9	MPIW↓	PICP≈0.9	MPIW↓	PICP≈0.9	MPIW↓
Feature Noise	0.1	0.89	0.66	0.76	0.78	0.85	0.78
	0.2	0.90	0.80	0.65	0.69	0.85	0.78
	0.3	0.92	0.83	0.71	0.81	0.84	0.76
Target Noise	0.1	0.96	0.44	0.49	0.52	0.95	0.87
	0.2	0.99	0.71	0.95	0.84	0.97	1.05
	0.3	1.00	1.05	0.98	0.99	0.98	1.32
Edge Dropout	0.2	0.84	0.21	0.53	0.44	0.86	0.80
	0.4	0.88	0.23	0.60	0.44	0.89	0.84
	0.6	0.91	0.21	0.82	0.50	0.89	0.84

Exchangeability under Splits

- Stable under random splits
- Maintains coverage; adapts width under non-exchangeable splits

Split Type	Education		Election		Income		Unemploy.		Twitch	
	PICP≈0.9	MPIW ↓	PICP≈0.9	MPIW ↓	PICP≈0.9	MPIW ↓	PICP≈0.9	MPIW ↓	PICP≈0.9	MPIW ↓
Random	0.99	0.90	1.00	0.97	1.00	0.72	1.00	0.93	0.98	0.54
Degree	0.99	0.75	0.99	0.87	0.99	0.44	0.99	0.93	0.94	0.68
Community	0.99	0.54	0.99	0.86	1.00	0.46	0.97	0.84	0.99	0.49

Split Type	Chameleon		Crocodile		Squirrel		Anaheim		Chicago	
	PICP≈0.9	MPIW ↓	PICP≈0.9	MPIW ↓	PICP≈0.9	MPIW ↓	PICP≈0.9	MPIW ↓	PICP≈0.9	MPIW ↓
Random	0.98	0.40	1.00	0.54	0.99	0.47	0.99	0.74	0.99	0.60
Degree	0.97	0.57	1.00	0.72	0.97	0.45	0.95	0.58	0.96	0.46
Community	0.95	0.58	0.98	0.74	0.97	0.47	0.97	0.60	1.00	0.53

Contents

Summary

- Quantile-Free UQ in GNNs
  - Dual-head architecture + quantile-free joint loss
  - no quantile input, resampling, or post-hoc calibration
- Theoretical Guarantees
  - Asymptotic & finite-sample coverage
  - near-optimal interval width
- Empirical Results
  - +22% coverage, −50% width across 19 datasets
  - robust to noise, structural shift, non-exchangeable splits

Limitations & Future Work

- Limitations
  - Approximate theoretical assumptions
  - Limited analysis on dense/high-degree graphs

- Future Work
  - Theory for graph-dependent settings
  - Aleatoric vs. Epistemic uncertainty
  - Dynamic graphs and broader graph tasks

Q&A