Bayesian Latent Space Models for Graphs Are Misspecified: Toward Robust Inference via Generalized Posteriors

Aldric Labarthe[1, 2]

Advisors: Roland Bouffanais[2], J. Randon-Furling[1]

[1]Ecole Normale Supérieure Paris-Saclay

[2]University of Geneva

May 2026, the 6th

- **The core idea:** Embed network nodes into a continuous metric space where connection probabilities decay with distance (Hoff et al. 2002).
- Highly effective for capturing topological features:
  - Homophily and transitivity.
  - Community structures.
- **The Geometric turn:** Recent shift toward non-Euclidean spaces (Hyperbolic with Krioukov et al. 2010, 2009, Spherical) to explain scale-free degrees and hierarchies.
- **The Bayesian advantage:** Provides principled uncertainty quantification for downstream tasks like link prediction (Newman 2018; Peixoto 2018; Young et al. 2020).

Before discussing how models fail, we must define what they try to achieve.

- **Latent Geometry:** Every node i has an unobserved coordinate z_i in a chosen metric space $(\mathcal{Z}, d_{\mathcal{Z}})$.
- **Link Function:** Dyads connect conditionally independently based on distance.

$$P(A_{ij} = 1 \mid z_i, z_j) = f(d_{\mathcal{Z}}(z_i, z_j))$$

- *Assumption:* We take a f which is strictly monotonically decreasing (closer nodes are more likely to connect).

As we are Bayesian, we do not just want a single best embedding but the **Posterior Distribution** to quantify uncertainty of the positions.

The peril of pure models: Standard Bayesian Random Geometric Graphs (RGGs) assume the true data-generating process aligns perfectly with the chosen geometry and link function.

In practice, real networks violate strict metric axioms:

- **Geometric Mismatch:** Dense hubs exceed the natural packing capacity of Euclidean spaces.
- **Link-Function Mismatch:** Heterophilic connections violate the strict monotonicity of distance-based decay.

The Consequence

Bayesian RGG models are almost universally misspecified.

What happens mathematically when a network contradicts its geometry?

- **Theorem 1 (Packing Capacity):** When volume growth exceeds the model's space, the KL-divergence minimizer is forced *outside* the pure model set $\mathcal{M}$.
- **Theorem 2 (Link Function):** When edges violate strict monotonicity, the KL-minimizer again resides strictly in the convex hull, $\text{conv}(\mathcal{M})$.

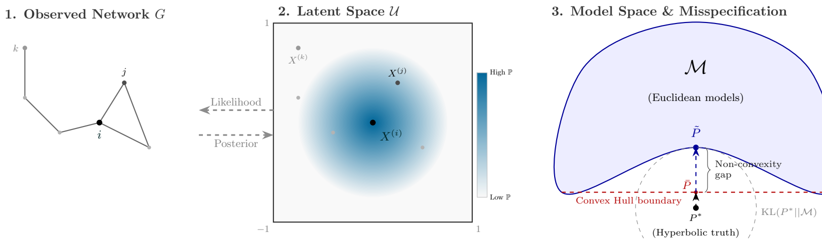


Figure: The true distribution $\tilde{P}$ minimizing KL-divergence resides strictly in $\text{conv}(\mathcal{M})$.

The Staircase Intuition:

- A pure model (e.g., hard-threshold) must commit to a single radius R.
- Real noise (edges outside R, non-edges inside R) incurs infinite log-loss penalties.
- A *mixture* of pure states can simulate a soft decay, avoiding these infinite penalties.

The Overconfidence Trap:

- Because $\tilde{P} \notin \mathcal{M}$, standard Bayesian updating $(\eta = 1)$ fails.
- The posterior asymptotically concentrates on a mixture of pure states.
- This leads to "sure-but-wrong" predictions on out-of-sample dyads.

To prevent the posterior from collapsing onto a single, highly-penalized pure state, we introduce a **generalized η-posterior** (Grünwald et al. 2017):

$$\pi_\eta(Z, r \mid A) \propto \pi(Z)\pi(r) \prod_{i<j} P(A_{ij} \mid z_i, z_j, r)^{\eta}$$

- $\eta \in (0, 1]$ is the learning rate.
- **The effect:** Setting $\eta < 1$ flattens the posterior, mathematically simulating a mixture over pure states in $\mathrm{conv}(\mathcal{M})$.
- **The challenge:** How do we select η when standard marginal likelihoods are inherently biased by the misspecification?

We adapt the SafeBayes framework (Grünwald et al. 2017) for network data by exploiting the *conditional independence* of dyads given the latent positions Z.

Algorithm Summary:

1. Partition the network's dyads into randomized, disjoint blocks.
2. Sequentially assimilate blocks, updating the η-posterior.
3. Predict connections in the *next* unseen block.
4. Measure cumulative prequential log-loss (Dawid 1984) across blocks.

$$\mathcal{L} = \frac{1}{|B_k|} \sum_{(i,j)\in B_k} [A_{ij} \log \hat{p}_{ij} + (1 - A_{ij}) \log(1 - \hat{p}_{ij})]$$

This directly quantifies the predictive penalty of structural mismatches, allowing us to select the optimal η^* without data leakage.

Empirical Result 1: Misspecification is Universal

Across all evaluated real-world networks (KONECT) the optimal learning rate was **always** $\eta^* < 1.0$ with a U-shaped loss curve balancing data and prior influences.

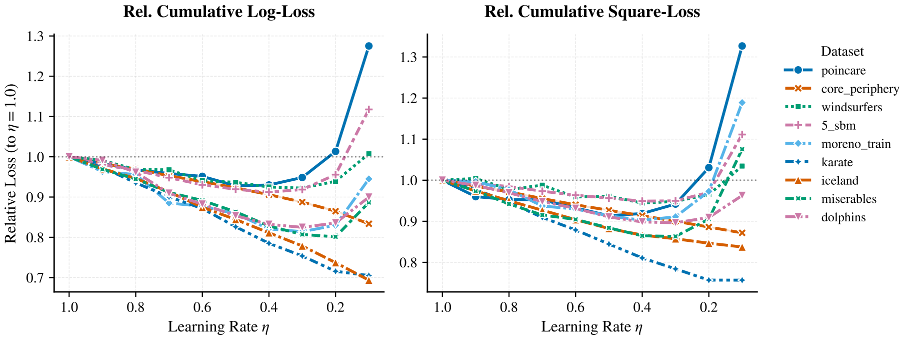
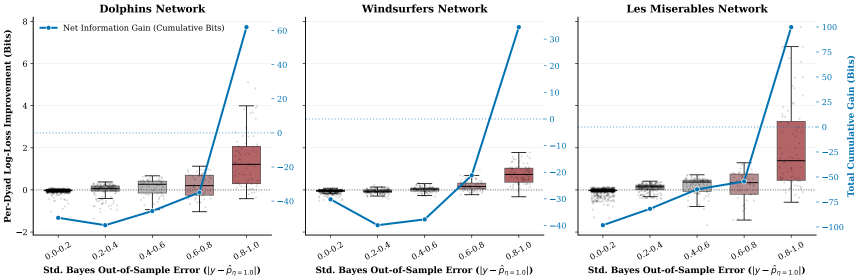


Figure: Relative cumulative losses.

By adopting a flattened posterior, R-SafeBayes prevents the model from making catastrophic, overconfident density estimations on topologically noisy dyads.



- **Outcome:** Average 15.6% reduction in log-loss and 11.6% reduction in squared-loss (Brier score) for link prediction in Euclidean settings.

The prequential risk evaluated by R-SafeBayes provides an unsupervised, highly reliable criterion for identifying a network's true underlying manifold.

- Standard marginal likelihoods are biased under misspecification.
- Prequential log-loss correctly distinguished between Euclidean, Spherical, and Hyperbolic spaces on all synthetic benchmarks.
- **Remarkable Robustness:** It correctly selected the Lorentz Hyperboloid for data generated via the Poincaré disk metric—recognizing the fundamental signature of negative curvature despite distinct isometries.

- **Computational Bottleneck:** $\mathcal{O}(N^2)$ likelihood evaluation combined with sequential blocks and grid search over η makes scaling to massive networks difficult.
- **Sampling Dynamics:** Even with η-tempering, unconstrained MCMC (NUTS) on non-Euclidean manifolds under severe mismatch can still trigger pathological chain behaviors.

- **The Reality:** Pure-state geometric models cannot perfectly assimilate the topological contradictions of real-world networks.
- **The Risk:** Standard Bayesian inference falls into an overconfidence trap, artificially inflating downstream prediction errors.
- **The Remedy:** Generalized η-posteriors, tuned via Link-Sequential R-SafeBayes, act as a vital defense mechanism—restoring calibration, improving link prediction, and revealing the true latent geometry.

Thank You.