Optimality of Message-Passing Architectures for Sparse Graphs

Aseem Baranwal, Kimon Fountoulakis, Aukosh Jagannath

Overview of Contributions

We study the node classification problem on feature-decorated graphs in the sparse setting — expected degree $O(1)$ in the number of nodes. Such graphs are typically known to be locally tree-like.

- We introduce the notion of asymptotic local Bayes optimality for node classification tasks and compute the optimal classifier according to this criterion for a fairly general statistical data model.
- We show that this optimal classifier is implementable using a family of message-passing GNN architectures.
- We then compute the generalization error of this classifier in terms of naturally identifiable SNRs in the data and compare it against existing learning methods and architectures.
- Extensive experiments demonstrate that the classifier is realizable via training using SGD, and is superior to both a simple MLP and a GCN.

Data Model and Definitions

n = # of nodes
d = # of features per node

y_u = class label of node u
C = # of classes

Edges:
$A = (a_{uv})_{u,v\in[n]} \sim \mathrm{SBM}(n, Q)$
$\Pr(a_{uv} = 1 \mid y_u = i, y_v = j) = q_{ij}$

Node features:
$X_u \sim \mathbb{P}_{y_u} \in \mathbb{R}^d$
$\mathbb{P}_c$ = Feature distribution for class c

$G_n \sim \mathrm{CSBM}(n, \mathbb{P}, Q)$ denotes a graph sampled from this model, with adjacency matrix $A \in \{0,1\}^{n\times n}$ and node features $X \in \mathbb{R}^{n\times d}$.

Questions

- What is the optimal classifier when $\mathbb{P}$ and Q are known?
- Can a message-passing architecture realize it by learning $\mathbb{P}$ and Q?

ℓ-local Classifiers ($\mathscr{C}_\ell$)

Input: a subgraph induced by nodes within the ℓ-hop neighbourhood of u, $\eta_\ell(u)$ and the features $\{X_v\}\ \forall v \in \eta_\ell(u)$.
Output: a class label for u.

Local Weak Convergence

For a uniform at random root node u_n, the sequence $(G_n, u_n) \xrightarrow{LWC} (G, u)$, a feature-decorated Poisson Galton-Watson tree.

Optimal Classifier

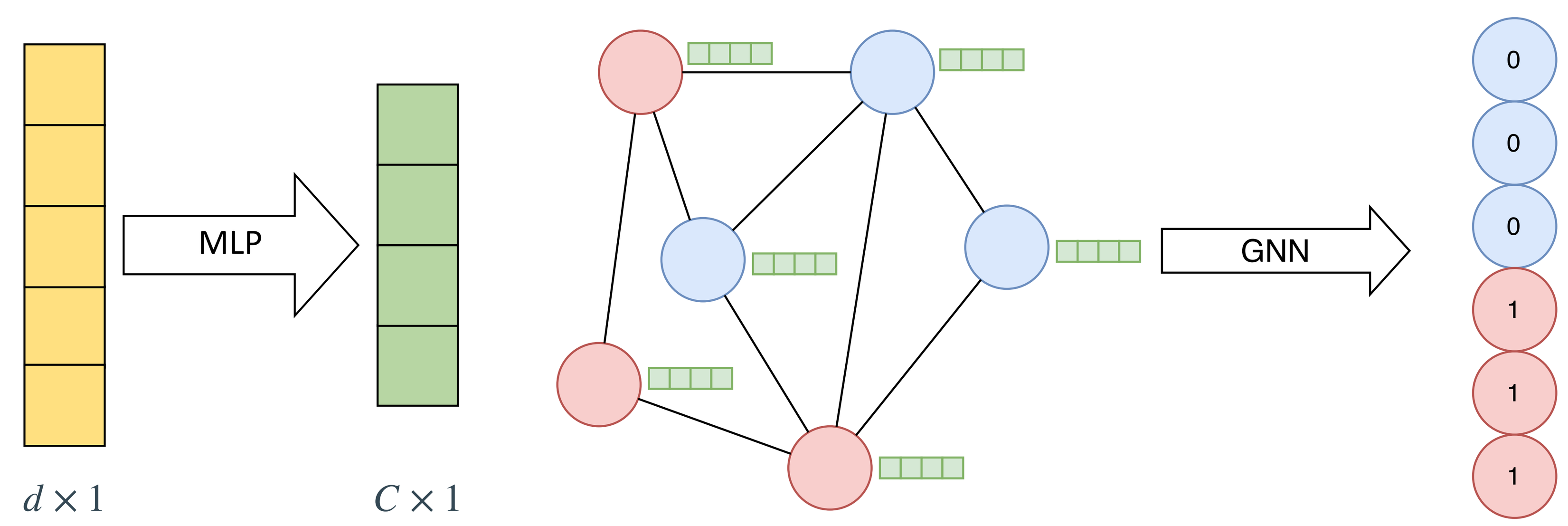
We say h^*_ℓ is the **asymptotically ℓ-locally Bayes optimal** classifier of the root for the sequence $\{(G_n, u_n)\}$ if it minimizes the misclassification probability of the root of the local weak limit (G, u) over $\mathscr{C}_\ell$.

For our data model:

$$h^*_\ell(u, \{X_v\}_{v\in\eta_\ell(u)}) = \operatorname*{argmax}_{i\in[C]}\left\{\log\rho_i(X_u) + \sum_{v\in\eta_\ell(u)\setminus\{u\}} M_{i\,d(u,v)}(X_v)\right\}$$

$$M_{ik}(x) = \max_{j\in[C]}\left\{\log\rho_j(x) + \log Q^k_{ij}\right\}$$

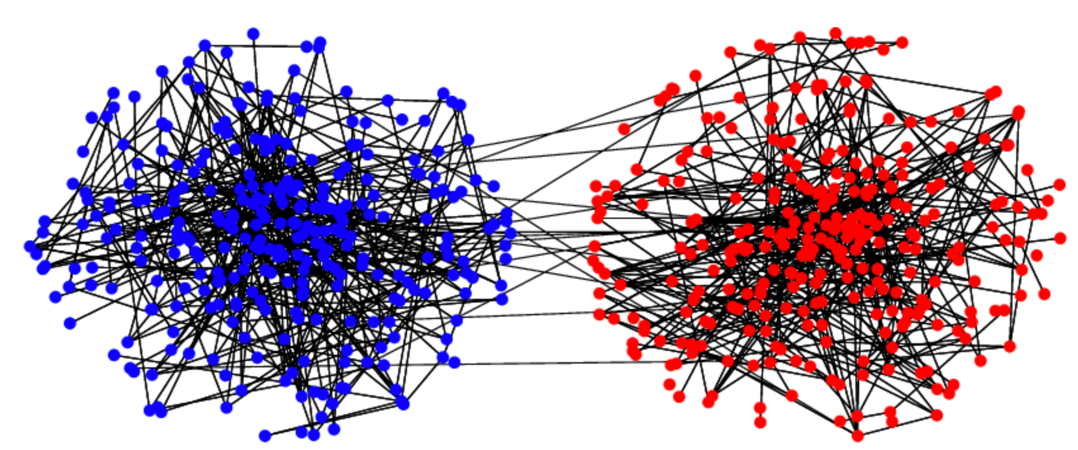
Implementation: Message-Passing GNN Architecture



Example: Binary Gaussian Mixture

$$\mathbb{P} = \{\mathcal{N}(\pm\mu, \sigma^2 I)\}$$

$$Q = \frac{1}{n}\begin{pmatrix} a & b \\ b & a \end{pmatrix}$$



Graph signal $\Gamma = \dfrac{a-b}{a+b}$ Feature signal $\gamma = \dfrac{2\|\mu\|}{\sigma}$

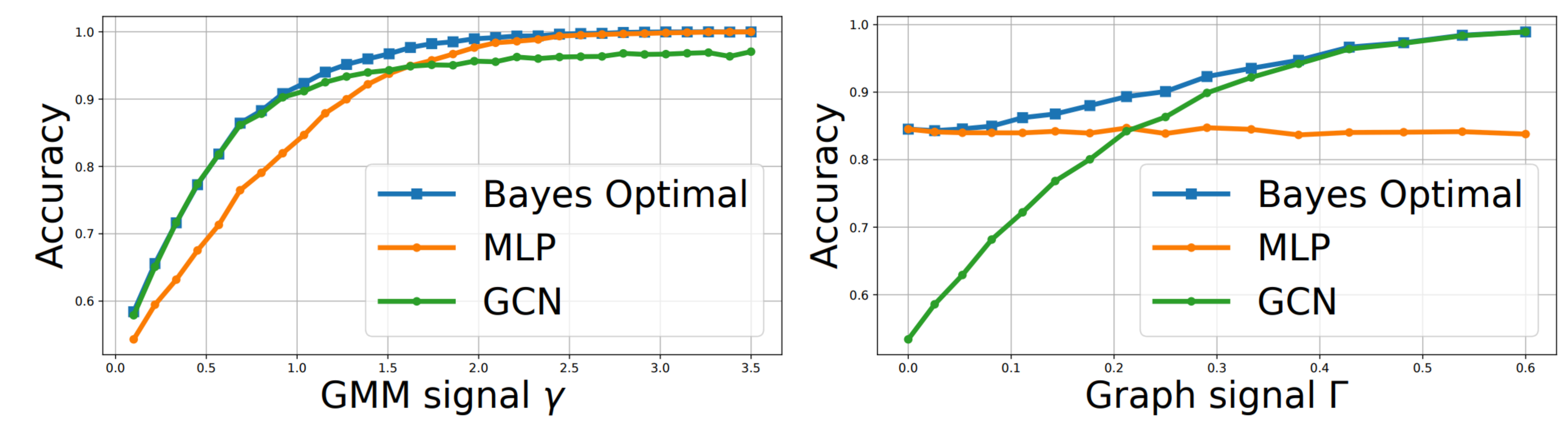
$$h^*_\ell(u, \{X_v\}_{v\in\eta_\ell(u)}) = \operatorname{sgn}\left(\langle X_u, \mu\rangle + \sum_{v\in\eta_\ell(u)\setminus\{u\}} M_{d(u,v)}(X_v)\right)$$

$$M_k(x) = \operatorname{sgn}(a-b)\cdot\mathrm{CLIP}(\langle x,\mu\rangle, \pm c_k), \qquad c_k = \log\left(\frac{1+\Gamma^k}{1-\Gamma^k}\right)$$

Characteristics

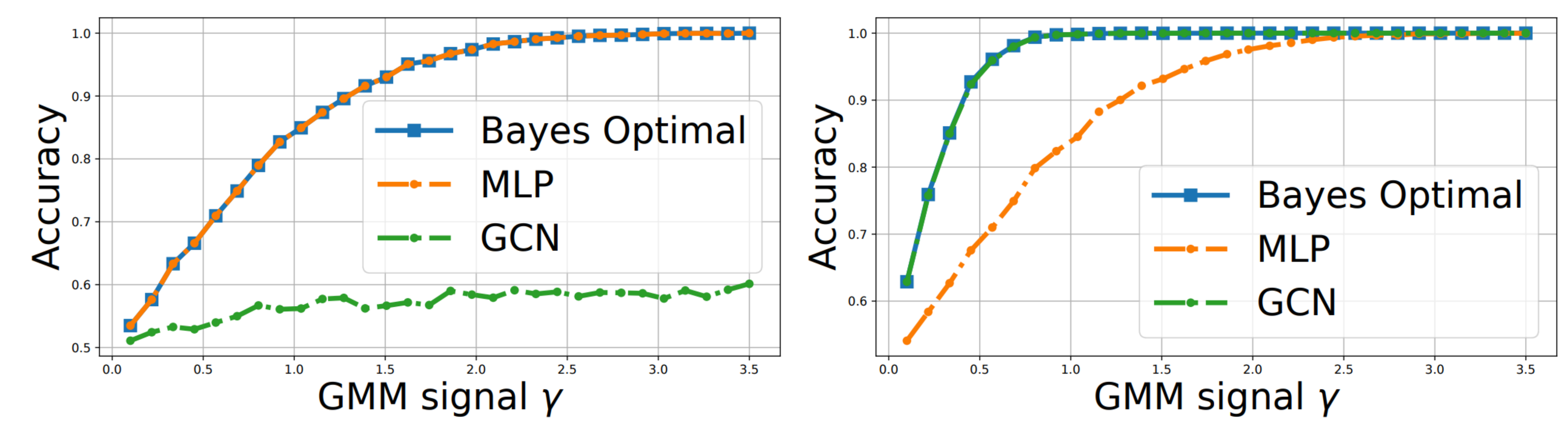
What happens in the limits of Graph SNR?

- When $\Gamma \to 0$, h^*_ℓ ignores all messages and collapses to a simple MLP.
- When $\Gamma \to 1$, h^*_ℓ collapses to a typical GCN.
- When $\Gamma \in (0,1)$, h^*_ℓ interpolates and is superior to MLP and GCN.



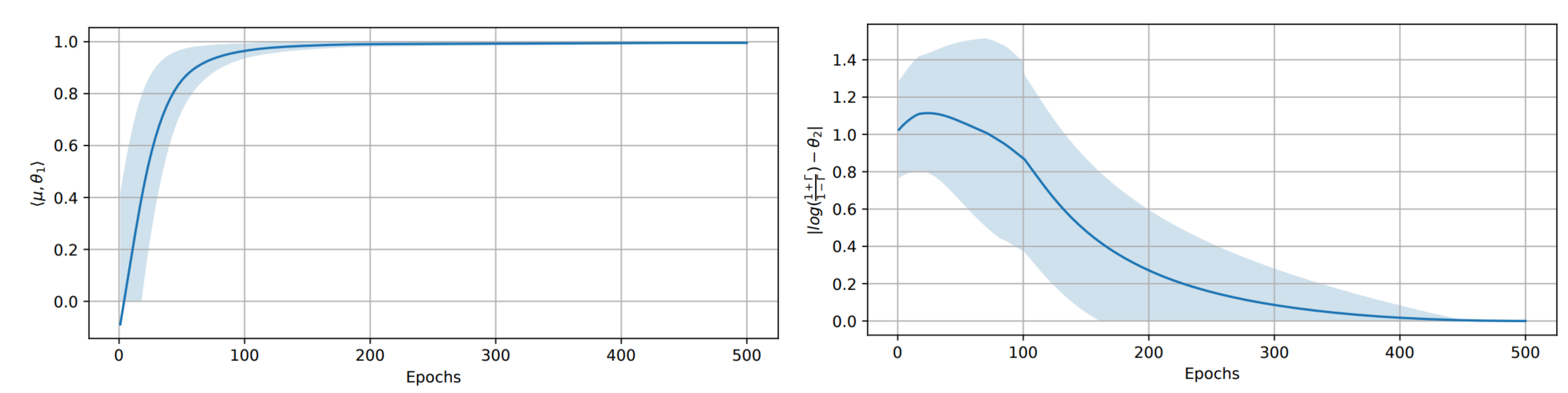
(a) Varying γ with fixed $\Gamma = 0.42$. (b) Varying Γ with fixed $\gamma = 1$.

Comparison with MLP and GCN (Kipf & Welling 2017).



(a) Fixed graph signal $\Gamma = 0$. (b) Fixed graph signal $\Gamma = 1$.

Demonstration of collapsing to MLP and GCN in the limits of graph SNR.



Convergence of model parameters to the ansatz during training.

Non-asymptotic setting

For fixed number of nodes n and $4\ell \le \log_{\mathbb{E}\deg}(n)$, the classifier h^*_ℓ is $o_n(1)$ away from the true optimal in terms misclassification probability.

Code: github.com/opallab/optimality-mp-archs-sparse-graphs.

Email aseemrb@gmail.com
Website aseemrb.me