Graph Neural Network Sensitivity Under Probabilistic Error Model

Xinjue Wang, Esa Ollila and Sergiy A. Vorobyov
Department of Signal Processing and Acoustics
Aalto University
Finland
{xinjue.wang,esa.ollila,sergiy.vorobyov}@aalto.fi

Abstract—Graph convolutional networks (GCNs) can successfully learn the graph signal representation by graph convolution. The graph convolution depends on the graph filter, which contains the topological dependency of data and propagates data features. However, the estimation errors in the propagation matrix (e.g., the adjacency matrix) can have a significant impact on graph filters and GCNs. In this paper, we study the effect of a probabilistic graph error model on the performance of the GCNs. We prove that the adjacency matrix under the error model is bounded by a function of graph size and error probability. We further analytically specify the upper bound of a normalized adjacency matrix with self-loop added. Finally, we illustrate the error bounds by running experiments on a synthetic dataset and study the sensitivity of a simple GCN under this probabilistic error model on accuracy.

Index Terms—Graph signal processing, graph neural network, probabilistic error model, stability

I. INTRODUCTION

Graph convolutional networks (GCNs) are recently developed machine learning tools for processing data with graph structures [1]. GCN consists of cascaded layers of graph convolutional filters and nonlinear activation functions. Graph convolutional filters couple the data and graph with the underlying graph matrix, named graph shift operator (GSO), such as the adjacency matrix and Laplacian, which encodes the arbitrary interactions between data samples [2]. Based on the GSO, the graph filter captures the structural information by aggregating the data propagated within its $k$-hop neighbourhoods, and feeds it to the next layer after processing techniques like nonlinearities, optional graph coarsening and pooling [3], [4]. As the key component of GCNs, GSO presents the graph structure, and is typically assumed to be perfectly estimated. The precise estimation of the hidden graph structure is essential for successfully performing feature propagation on a convolution layer [5].

In many real-world problems, the underlying graph is unknown and is usually estimated with errors [6]. Due to the close relation between GSO and graph filter, the stability of graph filter under perturbation of GSO is a critical question. It is also necessary to study the performance of GNN with a cascade of perturbed graph filters. Furthermore, it is essential to study the effect of different types of graph perturbation on outputs of a graph neural network (GNN).

The effect of general graph perturbation on the performance of GNN is studied in [7], where the authors prove the stability of GNN with an integral Lipschitz filter under relative graph perturbations, and derive an analytical upper bound controlled by the operator norm of error matrix. The bound in [7] is loose because of the general graph error model used, thus, leaving space for improvements by specifying the error model and network. In [8], a similar generalization bound is proposed, which is determined by the largest absolute eigenvalue of the graph convolution filter. The authors also consider the randomness of stochastic gradient descent (SGD) in the learning procedure.

Modelling and processing with the determined and stochastic topological perturbations on graphs have been shown in the existing literature on graph signal processing (GSP) and GNN. A probabilistic graph error model for partially correct estimation of the adjacency matrix is proposed in [6], where a perturbed graph is modelled as a combination of the true adjacency matrix and a perturbation term specified by Erdös-Rényi (ER) graph. In [9], a graph error model is constructed by successively connecting two nodes that are not connected for studying the relation between algebraic connectivity with an increasing number of edges. In contrast, [10] discusses the strategy of rewiring edge selection for the most significant increase in algebraic connectivity. With increased connectivity, the network is more robust against attacks and failures. In [11], closed-form expressions of perturbed eigenvector pairs under edge perturbation are derived by small perturbation analysis, formulating signal processing algorithms that are robust to graph errors. The work in [12] studies errors in graphs using random edge sampling, which is a scheme where existing edges are deleted randomly.

In this paper, we discuss the sensitivity of graph filters by deriving an upper bound for the difference between the original filter and a filter under perturbation (cf. Theorem 1). We further specify the type of graph perturbation with the probabilistic error model [6] for GSO, which is an adjacency matrix, and derive an upper bound for GSO difference in Theorem 2. Our bound is tighter than the general counterpart with a trivial Euclidean norm bound by exploiting a particular graph error model. We then extend the bound of the adjacency matrix to its augmented normalized counterpart in Theorem 3. After the key component of the GNN is specified, we discuss the sensitivity
of the simple GNN [13] both analytically and numerically. In general, we develop a methodology for GCN sensitivity analysis.

**Notation:** We use boldface lower case letters \(x\) for column vectors and boldface capital letters \(X\) for matrices. Moreover, \(1_N\) stands for a vector full of ones, the \(N \times N\) matrix full of ones is \(1_{N \times N} = 1_N1_N^\top\), while \(I_N\) denotes an identity matrix of size \(N \times N\). The notation \(\| \cdot \|\) stands for the Euclidean norm for vectors and spectral norm for matrices, i.e., \(\|A\| = \sigma_1(A) = \sqrt{\lambda_1(A^\top A)}\), where \(\sigma_1(A)\) and \(\lambda_1(A)\) denote the maximum singular value and maximum eigenvalue of matrix \(A\), respectively. Furthermore, \(c_1(A)\) represents the largest Euclidean column length of \(A\), \(\| \cdot \|_F\) denotes the Frobenius matrix norm \((\|A\|_F = \sqrt{\text{tr}(A^\top A)})\), \(\circ\) denotes the Hadamard product, and \(\mathbb{P}(\cdot)\) denotes probability.

**II. Preliminaries**

Consider a graph \(\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{W})\) with vertex set \(\mathcal{V} = \{v_1, \ldots, v_N\}\) of \(N\) nodes, edge set \(\mathcal{E} \subset \mathcal{V} \times \mathcal{V}\), and edge weighting function \(\mathcal{W} : \mathcal{E} \to \{0, 1\}\). We only discuss the undirected and unweighted graph, so \(\mathcal{W}(v_i, v_j) = 1\) for \((v_i, v_j) \in \mathcal{E}\). Signal on the graph is denoted by \(x \in \mathbb{R}^N\). Its \(i\)-th entry \([x]_i = x_i\) thus specifies the data value at vertex \(v_i\). We assume that the perturbation has no effects on the nodes of the true graph, thus, the perturbed graph \(\hat{\mathcal{G}} = (\mathcal{V}, \hat{\mathcal{E}}, \hat{\mathcal{W}})\).

**A. Graph Neural Network**

Graph shift operator (GSO) \(S \in \mathbb{R}^{N \times N}\), such as adjacency matrix \(A\), the Laplacian \(L\), or their normalized versions, represents the sparsity and connectivity of a graph \(\mathcal{G}\). As we constrain the graph to be undirected and unweighted, the GSOs are symmetric. Using GSO, signals can be shifted and averaged across their neighbourhoods. The one time shift of graph signal is simply \(Sx\), whose value at node \(v_i\) is \([Sx]_i = \sum_{j \in \mathcal{N}(v_i)} s_{ij}x_j\). A graph convolutional filter \(h(S)\) with \(K + 1\) taps is defined via polynomials of GSO and the filter weights \(h = \{h_k\}_{k=0}^K\) in the graph convolution

\[
y = h_0S^0x + h_1S^1x + \cdots + h_KS^Kx = h(S)x, \tag{1}
\]

where \(h(S) = [h_0S^0, h_1S^1, \ldots, h_KS^K]\) is a shift-invariant graph filter with \(K + 1\) taps, and denotes the weight of local information after \(K\)-hop data exchanges.

A general GCN consists of layers of graph convolutional filters followed by nonlinearities. We use matrix \(X\) to denote the input graph signal with \(d\) features \(X = [x_1, \cdots, x_d]\), and the feature propagation in a single GCNN layer is described as

\[
Y = \sigma\left(\sum_{k=0}^K S^kXH_k\right), \tag{2}
\]

where \(H_k \in \mathbb{R}^{d \times C}\) is the collection of \(k\)th entries of \(d \times C\) learnable graph filters, \(C\) represents the number of output features, and \(\sigma\) denotes the pointwise nonlinear activation function.

**B. Probabilistic Graph Error Model**

In [6], the Erdős-Rényi graph is used to model the perturbations on a graph adjacency matrix. The adjacency matrix of ER graph is a random \(N \times N\) matrix \(\Delta\), where \(\mathbb{P}(|\Delta|_{i,j} = 1) = \epsilon\) and \(\mathbb{P}(|\Delta|_{i,j} = 0) = 1 - \epsilon\) for all \(i \neq j\), and \(|\Delta|_{i,i} = 0\) for \(i = 1, \ldots, N\), where each element of the matrix is generated independently of the other elements. We revise the random matrix for the undirected case by using the lower triangular matrices \(\Delta'_l\), and then defining \(\Delta_e = \Delta'_l + (\Delta'_l)^\top\). The perturbed adjacency matrix of a graph signal is modelled as

\[
\hat{A} = A - \Delta_{e_1} \circ A + \Delta_{e_2} \circ (1_{N \times N} - A), \tag{3}
\]

where the 1st term is responsible for edge removal with probability \(\epsilon_1\), while the 2nd term is responsible for edge addition with probability \(\epsilon_2\). The error model can be seen as adding two Erdős-Rényi graphs on the top of the true graph.

**III. Main Results**

Let \(\hat{S}\) denote the perturbed GSO and the filter weights due to perturbation by \(h = \{\hat{h}_0, \hat{h}_1, \ldots, \hat{h}_K\}\). Denoting \(\hat{S} = S + E\), where \(E\) is the perturbation term, whose spectral norm

\[
d(\hat{S}, S) = \|\hat{S} - S\| = \|E\| \tag{4}
\]

defines a distance between the perturbed and the true GSOs. The following theorem investigates the error on the graph convolution filter \(h(S)\) due to the perturbed GSO and filter weights.

**Theorem 1.** Let \(S\) and \(\hat{S}\) be the GSOs for the true graph \(\mathcal{G}\) and the perturbed graph \(\hat{\mathcal{G}}\), respectively. The filter distance between \(h(S) = \sum_{k=0}^K h_kS^k\) and \(\hat{h}(\hat{S}) = \sum_{k=0}^K \hat{h}_k\hat{S}^k\) satisfies

\[
\|\hat{h}(\hat{S}) - h(S)\| \leq \sum_{k=0}^K (k|h_k||E|)^{k-1} + \epsilon_k|\lambda|^k, \tag{5}
\]

where \(\hat{h}_k = h_k + \epsilon_k, k = 0, \ldots, K\) denote the perturbed filter coefficients and \(\lambda\) denotes the largest of the maximum singular values of two GSOs, so \(\lambda = \max\{\|S\|, \|S\|\}\).

**Proof.** See Appendix A in [14].

Theorem 1 shows that the error in graph filter distance is controlled by the filter degree \(K\), the spectral distance between GSOs, \(\|E\|\), the maximum singular value of GSOs, \(\lambda\), and the filter coefficients \(h_k\). A graph filter with a higher degree thus suffers more from instability than a low-degree filter. Operator distance is our primary interest because it is related to structural perturbation analytically. Particularly, we further specify the bound in (5) when the underlying graph follows a stochastic block model (SBM) by bounding \(\|E\|\) under the probabilistic error model.

If we set the GSO as a normalized adjacency matrix, implying that \(\lambda \leq 1\), then the upper bound in (5) can be simplified as \(\sum_{k=0}^K (k|h_k||E| + |\epsilon_k|)\). The perturbation of the filter coefficients \(|\epsilon_k|\) is related to the difference of input GSOs, the activation functions between two network layers,
and the intrinsic randomness in stochastic gradient descent (SGD) of a GCN. For bounding the error $|e_k|$ in learnable filter weights, one needs to specify the GCN. We assume uniform stability for GCN weights [8], and in practice, we normalize the weight matrix on each layer by the standard deviation.

**A. Case 1: Adjacency Matrix as GSO**

Consider the case when the adjacency matrix is used as the GSO, so $\hat{S} = A$ and $S = A$. Based on the probabilistic graph error model (3), the perturbation term $E = S - \hat{S}$ is

$$E_A = \hat{A} - A = -\Delta_{e_1} \circ A + \Delta_{e_2} \circ (1_{N \times N} - A).$$

(6)

Next, we derive a bound on spectral norm of the error $\|E_A\|$.

**Theorem 2.** The distance $d(\hat{A}, A) = \|E_A\|$ is bounded by a constant $\varepsilon$ such that almost surely,

$$d(\hat{A}, A) \leq c_1(A)\rho(N, \varepsilon_1) + (1 + c_1(A))\rho(N, \varepsilon_2) = \varepsilon,$$

with function $\rho: \mathbb{Z}^+ \times [0, 1] \rightarrow \mathbb{R}^+_0$ defined by

$$\rho(n, p) = \frac{p}{n^2} + 2\sqrt{p(1-p)} + \frac{p}{n^2} \log n,$$

(7)

where $\varepsilon$ is a positive constant.

**Proof.** See Appendix B in [14].

(8)

Theorem 2 shows that the operator distance $d(\hat{A}, A)$ is controlled by a function of graph size $N$ and edge addition/removal probabilities ($\varepsilon_1, \varepsilon_2$) of the ER graphs, as well as the Euclidean column length of GSO $A$. Note that the function $\rho(n, p)$ is monotonically increasing if we set the constant $\xi = 1$, the number of nodes $n \geq 50$, and the probability $p \in [0, 1]$. The bound in (7) worsens when more randomness is added to the original graph. Furthermore, the maximal Euclidean length of the original adjacency matrix represents the sparsity of the original graph in a sense. The sparsity of the graph guarantees the tightness of the upper bound. We note that Theorem 2 provides a general bound (regardless of graph structure), and hence for a particular graph, it will be loose.

To show the improvement of our results, we make a comparison with the trivial bound only using the operator norm

$$d(\hat{A}, A) \leq \|\Delta_{e_2}\| + \|\Delta_{e_1} + \Delta_{e_2}\|\|A\|,$$

(9)

and a partially improved bound which uses the sparsity of two random graphs $\Delta_{e_1}, \Delta_{e_2}$ in (7)

$$d(\hat{A}, A) \leq \rho(N, \varepsilon_2) + c_1(\Delta_{e_1} + \Delta_{e_2})\|A\|.$$

(10)

The numerical results are shown in Section IV-A and the detailed settings of the underlying graph are discussed.

**B. Case 2: Augmented Normalized Adjacency Matrix as GSO**

Consider the augmented normalized adjacency matrix $\hat{A} = \hat{D}^{-\frac{1}{2}}\hat{A}\hat{D}^{-\frac{1}{2}}$, where $A = A + I$ is the adjacency matrix of the graph added with self-loops, and $\hat{D} = \text{diag}(d_1 + 1, \ldots, d_N + 1)$, with $d_i = \text{deg}(v_i)$, is the augmented degree matrix. The spectrum range of $A$ is $[-1, 1]$, thus the repeated application of $A$ in a GCN will not suffer from numerical instabilities.

We define the error matrix between the true and augmented perturbed degree matrices as

$$E_D = \hat{D} - D = \text{diag}(d_{e_1}, \ldots, d_{e_N}),$$

(11)

where $\{d_{e_i}\}_{i=1}^N$ denotes the degree difference at each vertex. For this diagonal matrix, its norm is the maximum element in the diagonal,

$$\|E_D\| = \max_i |d_{e_i}|, i = 1, \ldots, N.$$

Next let GSO be the augmented normalized adjacency matrix $S = A$, and let $\hat{S} = \hat{A}$ denote the perturbed GSO. The difference matrix $E_A$ is then

$$E_A = \hat{A} - A = \hat{D}^{-\frac{1}{2}}\hat{A}\hat{D}^{-\frac{1}{2}} - \hat{D}^{-\frac{1}{2}}A\hat{D}^{-\frac{1}{2}} = (\hat{D} + E_D)^{-\frac{1}{2}}(\hat{A} + E_A)(\hat{D} + E_D)^{-\frac{1}{2}} - (\hat{D} + E_D)^{-\frac{1}{2}}A(\hat{D} + E_D)^{-\frac{1}{2}}.$$

(13)

Next theorem provides a bound on $\|E_A\|$.

**Theorem 3.** The distance $d(\hat{A}, A) = \|E_A\|$ is bounded by

$$\|E_A\| \leq \alpha\|A\| + \beta\varepsilon$$

(14)

where $\varepsilon$ is the upper bound of $\|E_A\|$ in (7), $\delta = \max_i (d_i + d_{e_i} + 1)^{-1/2} - (d_i + 1)^{-1/2}$,

$$\alpha = \delta^2 + \frac{1}{d_{\text{min}} + 1} + \frac{\delta}{\sqrt{d_{\text{min}} + 1}}$$

and $\beta = \alpha + \frac{\delta}{\sqrt{d_{\text{min}} + 1}}$

with $d_{\text{min}} = \min_i d_i$ being the minimum degree of the original graph.

**Proof.** See Appendix C in [14].

(15)

Theorem 3 establishes that the perturbation on the augmented normalized adjacency matrix is bounded by the largest singular values of the augmented adjacency matrix $A$ and the perturbed adjacency matrix $E_A$. The derived bound is not necessarily tight as it does not exploit any a priori information about the graph (such as sparsity or other possible structures).

**C. The stability of SGCN**

The simple GCN (SGCN) [15] is motivated by considering a multilayered GCN with an affine approximation of graph convolution filter and the removal of activation functions between layers. By simplifying a GCN in [16] with $K$ layers, the graph filter in SGCN can be represented by keeping only the $K$th order GSO in (1), so the output of the filter is $y = h_K S^K x = h(S) x$ and the output of a SGCN with a single linear logistic regression layer is

$$Y = \sigma(S^K X H_K),$$

(15)

where $[\sigma(x)]_i = e^{x_i}/\sum_{j=1}^C e^{x_j}$ is a softmax function applied to row vectors of the input matrix, $X \in \mathbb{R}^{N \times d}$ is the graph data, and $H_K \in \mathbb{R}^{d \times C}$ is a learnable weight matrix with $C$ classes for prediction.

Let the output of SGCN with perturbed GSO be $Y = \sigma(\hat{S}^K \hat{X} H)$, where $\hat{S}$ is the distorted GSO, and $\hat{H}$ is the weight matrix learned under the perturbed GSO. We can derive the
following upper bound on the difference between the outputs of original and perturbed GSOs.

**Corollary 1.** The distance between the original and perturbed SGCN outputs is

\[
d(\hat{Y}, Y) = ||\sigma(S^KXH) - \sigma(S^KXH)|| \\
\leq (||S||^K\delta_H + K\lambda^{K-1}||E_S||\|H\|)\sqrt{d},
\]

where \(E_S = \hat{S} - S\), and \(\delta_H \geq \|\hat{H} - H\|\).

**Proof.** See Appendix D in [14].

### IV. Numerical Simulations

#### A. Theoretical Bound Corroboration

We execute this experiment to corroborate the upper bound in Theorem 2. We consider a two-group planted partition model (PPM), which is a special case of the SBM by setting all the in-group probabilities to \(p_{in} = 0.8\), and the between-group possibility to \(p_{net} = 0.5\). The PPM graph is perturbed by the error model in (3) with \(\epsilon_1 = 0.1\), and \(\epsilon_2 = 0.01\), where the small edge removal/addition probabilities are designed to perturb the graph without destroying the original structure. We carry out 50 Monte Carlo simulations for graph size \(N \in [50, 1000]\) with step size of 50. We study how our theoretical bounds in (9), (10) and Theorem 2 change with the increasing graph size, \(N\), and compare the bounds with the empirical difference between the true graph and its perturbed counterpart.

Fig. 1 shows the results. First, the results corroborate the impact of the probabilistic graph error model on the original graph. The graph difference increases with graph size, and so does the analytical bound. Second, we note that the theoretical analysis in Theorem 2 offers a better upper bound than the other two as it exploits the sparsity of the random matrices in the error model. One can also note that the upper bound is not tight, intrinsically because of its generalization ability for all graphs. The main goal of the bound analysis is to show that the probabilistic error model does not ruin the structure of the original graph and indicates the role played by the graph size on the upper bound and the graph difference.

#### B. Error Effect on SGCN

We evaluate the effect of the probabilistic error on the node classification performance on the Cora citation graph [17], which has \(C = 7\) classes, \(N = 2708\) vertices. Assuming the indirect nature of the underlying graph, we modify the original Cora graph from a directed to an undirected one. The undirected Cora graph has \(|E| = 5278\) edges. We follow the data splitting scheme in [13]. For the probabilistic error model, we distort the original graph with all combinations of edge deleting probabilities \(\epsilon_1 = [0, 3, 6, 9] \times 10^{-2}\), and edge adding probabilities \(\epsilon_2 = [0, 3, 6, 9] \times 10^{-4}\). Each error combination is run for 100 times to capture the randomness introduced by the error model. The GSO \(S\) is the augmented normalized adjacency matrix, and the orders of graph filters are \(K = 2\) and \(K = 3\).

Effects of probabilities of deleting edges on the probabilities of adding edges are shown in Fig. 2(a), and vice versa in Fig. 2(b). In Fig. 2(a), when more edges are deleted, the test accuracy of SGCN keeps relatively stable, and the largest drop is approximately 5% with order \(K = 2\) and \(\epsilon_2 = 0.0009\). The
overall performance of SGCN is better when the order of the graph filter increases from 2 to 3. The error bar indicates the sensitivity of SGCN under the perturbation. The increasing probability of adding edges breaks the network, and the results are lower accuracies along with larger variations. Fig. 2(b) shows that the effect of $\epsilon_2$ on accuracy is greater than that of $\epsilon_1$. For further explanation, the original graph of Cora $G_{ori}$, its edge deleted counterpart $G_{del}$ with $\epsilon_1 = 9 \times 10^{-2}$, and its edge added version $G_{add}$ with $\epsilon_2 = 9 \times 10^{-4}$ are shown in Fig. 3(a), (b) and (c), respectively. The sparsity of the original graph could partly explain the effect difference between $\epsilon_1$ and $\epsilon_2$, that is, the number of added edges is larger than the number of deleted edges, and therefore $G_{add}$ gives more errors than those of $G_{del}$ in a GNN task.

V. CONCLUSION AND DISCUSSION

In this paper, we derived a theoretical upper bound for the graph filter distance under graph topology perturbation. By specifying the perturbation with the probabilistic error model, we gave upper bounds for GSOs as an adjacency matrix and its augmented normalized adjacency counterpart. Numerical results validated our results.

It would be promising to extend the analytical bound of SGCN to a general case of GNN once the key component, a tight bound of GSO distance under a particular graph topology, is ready. For simplicity, we considered in this paper the unweighted and undirected graph. A future research topic is to extend the analysis to weighted and directed graphs, whose changes can be modelled by a generalized form of probabilistic error model in [6].

REFERENCES