Northeastern University

REGE: A Method for Incorporating Uncertainty in Graph Embeddings

Zohair Shafi, Germans Savcisens and Tina Eliassi-Rad

Germans Savcisens
Northeastern University

Tina Eliassi-Rad
Northeastern University

2025 SIAM International Conference on Data Mining

Preliminaries

- Graph embedding
  - Let $G = (V, E, X)$ be a graph with $|V| = n$ nodes and $|E| = m$ edges.
  - $X \in \mathbb{R}^{n \times r}$ is a node feature matrix with r features per node.
  - $f(G(V, E, X)) = Z \in \mathbb{R}^{n \times d}$ is a graph embedding function. It maps each node onto a d-dimensional space.
- Uncertainty
  - Data: Uncertainty due to noisy or incomplete data
  - Model: Uncertainty due to parameters, optimization strategy, lack of training knowledge, *etc*.

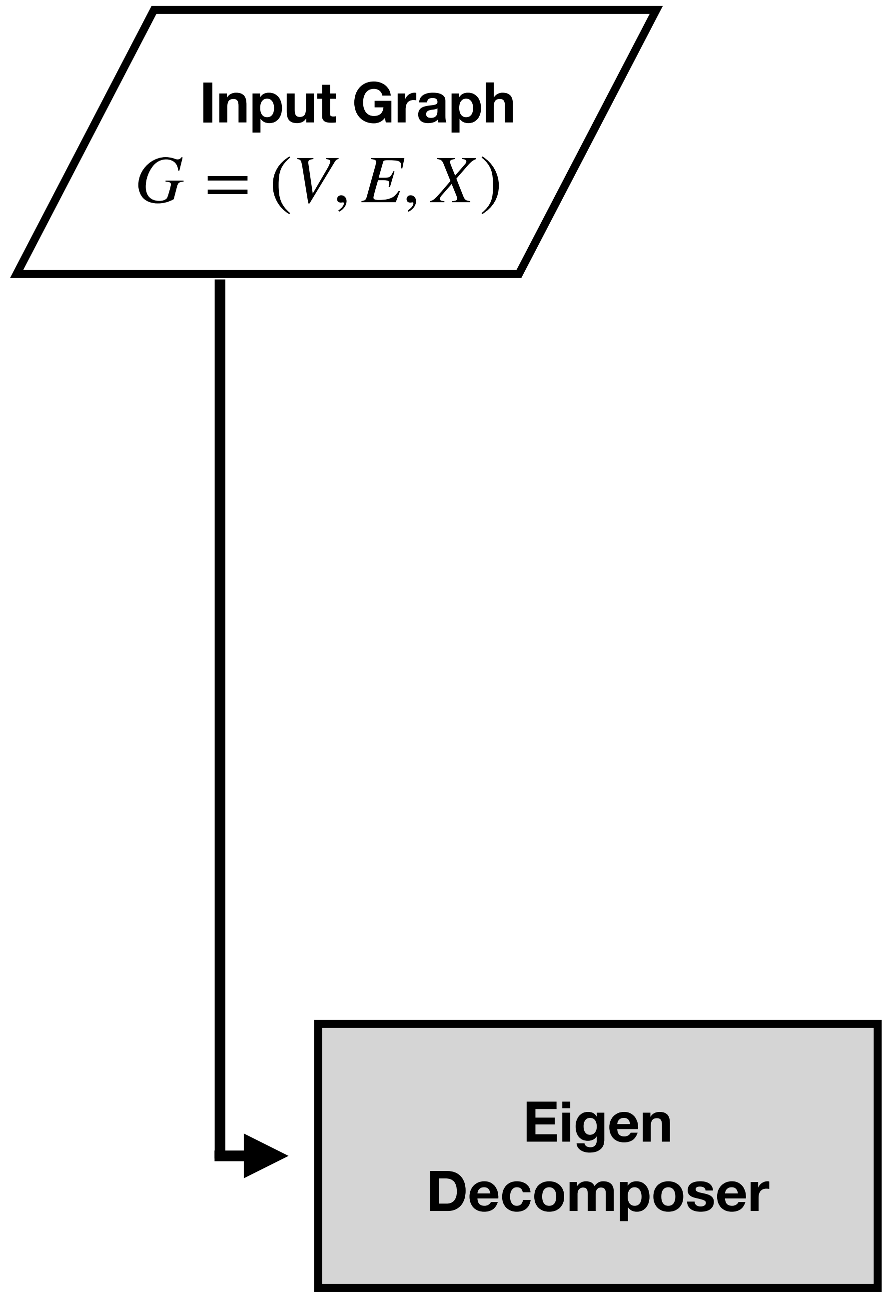
Motivation

- Why should we expect a node to embed in an exact spot in a *d*-dimensional space?
- Can we create a notion of a “radius” around each node where the node may embed?
- Could this "radius" help make node embeddings more robust to adversarial attacks?

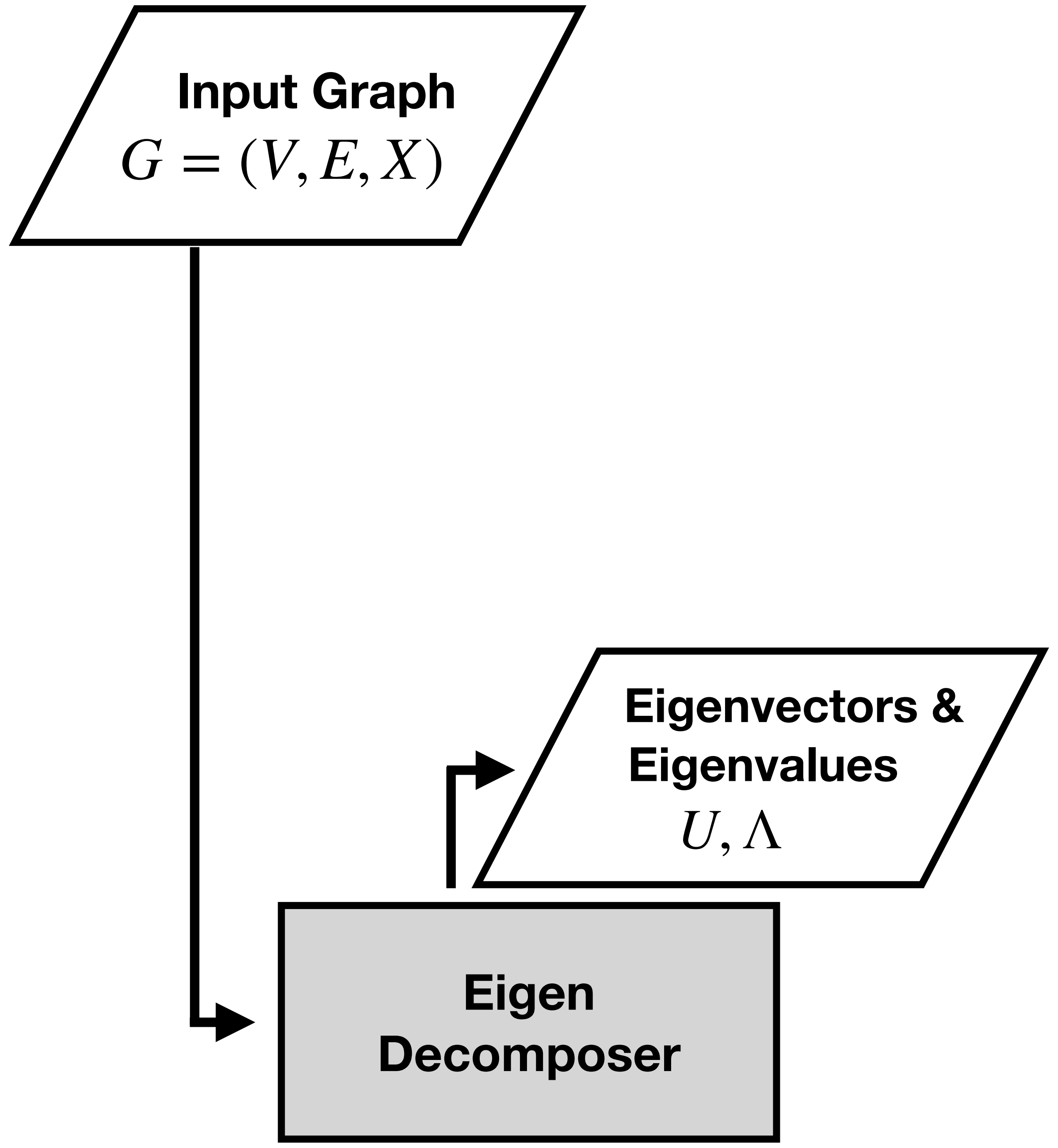
REGE: Radius Enhanced Graph Embeddings

- What does REGE do?
- How does it measure uncertainty in data?
- How does it measure uncertainty in the model?
- How does it incorporate uncertainty?
- How effective is it?

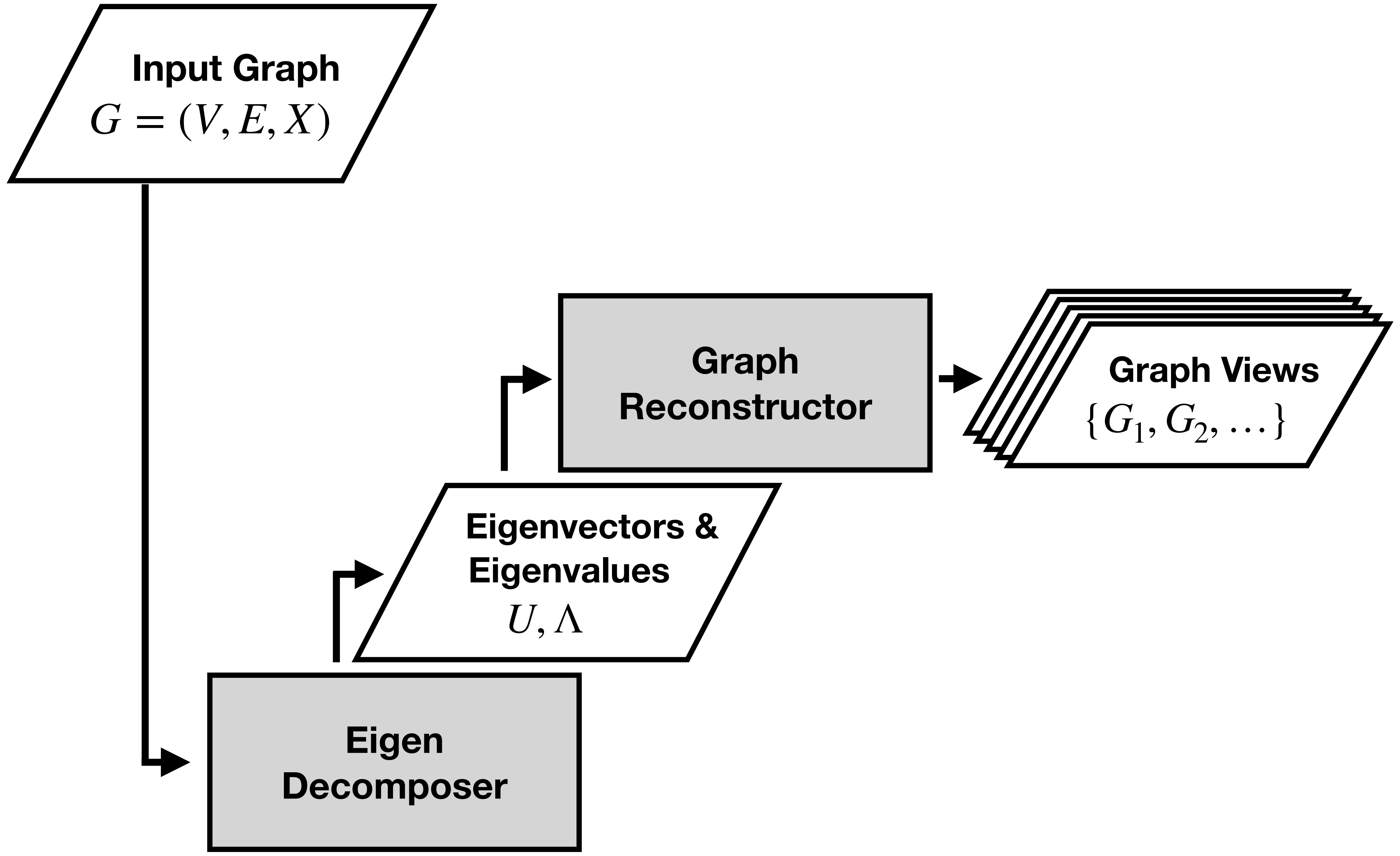
How does REGE measure uncertainty in data?



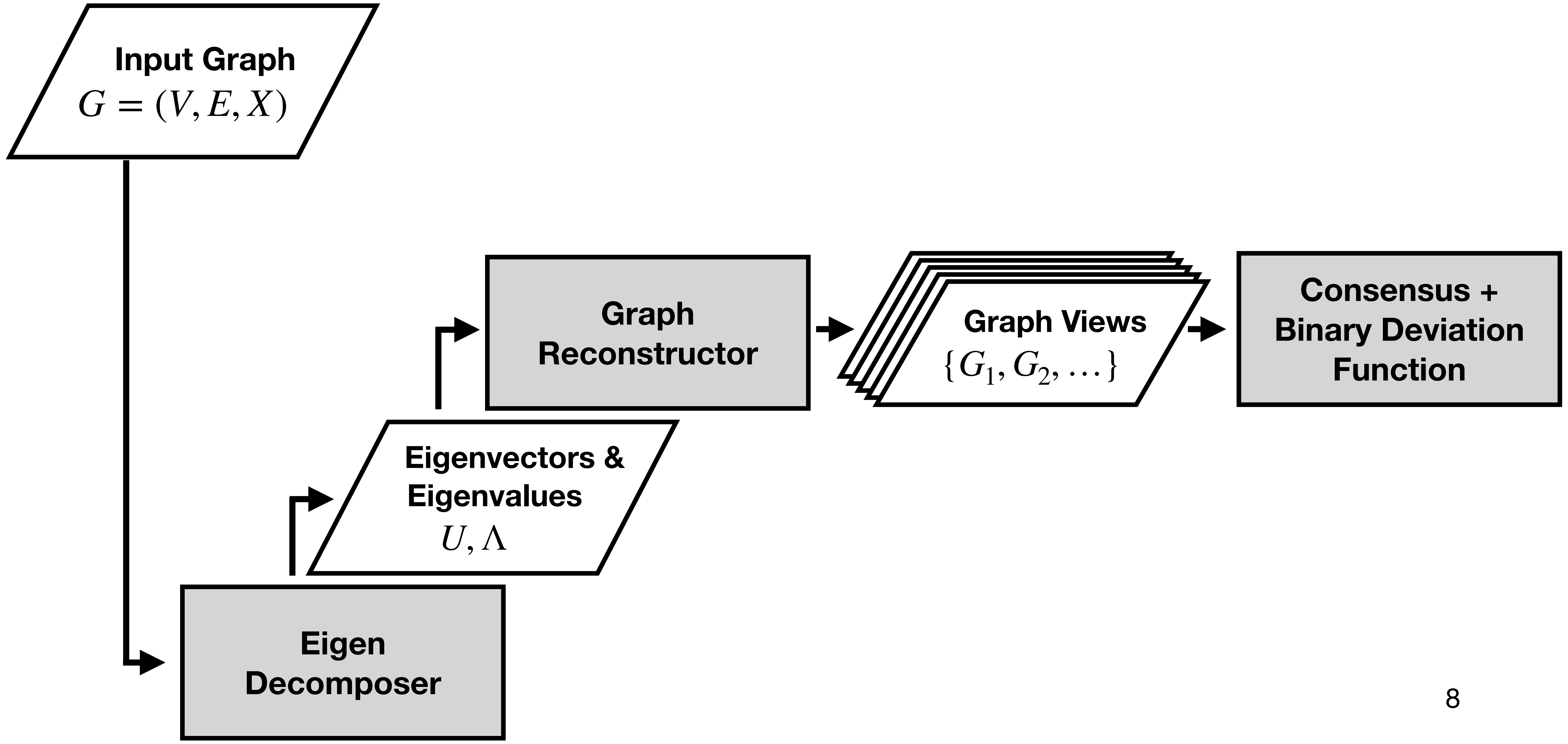
How does REGE measure uncertainty in data?



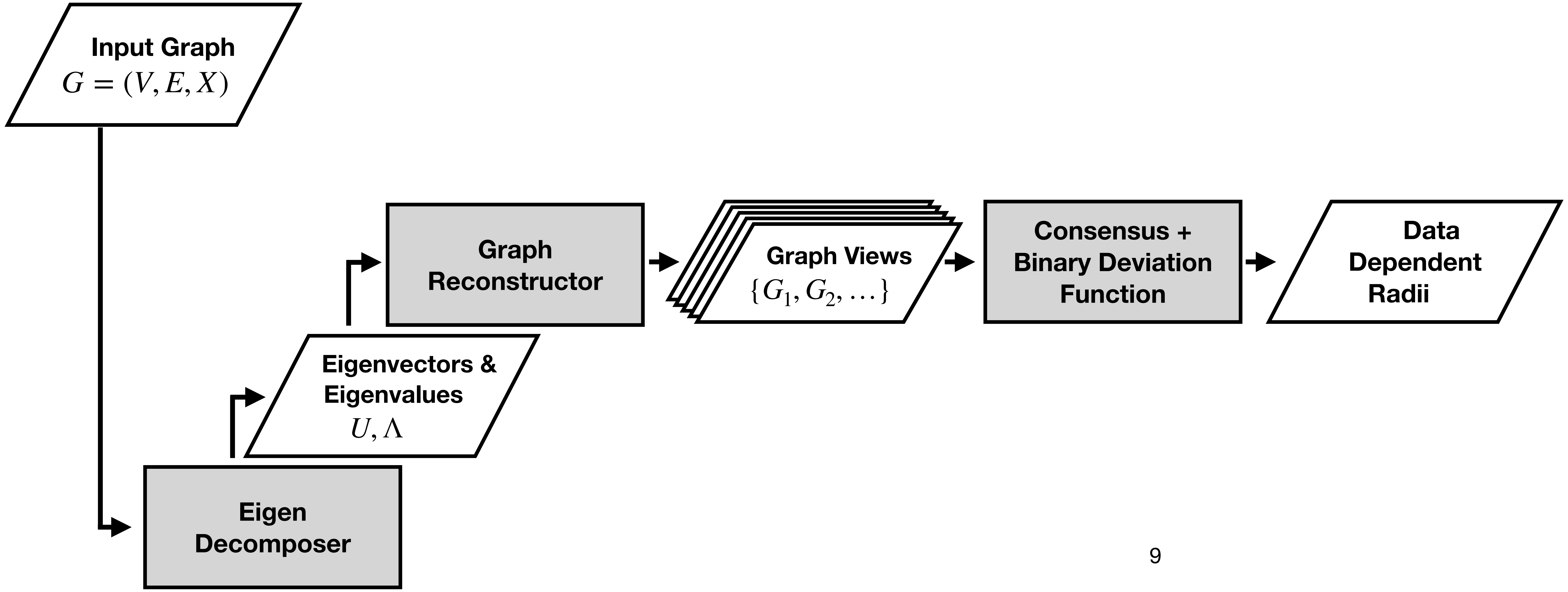
How does REGE measure uncertainty in data?



How does REGE measure uncertainty in data?

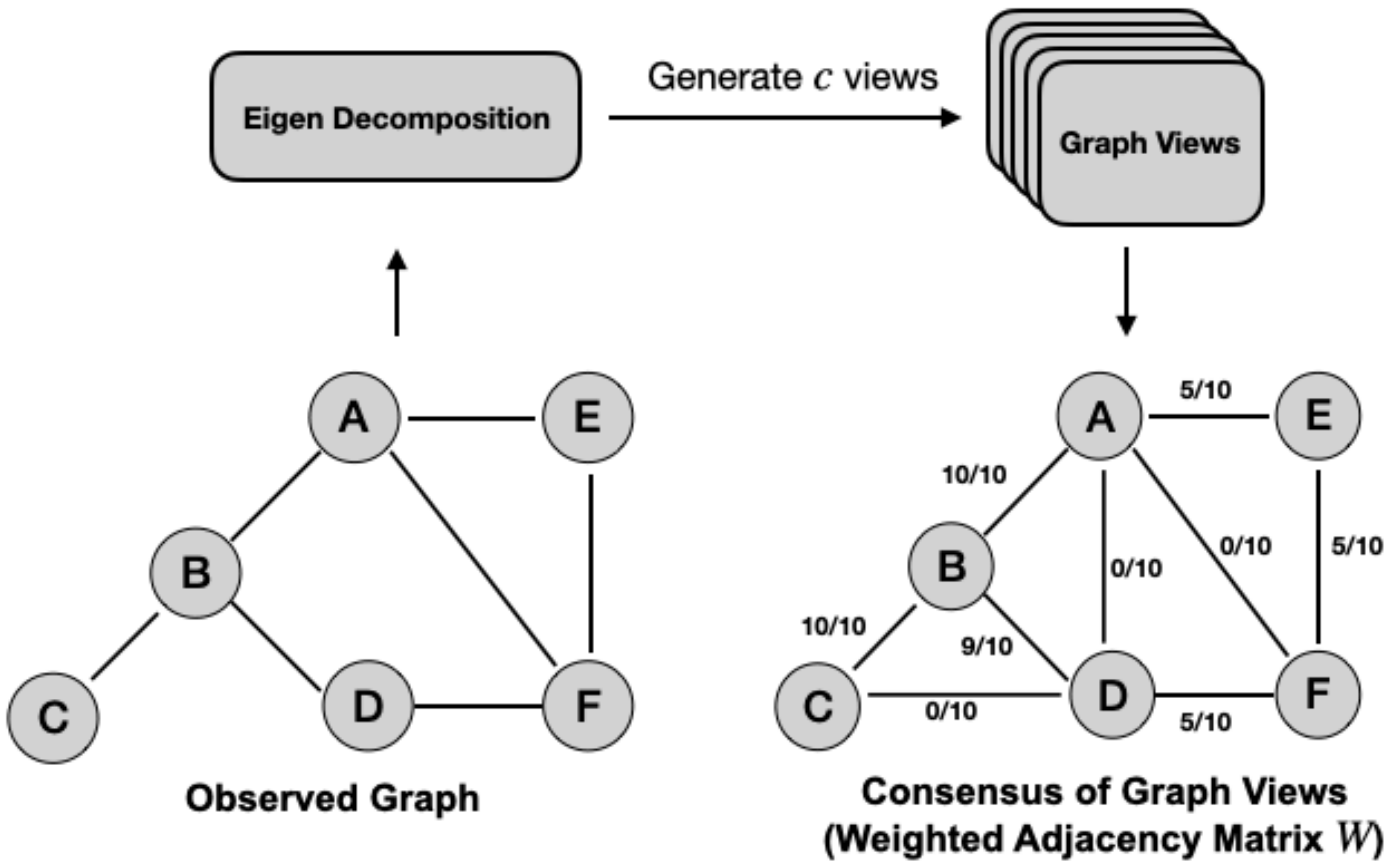


How does REGE measure uncertainty in data?



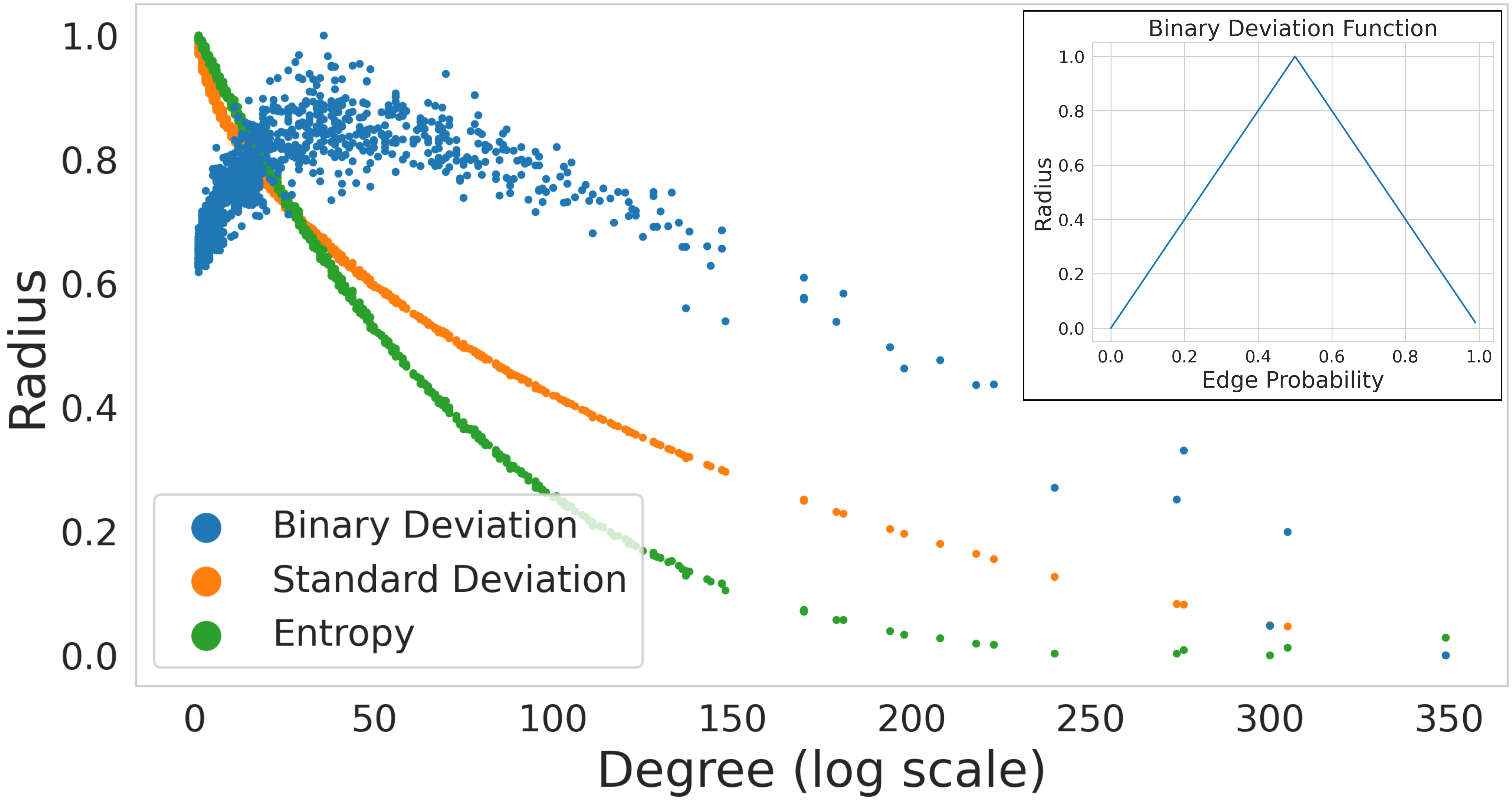
Data-dependent Radii (1/2)

- Given a graph G, compute its eigen-decomposition
- Reconstruct views of the graph $(G_1, G_2, G_3, \ldots)$
- Compute weighted adjacency matrix W, by averaging the adjacency matrices of each reconstructed graph

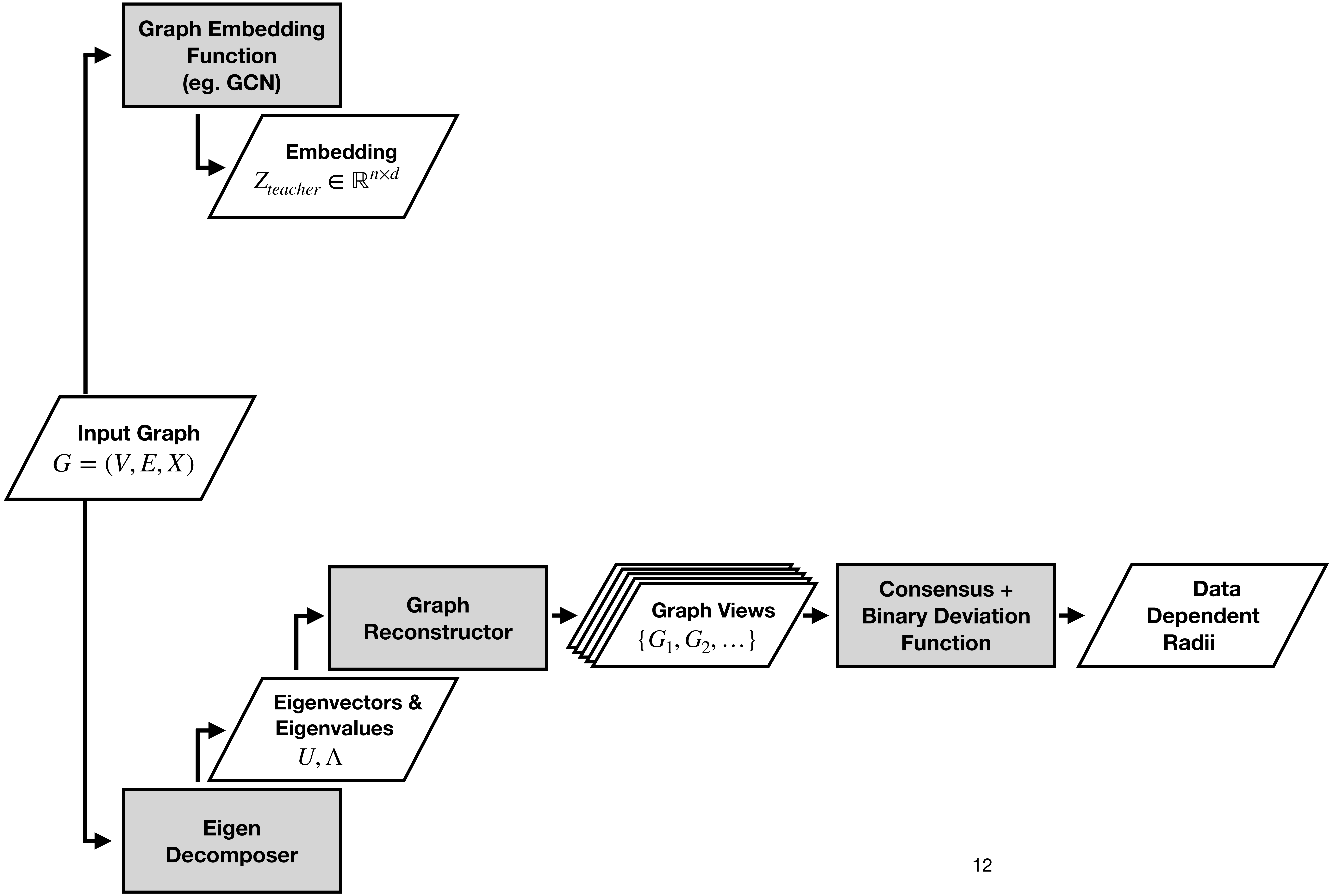


Data-dependent Radii (2/2)

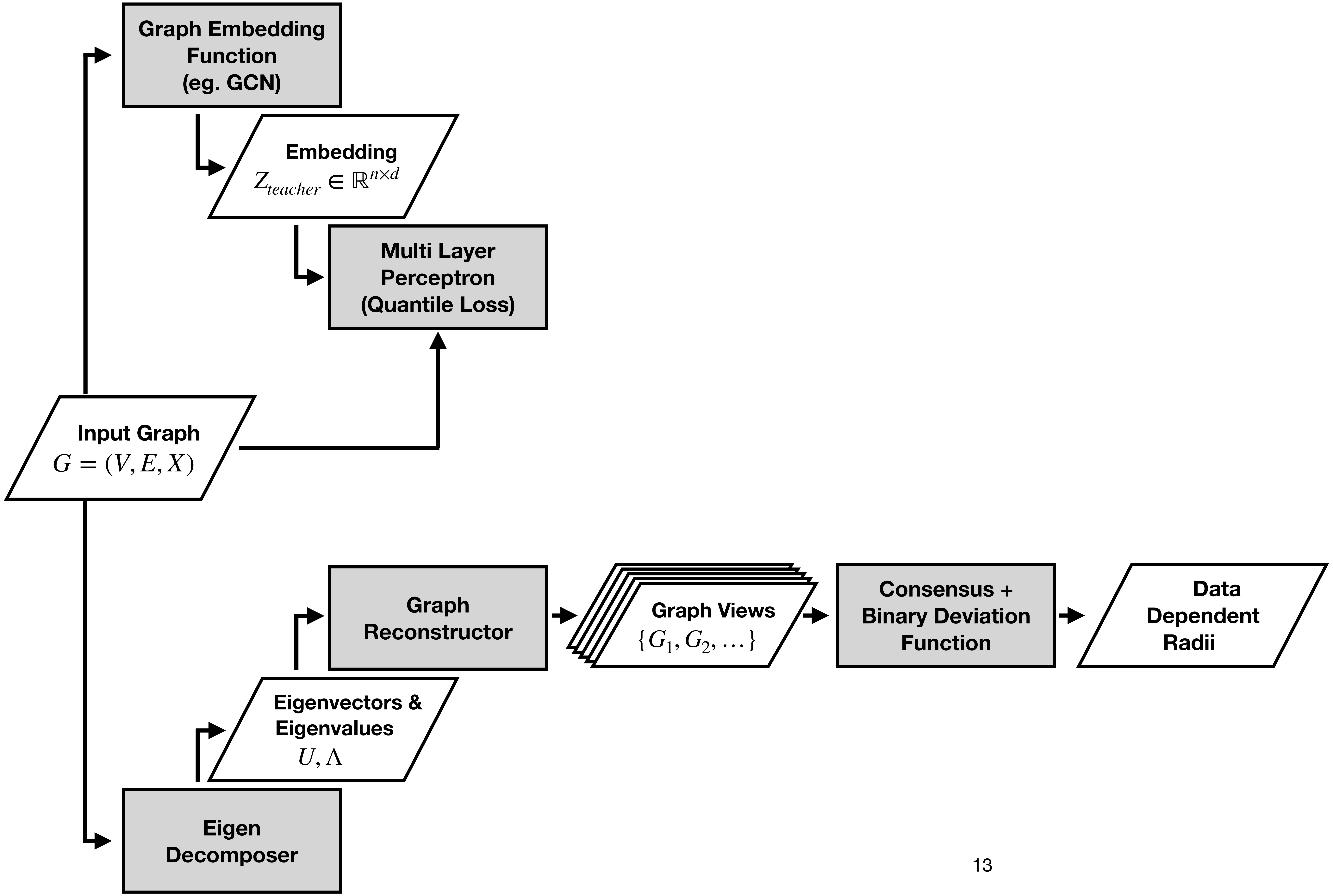
Given W, compute uncertainty for each edge e between nodes i, j using the binary deviation function: $u_e = 1 - |W_{ij} - (1 - W_{ij})|$



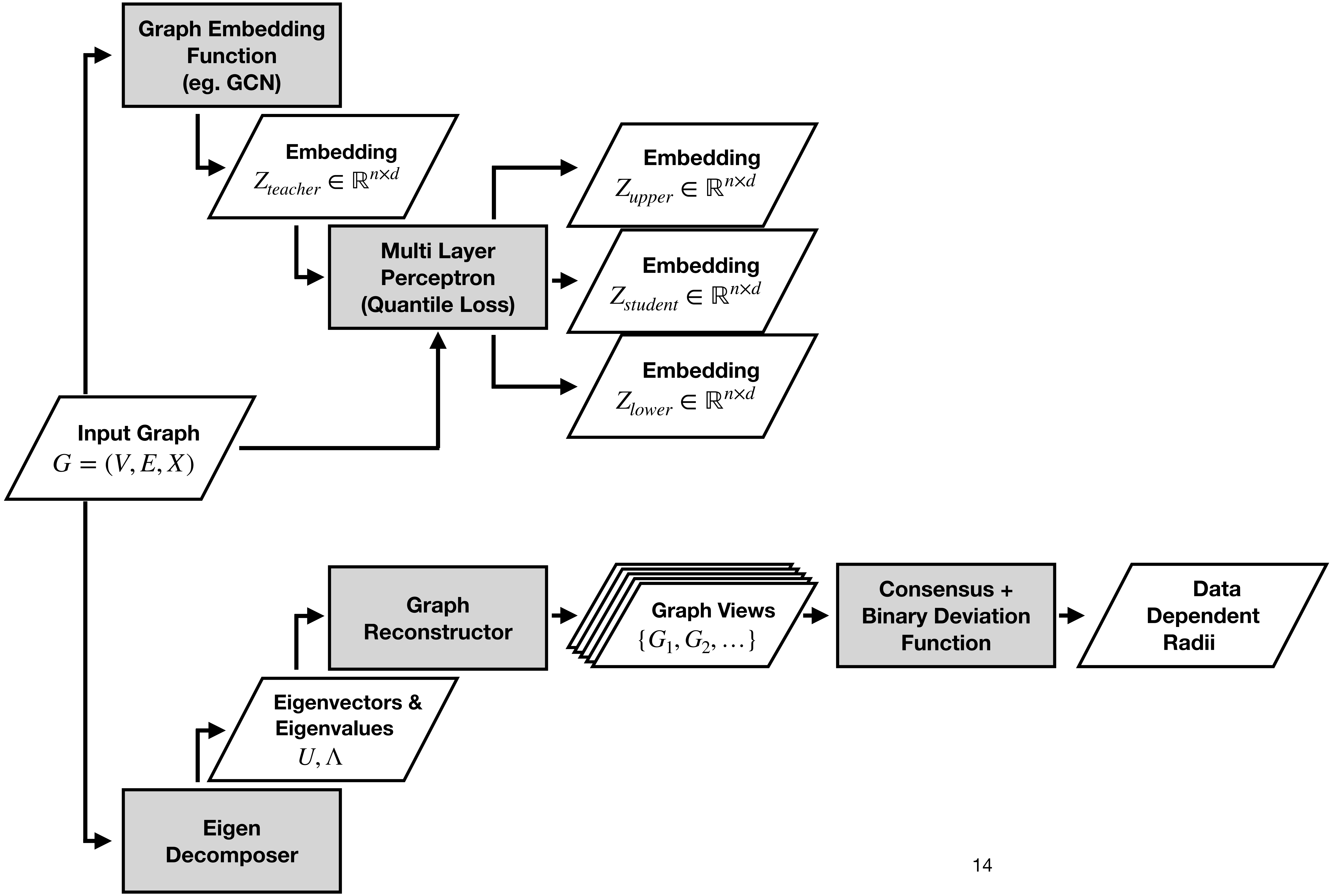
How does REGE measure uncertainty in the model?



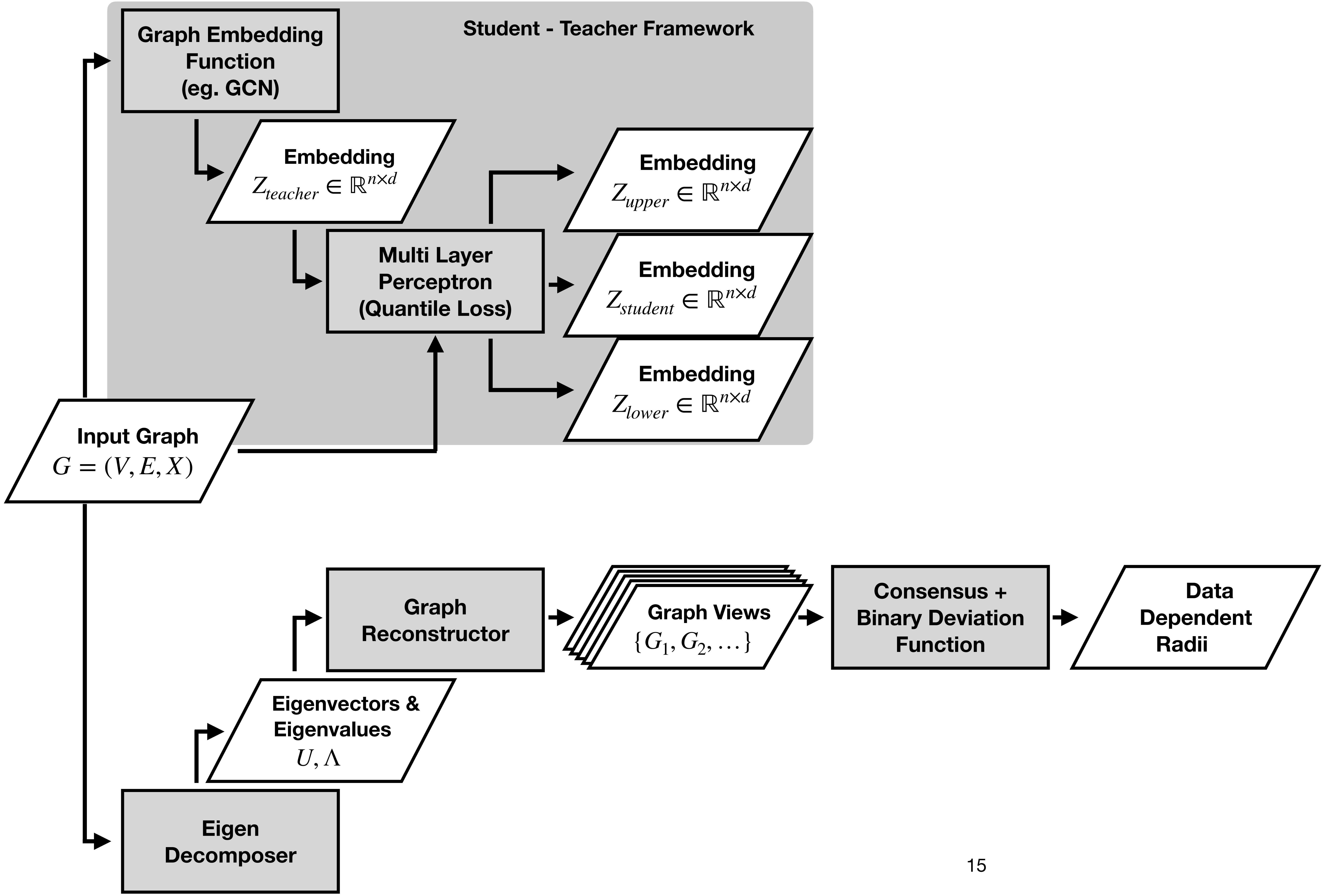
How does REGE measure uncertainty in the model?



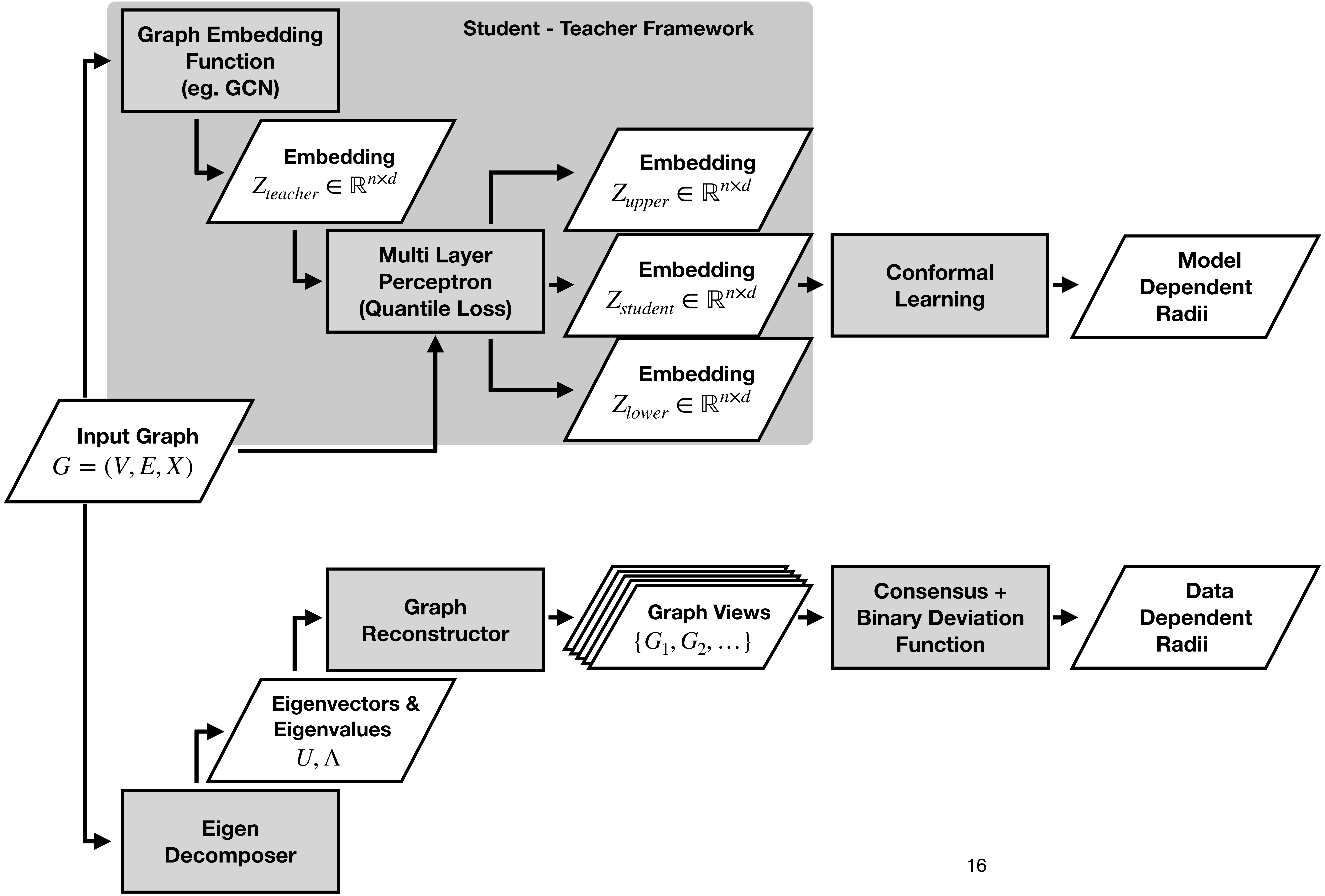
How does REGE measure uncertainty in the model?



How does REGE measure uncertainty in the model?



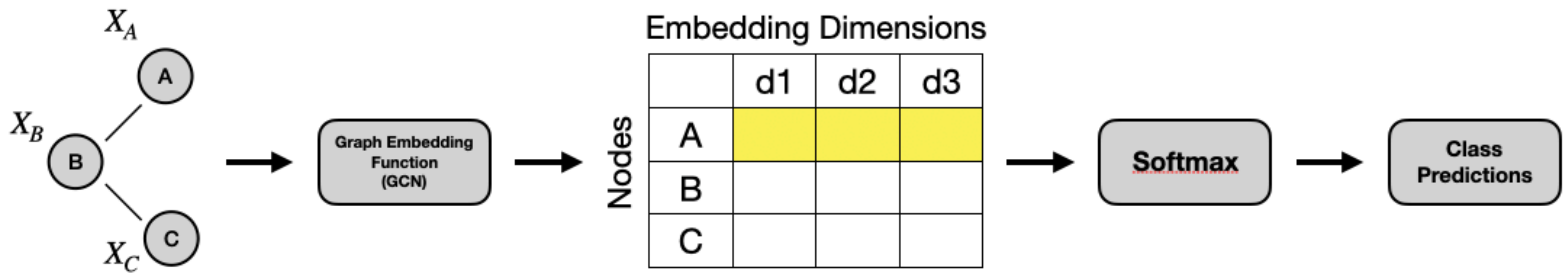
How does REGE measure uncertainty in the model?



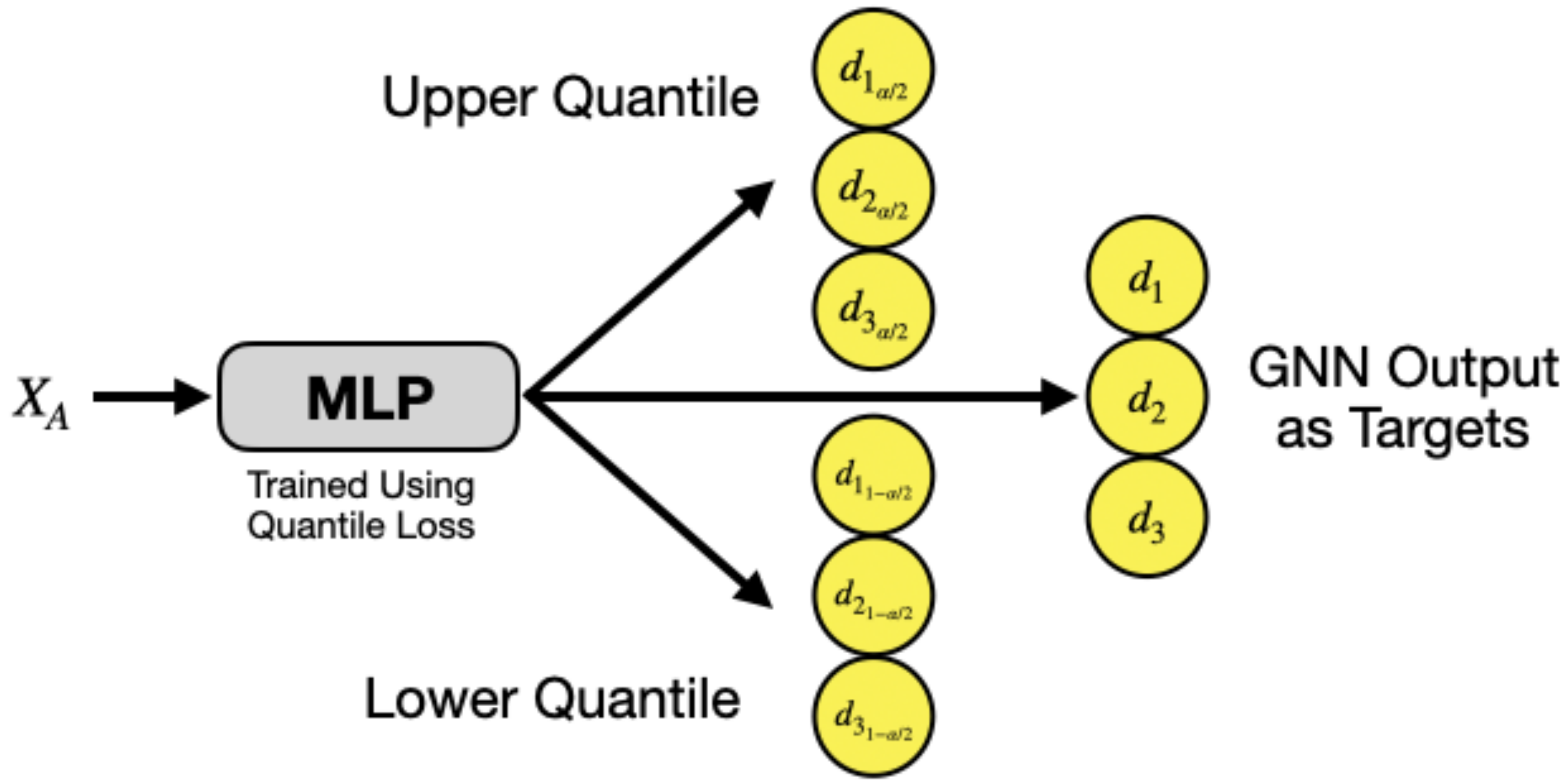
Model-dependent Radii

- Capture uncertainty around **<u>each embedding dimension</u>**
- Student-Teacher Framework
  - Learn an MLP to predict dimensions of a pre-trained GCN using quantile regression
  - This MLP predicts **<u>upper</u>** and **<u>lower</u>** quantiles
- Conformal learning used to <u>refine distance</u> between quantiles
- <u>Distance between upper and lower quantile</u> is considered as the uncertainty for that dimension

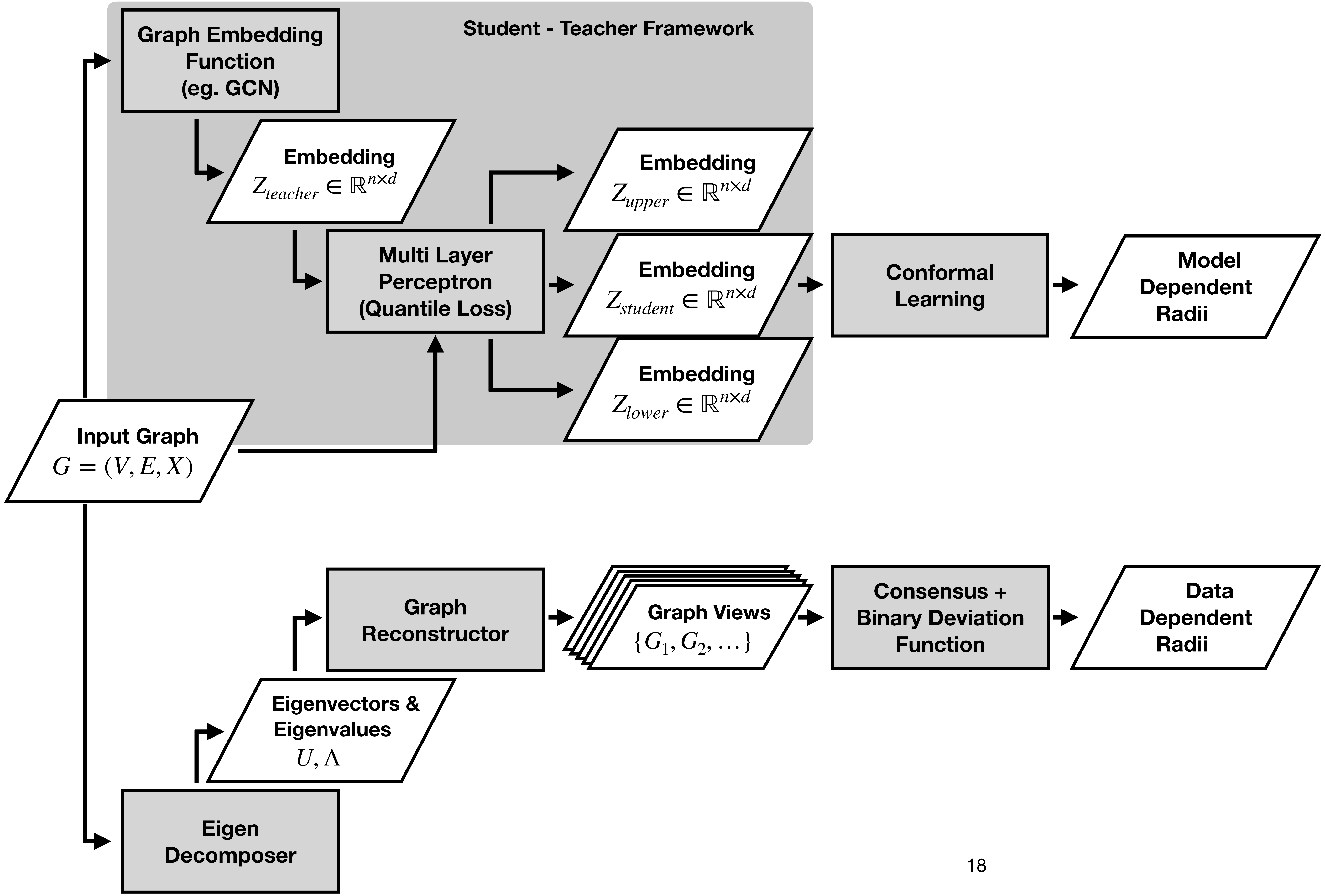
Teacher Model - A Standard GNN



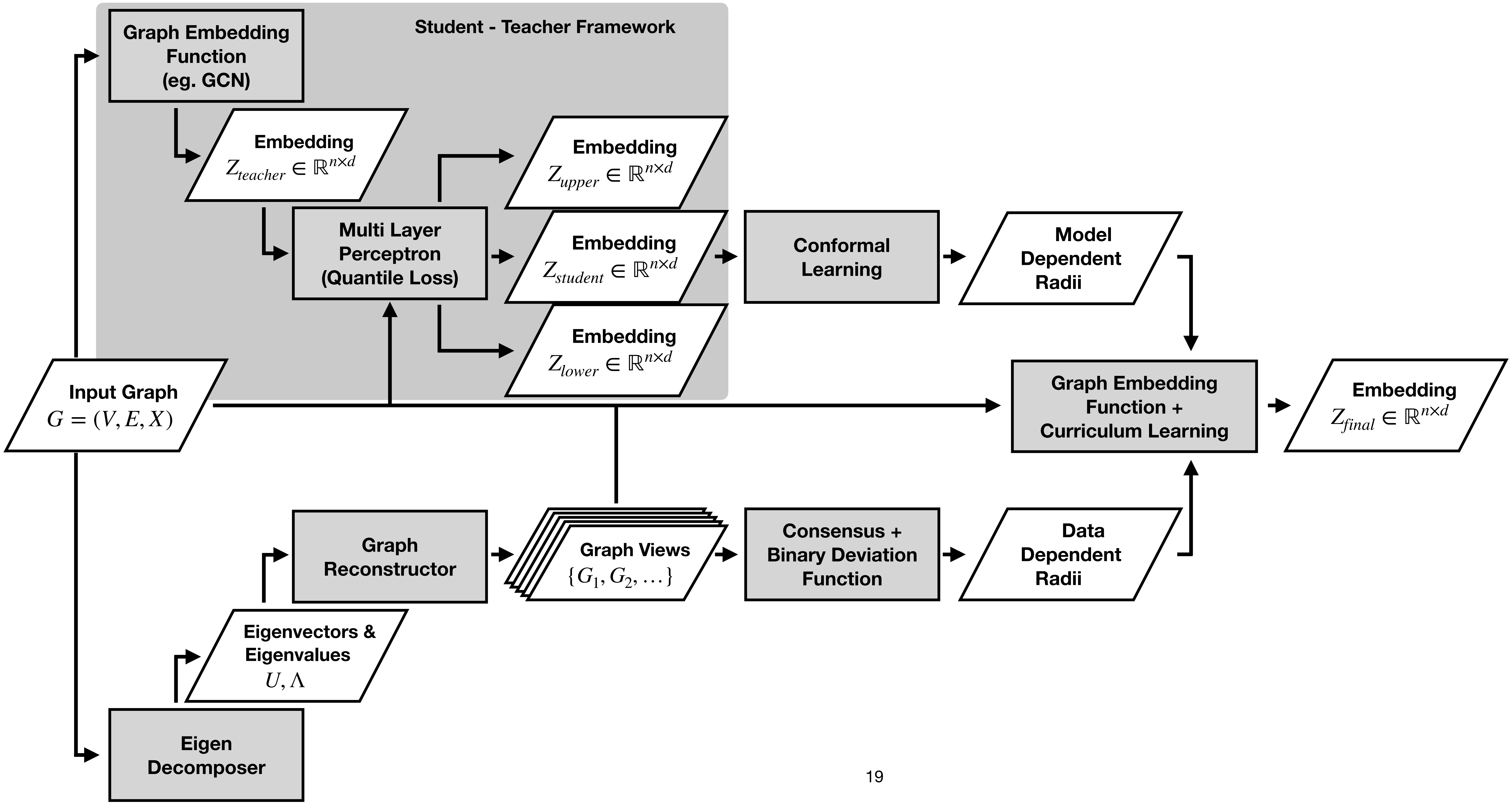
Student Model - Multi-layer Perceptron



How does REGE incorporate uncertainty?



How does REGE incorporate uncertainty?



How does REGE incorporate uncertainty?

Noise

- REGE **adds noise** to hidden layer representations of each node.
- This noise is proportional to the radius value of each node.
  - Nodes with low radius values have relatively stable embeddings.
  - Nodes with large radii have relatively unstable embeddings.
- This controlled instability makes the model learn robust representations for the nodes.

$$x_i^l \leftarrow x_i^l + \mathcal{N}(0, r_i)$$

How does REGE incorporate uncertainty?

Noise

- REGE **adds noise** to hidden layer representations of each node.
- This noise is proportional to the radius value of each node.
  - Nodes with low radius values have relatively stable embeddings.
  - Nodes with large radii have relatively unstable embeddings.
- This controlled instability makes the model learn robust representations for the nodes.

$$x_i^l \leftarrow x_i^l + \mathcal{N}(0, r_i)$$

Curriculum Learning

- Recall that we reconstruct multiple views of a graph.
- Graph G_1 is reconstructed by using the fewest components is the simplest graph with edges with high certainty
- As more components are added, so is more detail and smaller communities [1][2][3].
- We train the model starting on the simplest graph G_1 followed by G_2 and so on.

[1] S. Sawlani, L. Zhao, and L. Akoglu, "Fast attributed graph embedding via density of states," in *ICDM* (2021)

[2] M. Cucuringu and M. W. Mahoney, "Localization on low-order eigenvectors of data matrices," *arXiv preprint* arXiv:1109.1355 (2011)

[3] M. Mitrovi´c and B. Tadi´c, "Spectral and dynamical properties in classes of sparse networks with mesoscopic inhomogeneities," *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics* (2009)

All together

REGE: Radius Enhanced Graph Embeddings

- What does REGE do?
- How does it measure uncertainty in data?
- How does it measure uncertainty in the model?
- How does it incorporate uncertainty?
- How effective is it?

- REGE (D): Data-dependent Radii
- REGE (M): Model-dependent Radii

Evaluation on PolBlogs

(See paper for more results.)

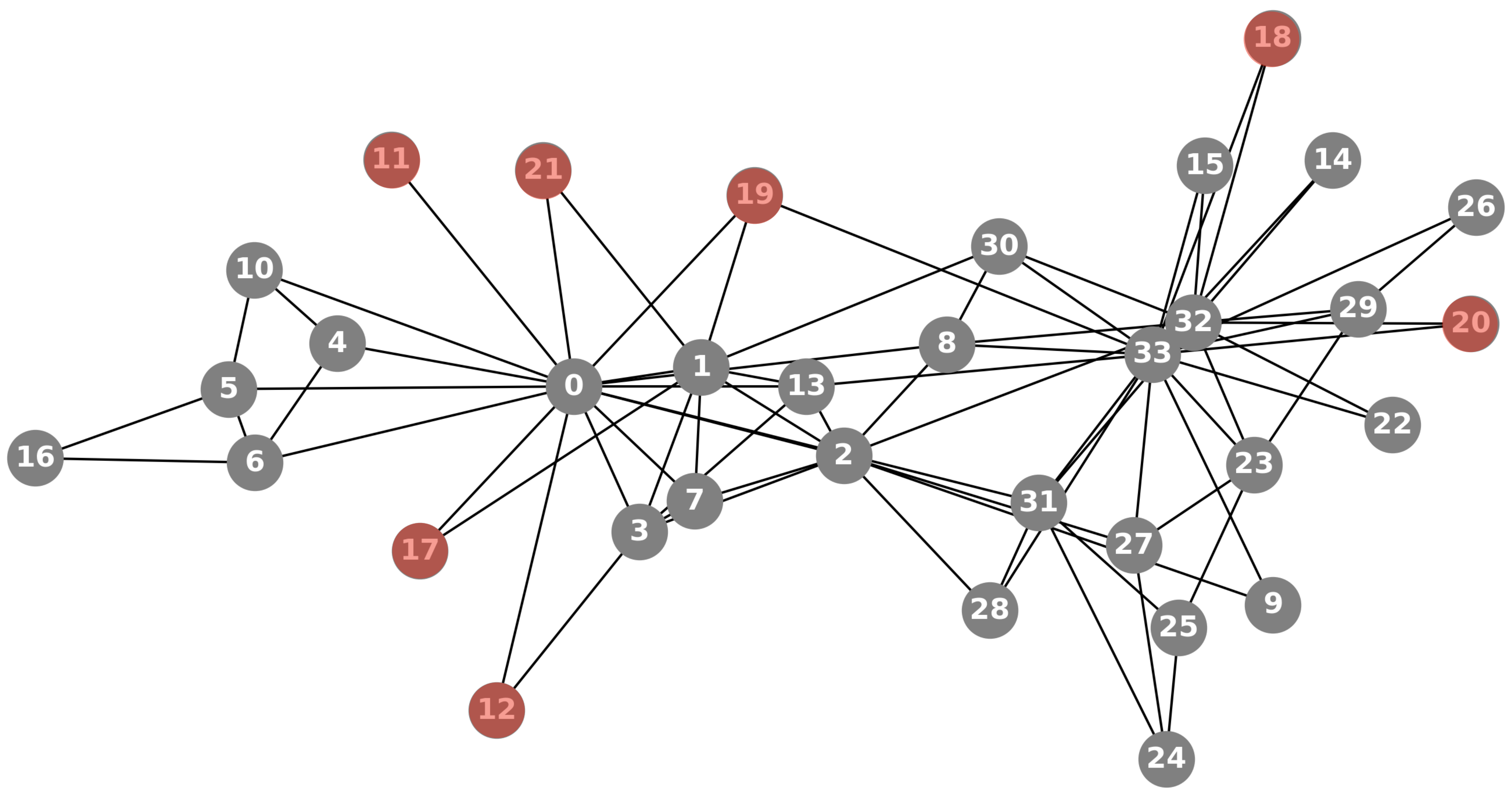
- Methods in the red box are some of the well-established defense methods
- Highlighted in blue are recent methods
- Best results are shown in **bold**, with second best underlined.

Method	MinMax (1%)	MinMax(10%)	Meta (1%)	Meta(10%)	GraD (1%)	GraD (10%)
GCN	$.944 \pm .001$	$.871 \pm .002$	$.859 \pm .002$	$.726 \pm .004$	$.876 \pm .005$	$.795 \pm .002$
RGCN	$.936 \pm .002$	$.854 \pm .002$	$.850 \pm .002$	$.699 \pm .007$	$.866 \pm .003$	$.811 \pm .003$
GCN-SVD	$.939 \pm .005$	$\underline{.885 \pm .002}$	$.926 \pm .002$	$.894 \pm .007$	$.883 \pm .004$	$\mathbf{.865 \pm .003}$
GNNGuard	$\mathbf{.950 \pm .004}$	$.861 \pm .001$	$.854 \pm .002$	$.707 \pm .014$	$.855 \pm .005$	$.812 \pm .002$
ProGNN	$.935 \pm .017$	$.869 \pm .029$	$\underline{.936 \pm .023}$	$.823 \pm .055$	$.829 \pm .029$	$.859 \pm .005$
GADC	$.512 \pm .008$	$.512 \pm .008$	$.512 \pm .008$	$.512 \pm .008$	$.498 \pm .009$	$.497 \pm .014$
GraphReshape	$.935 \pm .007$	$.847 \pm .002$	$.850 \pm .006$	$.694 \pm .002$	$.851 \pm .003$	$.803 \pm .004$
Ricci-GNN	$.941 \pm .004$	$.874 \pm .004$	$.932 \pm .003$	$.928 \pm .010$	$.875 \pm .011$	$\mathbf{.865 \pm .008}$
REGE (D)	$\underline{.946 \pm .004}$	$\mathbf{.890 \pm .004}$	$\mathbf{.946 \pm .007}$	$\mathbf{.950 \pm .005}$	$\underline{.887 \pm .002}$	$\mathbf{.865 \pm .003}$
REGE (M)	$\underline{.929 \pm .009}$	$.880 \pm .006$	$.931 \pm .017$	$\underline{.942 \pm .017}$	$\mathbf{.889 \pm .002}$	$\underline{.861 \pm .004}$

REGE on Karate Club Network

Data vs. Model Dependent Radii

- DDR - Data-dependent radii
- MDR - Model-dependent radii
- Low degree nodes (red) show low DDR - possibly due to consistent edge reconstruction
- The same nodes show high MDR - indicating that GNNs may not do as well for low degree nodes



Node	Degree	DDR	MDR
0	16	1.0	0.25
1	9	0.41	0.13
2	10	0.5	0.21
3	6	0.45	0.55
4	3	0.16	0.33
5	4	0.22	0.42
6	4	0.19	0.37
7	4	0.24	0.1
8	5	0.29	0.06
9	2	0.0	0.0
10	3	0.16	0.4
11	**1**	**0.01**	**0.34**
12	**2**	**0.02**	**0.26**
13	5	0.31	0.19
14	2	0.07	0.24
15	2	0.08	0.3
16	2	0.12	0.39
17	**2**	**0.0**	**0.3**
18	**2**	**0.0**	**0.27**
19	**3**	**0.06**	**0.23**
20	**2**	**0.04**	**0.35**
21	**2**	**0.03**	**0.32**
22	2	0.06	0.0
23	5	0.25	0.12
24	3	0.13	0.06
25	3	0.02	0.03
26	2	0.08	0.12
27	4	0.21	0.16
28	3	0.1	0.09
29	4	0.23	0.55
30	4	0.26	0.22
31	6	0.2	0.12
32	12	0.73	1.0
33	17	0.98	0.71

REGE: Takeaway points

Paper

Slides

- REGE improves the robustness of graph embeddings.
- How?
  - It incorporates data and model uncertainty during training.
- How effective is it?
  - It outperforms state of the art methods in terms of node classification accuracy on adversarially attacked datasets.