



THE UNIVERSITY OF
CHICAGO

DEPARTMENT OF STATISTICS

Master's Thesis Presentation

Tommaso Castellani

Department of Statistics
The University of Chicago

“Bias, Measurement Error, and Double-Dipping: When Can GNN
Convolutions Help Brain Connectome Prediction?”

May 11, 2026, at 2:30 PM
Jones 111, 5747 S. Ellis Avenue

Abstract

Graph Neural Networks (GNNs) are an important tool for fMRI-based prediction, where empirical functional connectivity matrices serve both as node features and as the basis for graph construction. However, recent studies (and our own experiments) show that these methods often fail to outperform simpler graph-agnostic baselines. We propose a statistical explanation for this failure by formulating connectome prediction as an errors-in-variables (EIV) problem: the true connectome is latent, while the observed covariance or correlation matrix is a noisy finite-time-series estimate. In a linear asymptotic setting, we decompose the effect of graph convolution into three mechanisms: smoothing-induced bias and variance reduction, graph estimation error, and a double-dipping effect arising when the same noisy matrix determines both topology and features. Under a latent community model adapted to this neuroscience problem, we identify a narrow favorable regime in which message passing can improve prediction: the regression target must be community-smooth, and the time-series length must lie in a window between the graph-estimation and raw-noise limits. Simulations validate this decomposition by isolating oracle, independently estimated, and same-sample graphs, while real-data experiments illustrate that standard connectome GNN pipelines often fall outside this regime. Our results clarify when graph convolutions can help brain connectome prediction, and why they often do not.