Approximation Theory and Special Functions
ATSF 2026 Conference - 9th Series
September 2-5, 2026, Ankara - Türkiye

ℓ_0-Regularized Random Graphical Modeling for Sparse Microbiome Network Inference on AGP and HMP Data

Mahdieh Farzin Asanjan(1), Vilda Purutçuoğlu(2) and Ernst Wit(3)

(1) *Department of Biomedical Engineering, Middle East Technical University, Ankara, Türkiye*
e-mail: mah.farzin2010@gmail.com

(2) *Department of Statistics, Middle East Technical University, Ankara, Türkiye*
e-mail: vpurutcu@metu.edu.tr

(3) *Institute of Computational Science, Università della Svizzera Italiana, Lugano, Switzerland*
e-mail: e.c.wit@rug.nl

Abstract

Microbiome interaction network inference is a challenging problem due to high dimensionality, compositional constraints, and sparsity observed in microbial abundance data. In graphical modeling frameworks, microbial interactions are represented through conditional dependency structures encoded by the precision matrix $\Theta = \Sigma^{-1}$, where zero off-diagonal elements indicate conditional independence relationships between microbial taxa.

In this study, we investigate sparse precision matrix estimation within an ℓ_0-regularized Random Graphical Modeling (RGM) framework for microbiome network inference. The microbial interaction network is represented as a random graph $G = (V, E)$, where V denotes microbial taxa and E represents microbial interactions. Unlike conventional Gaussian graphical models based on ℓ_1-penalized graphical lasso approaches, the proposed framework promotes exact sparsity by directly penalizing the number of nonzero elements in the precision matrix. The estimation procedure is formulated through the penalized likelihood optimization problem

$$\log|\Theta| - \mathrm{tr}(S\Theta) - \lambda\|\Theta\|_0,$$

where S denotes the sample covariance matrix, $\mathrm{tr}(S\Theta)$ measures data fidelity, and the ℓ_0 penalty explicitly controls graph sparsity. The sparsity structure is modeled probabilistically through an edge inclusion parameter and estimated under sparse graphical assumptions. Since ℓ_0-regularized optimization leads to a non-convex optimization problem, sparse-step approximation techniques are employed to obtain stable sparse solutions. The proposed methodology is compared against SPIEC-EASI and conventional ℓ_1-based graphical modeling approaches for sparse microbial network inference. Previous studies demonstrated promising performance of the proposed ℓ_0-based framework on simulated toy datasets under different sparsity scenarios. In this work, the method is applied for the first time to large-scale real microbiome datasets, including stool samples from the American Gut Project and multiple body-site samples from the Human Microbiome Project. A unified preprocessing pipeline involving filtering, centered log-ratio transformation, and variance-based feature selection is employed to ensure consistent comparison across methods.

Experimental results indicate that the proposed framework captures sparse topological structures more effectively than conventional ℓ_1-based graphical models, producing more interpretable and biologically plausible microbial interaction networks while preserving core conditional dependency relationships. The findings suggest that sparsity-aware random graphical modeling provides a flexible mathematical framework for high-dimensional microbiome network inference.

Keywords: Microbiome networks, random graphical models, graphical lasso, SPIEC-EASI.