Shuyang Gong

PhD candidate, School of Mathematical Sciences, Peking University, Beijing, China
gongshuyang@stu.pku.edu.cn

Research Interests

My research interest is probability theory, and its applications into statistics, statistical physics and theoretical computer science.

Education

Peking University, Beijing, China — September, 2021 — June, 2026(expected)
PhD in Probability and Statistics

Shandong University, Jinan, China — September, 2017 — June, 2021
Bachelor in Statistics (with honor): GPA ranked 1st/132

Publications

- **A polynomial-time approximation scheme for the maximal overlap of two independent Erdős-Rényi graphs.**
  Random Structures and Algorithms (2024), 1-38. https://doi.org/10.1002/rsa.21212
  Coauthors: Jian Ding and Hang Du
  Abstract: We presented a polynomial-time algorithm that finds a vertex correspondence which maximizes the overlap of two independent Erdős-Rényi graphs with a constant arbitrarily close to 1 compared with the asymptotic of the maximal overlap. This result gives a new example to the few problems that efficient algorithms exist for random instances while worst-cases are known to be NP-hard.
- **The algorithmic phase transition of random graph alignment problem.**
  Preprint: https://arxiv.org/abs/2307.06590, *submitted*
  Coauthors: Hang Du and Rundong Huang
  Abstract: We study the graph alignment problem over two independent Erdős-Rényi graphs on n vertices, with edge density p falling into two regimes separated by the critical window around $p_c = \sqrt{\log n/n}$. Our result reveals an algorithmic phase transition for this random optimization problem: polynomial-time approximation schemes exist in the sparse regime, while statistical-computational gap emerges in the dense regime. Additionally, we establish a sharp transition on the performance of online algorithms for this problem when p lies in the dense regime, resulting in a $\sqrt{8/9}$ multiplicative constant factor gap between achievable and optimal solutions.
- **The Umeyama algorithm for matching correlated Gaussian geometric models in the low-dimensional regime.**
  Preprint: https://arxiv.org/abs/2402.15095, *submitted*
  Coauthor: Zhangsong Li
  Abstract: Motivated by the problem of matching two correlated random geometric graphs, we study the problem of matching two Gaussian geometric models correlated through a latent node permutation. Specifically, given an unknown permutation π^* on $\{1,\ldots,n\}$ and given n i.i.d. pairs of correlated Gaussian vectors $\{X_{\pi^*(i)}, Y_i\}$ in R^d with noise parameter σ, we consider two types of (correlated) weighted complete graphs with edge weights given by $A_{i,j} = \langle X_i, X_j\rangle$, $B_{i,j} = \langle Y_i, Y_j\rangle$. The goal is to recover the hidden vertex correspondence π^* based on the observed matrices A and B. For the low-dimensional regime where $d = O(\log n)$, Wang, Wu, Xu, and Yolou established the information thresholds for exact and almost exact recovery in matching correlated Gaussian geometric models. They also conducted numerical experiments for the classical Umeyama algorithm. In our work, we prove that this algorithm achieves exact recovery of π^* when the noise parameter $\sigma = o(d^{-3}n^{-2/d})$, and almost exact recovery when $\sigma = o(d^{-3}n^{-1/d})$. Our results approach the information thresholds up to a $poly(d)$ factor in the low-dimensional regime.
- **A computational transition for detecting correlated stochastic block models by low-degree polynomials**
  Preprint: https://arxiv.org/abs/2409.00966, *submitted*
  Coauthor: Guanyi Chen, Jian Ding and Zhangsong Li
  Abstract: Detection of correlation in a pair of random graphs is a fundamental statistical and computational problem that has been extensively studied in recent years. In this work, we consider a pair of correlated (sparse) stochastic block models $\mathcal{S}(n, \frac{\lambda}{n}; k, \epsilon; s)$ that are subsampled from a common parent stochastic block model $\mathcal{S}(n, \frac{\lambda}{n}; k, \epsilon)$ with $k = O(1)$ symmetric communities, average degree $\lambda = O(1)$, divergence parameter ϵ, and subsampling probability s.
  For the detection problem of distinguishing this model from a pair of independent Erdős-Rényi graphs with the same edge density $\mathcal{G}(n, \frac{\lambda s}{n})$, we focus on tests based on *low-degree polynomials* of the entries of the adjacency matrices, and we determine the threshold that separates the easy and hard regimes. More precisely, we show that this class of tests can distinguish these two models if and only if $s > \min\{\sqrt{\alpha}, \frac{1}{\lambda\epsilon^2}\}$, where $\alpha \approx 0.338$ is the Otter's constant and $\frac{1}{\lambda\epsilon^2}$

is the Kesten-Stigum threshold. Our proof of low-degree hardness is based on a conditional variant of the low-degree likelihood calculation.

Awards

- National Scholarship — October, 2019/Shandong University
- National Scholarship — October, 2020/Shandong University
- Principal Scholarship — October, 2020/Shandong University
- Schlumberger Scholarship — October, 2023/Peking University
- Principal Scholarship — May, 2024/Peking University

LANGUAGE

Chinese(native), English(fluent)