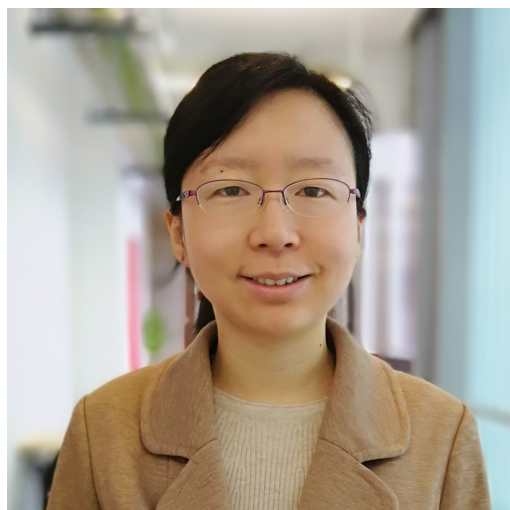


Department of Statistics Seminar Series



Tracy Ke
Department of Statistics
Harvard University

Estimating Weak Spikes by Counting Cycles

Thursday, February 26th, 2026

12:10-1:30 PM

Hill Center - 552

Zoom Meeting: Meeting ID: 926 1608 1995

Password: 153525

<https://rutgers.zoom.us/j/92616081995?pwd=TqPVh6Y1O6aLUt8i2uUBCRhrgH6Ayx.1>

Light refreshments will be served in Hill 452 11:40am

Abstract:

In low-rank plus noise models, cycle count statistics provide estimators of spectral moments and can be used for parameter estimation and statistical inference.

In the first part of this talk, I will present recent work on estimating spiked eigenvalues (spikes) using cycle count statistics. The commonly used approach based on empirical eigenvalues is known to be biased, and bias correction becomes challenging when noise variances are heterogeneous and unknown. We propose two alternative methods: an optimization-based approach and a Newton–Prony approach. By interpreting cycle count statistics as estimators of the power sums of the spikes, we apply Newton’s identities and Prony’s equations to recover individual spike values. We show that, in the weak spike regime where empirical eigenvalue–based estimators are inconsistent, the proposed methods are consistent and achieve algebraic convergence rates. Simulation studies demonstrate improved performance over existing methods in weak-signal settings, and an application to a Facebook network dataset illustrates the use of the proposed approach in community detection.

In the second part of the talk, I will discuss a problem motivated by the first part: deriving computationally efficient equivalent forms (CEEFs) for cycle count statistics. General solutions to the CEEF problem are not currently available and typically require involved combinatorial arguments. We address this problem by combining a new proof strategy with AI-assisted reasoning. The resulting analysis uses graph-theoretic techniques and leads to new explicit formulas in general settings. We also examine the role of AI in this process, showing that, while AI alone is insufficient to solve the problem, it can be effective when guided by a clearly specified strategy and structured prompts. For illustration, we focus on DeepSeek-R1 and include comparisons with other AI-based approaches.

Bio:

Tracy Ke is currently Associate Professor of Statistics at Harvard University. She received her PhD from Princeton University in 2014. Prior to joining Harvard in 2018, she was Assistant Professor of Statistics at The University of Chicago. Her research interests include network data analysis, high-dimensional statistics, text mining, and machine learning. One of her recent research interests is developing statistically optimal methods for analysis of network and text data. She was the recipient of NSF CAREER Award, ASA Noether Young Scholar Award, IMS Peter Hall Prize, COPSS Emerging Leader Award, and Sloan Research Fellowship.