Shanyun Gao

Ph.D. Student in Statistics, Purdue University · West Lafayette, IN
gao565@purdue.edu · shanyun-gao-l.github.io · Google Scholar · GitHub

Research Focus

Causal discovery for time series (periodic, stationary, and change-point settings): three first-author NeurIPS/AISTATS papers with open-source code. Currently studying causal in-context learning in transformers.

Education

Ph.D. in Statistics, Purdue University Aug. 2021 – Present
Advisors: Qifan Song (Purdue) & Murat Kocaoglu (Johns Hopkins) *GPA 3.93/4.00 · Qualifying Exam ranked 1st*
M.S. in Applied Statistics, Purdue University (GPA 4.00/4.00) Aug. 2019 – May 2021
B.S. in Applied Statistics, Shanghai University of Finance and Economics Sep. 2015 – Jun. 2019

Research Experience

Research Collaboration with Adobe Research Aug. 2022 – Jan. 2026
Collaborators: R. Addanki, T. Yu, R. A. Rossi

- Designed constraint-based causal discovery algorithms for time series (NeurIPS'23, AISTATS'25, AISTATS'26).
- **MSSD**: exploits the structure of minimal separating sets under causal stationarity; highest F1 and lowest SHD among baselines on synthetic and river-flow data.
- **Causal-RuLSIF**: two-stage change point detection that targets shifts in a variable's causal mechanism, relaxing the usual IID assumption via the Causal Markov Condition.
- **PCMCI$_\Omega$**: non-parametric discovery for periodically alternating causal mechanisms; provably sound for discrete-valued series and applied to climate data.

Graduate Research Assistant, CausalML Lab, Purdue University Aug. 2022 – Aug. 2023
PI: Dr. Murat Kocaoglu

- Formulated the semi-stationary structural causal model; implemented and evaluated PCMCI$_\Omega$.

Graduate Research Assistant, College of Health & Human Sciences, Purdue Aug. 2020 – May 2021

- Developed simulation-based power analysis procedures for multilevel models in a five-year NIH-funded project.

Selected Publications

S. Gao, R. Addanki, T. Yu, R. A. Rossi, Q. Song, M. Kocaoglu. *Stationarity-Aware Causal Discovery in Time Series via Minimal Separating Sets.* **AISTATS 2026**. [Paper]

S. Gao, R. Addanki, T. Yu, R. A. Rossi, M. Kocaoglu. *Causal Discovery-Driven Change Point Detection in Time Series.* **AISTATS 2025**. [Paper]

S. Gao, R. Addanki, T. Yu, R. A. Rossi, M. Kocaoglu. *Causal Discovery in Semi-Stationary Time Series.* **NeurIPS 2023**; oral at UAI 2023 Causal Inference for Time Series Workshop. [Paper]

S. Gao, M. Kocaoglu, Q. Song. *Causal In-Context Learning in Transformers: Training Dynamics Across Heterogeneous Interventional Data.* AISTATS 2026 Workshop on Causality in the Age of AI Scaling.

Teaching

TA, Administrative TA & Head TA, Dept. of Statistics, Purdue Aug. 2023 – Present

- Lead lab sessions for STAT 301 (Elementary Statistical Methods); guide students on statistical concepts, SPSS, and data analysis.

Honors, Service & Skills

Honors Bilsland Dissertation Fellowship (2026–27); Qualcomm Innovation Fellowship proposal abstract accepted (2026); Burr Award (2025–26); WISP Fall Travel Grant (2023–24); McAuliffe Graduate Student Scholar Award (2020–21); Applied Statistics Scholarship (2020–21)

Reviewing NeurIPS, ICML, ICLR, UAI, AAAI, AISTATS, TMLR

Skills Python, R; LaTeX, MATLAB, SPSS; causal discovery, causal in-context learning, conditional independence testing, change point detection