

♀ Baltimore, MD 🛛 🖾 awang87@jhu.edu

Education	
 Ph.D. Johns Hopkins University, Applied Mathematics & Statistics M.S.E Johns Hopkins University, Data Science B.S. Johns Hopkins University, Applied Mathematics & Statistics GPA: 4.0/4.0 Undergraduate Dean's List: Fall 2019 - Spring 2021 	Sept 2022 – Now Sept 2021 – May 2022 Sept 2019 – May 2021
Honors	
RECOMB Conference Travel Fellowship Award Rufus S. Issacs Graduate Fellowship	2025 2023 - 2024
Industry Experience	
 Center for Data Science in Emergency Medicine (JHUSOM), Research Intern Developed XGBoost model to predict the risk of acute kidney injury upon emergency department visits Enhanced the model interpretability by quantifying the importance of each clinical covariate via the Shapley value 	Baltimore, MD Jan 2022 – May 2022
 Accenture, Strategy Intern Surveyed the domestic and global natural rubber industry by analyzing the growth rate of each section of the supply chain Completed business portfolio diagnosis on twenty business segments of corporate customers 	Shanghai, China June 2020 – Aug 2020

Publications _

Wang, An, Stephanie C. Hicks, Donald Geman, and Laurent Younes. "GeneCover: A Combinatorial Approach for Label-free Marker Gene Selection." bioRxiv (2024): 2024-10. (Accepted by RECOMB 2025)

Ji, Lanlan, **An Wang**, Shreyash Sonthalia, Daniel Q. Naiman, Laurent Younes, Carlo Colantuoni, and Donald Geman. "CellCover Captures Neural Stem Cell Progression in Mammalian Neocortical Development." bioRxiv (2024).

Yu, Doudou, Manlin Li, Guanjie Linghu, Yihuan Hu, Kaitlyn H. Hajdarovic, **An Wang**, Ritambhara Singh, and Ashley E. Webb. "CellBiAge: Improved single-cell age classification using data binarization." Cell Reports 42, no. 12 (2023).

Research Projects

GeneCover: A Combinatorial Approach for Label-free Marker Gene Selection

- Proposed geneCover, a label-free combinatorial method that selects an optimal panel of minimally redundant marker genes based on gene-gene correlations
- Demonstrated that geneCover captures distinct correlation structures across the transcriptome and enhances the detection of rare cell types and highly refined spatial organizations
- Showcased the excellent scalability of geneCover, enabling efficient marker gene selection in large omics datasets

Markov Random Field Modeling of Spatial Transcriptomics Data

• Modeled the probabilistic distribution of spatial transcriptomics via Gibbs distribution and estimated the model

parameters by stochastic approximation and Markov chain Monte Carlo

• Applied the model to simulate spatial transcriptomics data, infer cell states, and conduct hypothesis testing on the spatial organization of cell types

CellCover: Minimal Coverings of Cell Types in scRNA-seq Data

- Proposed marker gene panel selection as a variation of the minimal set-covering problem on the space of cells and solved with integer programming
- Applied CellCover to generate gene panels that characterize stages in mammalian neocortical development and demonstrate they can capture cell-type specific signals in the neurogenesis

Teaching Experience _____

Teaching Assistant , Johns Hopkins University	
EN.553.633: Monte Carlo Methods	Fall 2021, 2022
EN.553.650: Computational Molecular Medicine	Spring 2022, 2023
Directed Reading Program Mentor, Johns Hopkins University	Spring 2024
 Mentee: Hyun Woo (Eric) Kim, B.S./M.S.E in Applied Mathematics and Statistics, Class of 2024 	
 Project: Markov Chains and Mixing Time 	
Directed Reading Program Mentor, Johns Hopkins University	Fall 2023
 Mentee: Tuxun (Nick) Lu, B.S. in Computer Science, Class of 2024 	
 Project: Maximum Likelihood Estimation of Energy-based Model and Diffusion Probabilistic Model 	
Presentations	
GeneCover: A Combinatorial Approach for Label-free Marker Gene Selection	
JHU Center for Imaging Science Retreat	Dec 2024
JHU Institute for Computational Medicine Retreat	Dec 2024
Learning the Probabilistic Distribution of Spatial Transcriptomics Data via Markov Randon	n Field
JHU Center for Imaging Science Retreat	Dec 2023
Minimal Coverings of Cell Types in scRNA-seq Data	
JHU AMS Graduate Student Seminar	Mar 2023
JHU Center for Imaging Science Retreat	Dec 2022

Core Qualifications _

Programming: Python, R

Language: Chinese Mandarin (native), Enghlish