

# Huan He

☎ (+1) 860-617-6686  
✉ Huan.He@Pennmedicine.upenn.edu  
📄 hehuannb.github.io

My name is Huan He, an assistant professor in the Department of Mathematics and Statistics at Auburn University. Previously, I worked as a postdoctoral researcher at University of Pennsylvania, working with Professor Yong Chen and Harvard University with Professor Marinka Zitnik. I received my PhD in Computer Science from Emory in 2022, where I was co-supervised by Professors Joyce Ho and Yuanzhe Xi.

A main thrust of my research focuses on AI for accelerating healthcare discoveries. Thus far, I have focused on two complementary directions: (1) **Scientific Computing for Machine Learning** (e.g., optimization, numerical analysis, parallel computing) (2) **AI-Accelerated Healthcare** (e.g., Electronic Health Records Phenotyping, Causal Inference, Drug Discovery)

## Professional Experience

- 08/2024-Present **Assistant Professor**, Auburn University, Department of Mathematics and Statistics.
- 09/2023-08/2024 **Postdoc**, University of Pennsylvania, Advisor: Dr. Yong Chen,  
Role: Research on Causal Inference and Machine Learning.
- 06/2022-06/2023 **Postdoc**, Harvard University, Advisor: Dr. Marinka Zitnik,  
Role: Research on Machine Learning for time series.

## Education

- 2016-2022 **Ph.D in Computer Science**, Emory University, Advisor: Dr. Joyce C. Ho and Yuanzhe Xi.
- 2014-2016 **M.Sc in Financial Mathematics**, University of Connecticut, Advisor: Dr. James G. Bridgeman.
- 2010-2014 **B.S in Financial Engineering**, Shanghai Finance University.

## Publications

- SIAMX 2024 Ziyuan Tang, Tianshi Xu, **Huan He**, Yousef Saad, Yuanzhe Xi  
AATGS: Anderson Acceleration with Truncated Gram-Schmidt  
*Accepted by SIAM Journal on Matrix Analysis and Applications*
- MLHC 2024 Yijie Hao, Joyce C. Ho, **Huan He**  
LLMSYN: Generating Synthetic Electronic Health Records Without Patient-Level Data  
*Machine Learning for Healthcare 2024*
- ICLR-W 2024 Yaoyao Xu, Xuxi Chen, Tong Wang, **Huan He**, Tianlong Chen, Manolis Kellis  
Demystify the Secret Function in Protein Sequence via Conditional Diffusion Models  
*The 12th International Conference on Learning Representations Workshop on Generative and Experimental Perspectives for Biomolecular Design*
- ICLR 2024 **Huan He**, William hao, Yuanzhe Xi, Yong Chen, Bradley Malin, Joyce Ho  
A Flexible Generative Model for Heterogeneous Tabular EHR with Missing Modality  
*The 12th International Conference on Learning Representations*
- SISC 2023 Ru Huang, Kai Chang, **Huan He**, Ruipeng Li, Yuanzhe Xi  
Reducing Operator Complexity in Algebraic Multigrid with Machine Learning  
*Accepted by SIAM Journal on Scientific Computing*
- SIMAX 2023 **Huan He**, Shifan Zhao, Ziyuan Tang, Yousef Saad, Yuanzhe Xi  
NLTGCR: A class of Nonlinear Acceleration Procedures based on Conjugate Residuals  
*Accepted by SIAM Journal on Matrix Analysis and Applications*

- NeurIPS 2023 Owen Queen, Thomas Hartvigsen, Teddy Koker, **Huan He**, Theodoros Tsiligkaridis, Marinka Zitnik  
Encoding Time-Series Explanations through Self-Supervised Model Behavior Consistency  
*The 37th Conference on Neural Information Processing Systems*
- ICML 2023 **Huan He**, Owen Queen, Teddy Koker, Consuelo Cuevas, Theodoros Tsiligkaridis, Marinka Zitnik. Domain Adaptation for Time Series Under Feature and Label Shifts.  
*The 40th International Conference on Machine Learning (ICML)*
- ICLR 2023 Jiali Cheng, **Huan He**\*, George Dasoulas\*, Chirag Agarwal, Marinka Zitnik  
GNNDelete: A General Unlearning Strategy for Graph Neural Networks. (\*: equal contribution)  
*The 11th International Conference on Learning Representations*
- UAI 2022 Yuliang Ji, Difeng Cai, **Huan He**, Yuanzhe Xi  
AUTM Flow: Atomic Unrestricted Time Machine for Monotonic Normalizing Flows  
*The 38th Conference on Uncertainty in Artificial Intelligence*
- ICLR 2022 **Huan He**, Shifan Zhao, Yuanzhe Xi, Joyce C Ho, Yousef Saad  
GDA-AM: Solve Minimax Optimizatton by Anderson Mixing  
*The 10th International Conference on Learning Representations*
- NeurIPS-W 2021 **Huan He**, Yuanzhe Xi, Joyce C Ho  
AGE: Enhancing the Convergence on GANs using Alternating extra-gradient with Nonlinear Gradient Extrapolation  
*The 35th NeurIPS Workshop on Deep Generative Models and Downstream Applications (Oral)*
- IEEE BigData 2020 **Huan He**, Yuanzhe Xi, Joyce C Ho  
Fast and Accurate Tensor Decomposition without a High-Performance Computing Machine  
*International Conference on BigData*
- ICDM-W 2020 **Huan He**, Yuanzhe Xi, Joyce C Ho. Accelerated Stochastic Gradient Decent(SGD) for Sparse Tensor Decomposition  
*IEEE International Conference on Data Mining Workshop*
- ACM WWW 2019 **Huan He**, Jette Henderson, Joyce C Ho  
SGranite: Distributed Tensor Decomposition for Large Scale Health Analytics  
*Proceedings of The Web Conference*
- AMIA 2018 Jette Henderson, **Huan He**, Bradley A Malin, Joshua C Denny, Abel N Kho, Joydeep Ghosh, and Joyce C Ho  
Phenotyping through Semi-Supervised Tensor Factorization (PSST)  
**AMIA 2018** —*AMIA Annual Symposium*

---

## Awards and Grants

- 2023 **NSF Travel Grant**, *University of Illinois Urbana-Champaign*,  
Invited Presentation at the Workshop on Sparse Tensor Computations.
- 2023 **NSF Travel Grant**, *Brown University* ,  
Invited Presentation at the fourth Mathematical and Scientific Machine Learning conference.
- 2019 **Best Poster Mention Award**, *Georgia Scientific Computing Symposium*,  
Topic: Distributed Tensor Decomposition for Large Scale Health Analytics .
- 2013 **Outstanding Winner of 5th Challenge Cup**, *Shanghai Finance University*,  
Topic: Volatility Modeling of Junk Bonds using GARCH.
- 2012, 2013 **Merit Student Award**, *Shanghai Finance University*.

2011, 2012, 2013, **First-Class Scholarship**, *Shanghai Finance University*.  
2014

---

## High-level Research Projects

2018-2023 **Scientific Computing for Machine Learning** .

- **Nonlinear Acceleration Methods and Optimization Techniques:**
  - Focused on optimization methods for complex problems
  - Developed a nonlinear Truncated Generalized Conjugate Residual method (nlTGCR) to reduce memory usage by exploiting the symmetry of the Hessian
  - Developed the **first work** that solves the diverging or cycling behavior of gradient descent-ascend on minimax optimization problems by tapping into Anderson Acceleration
  - Proved accelerated convergence rates and demonstrated improved convergence compared to SOTA methods on classical and deep learning problems
- **Generative Models with Enhanced Convergence**
  - Focus on training generative models with enhanced convergence
  - Developed the **first work** that solves the diverging or cycling behavior of GDA on minimax optimization problems by tapping into Anderson Acceleration
  - Developed a normalizing flow that is able to learn both latent feature and tractable marginal likelihood estimation
  - Developed diffusion models to generate high-fidelity synthetic electronic health records
- **Parallel Computing:**
  - Designed smart parallel computing algorithms to enhance the performance of machine learning models training
  - Achieved up to  $5\times$  speedup compared with the state-of-the-art GAN models
  - Achieved at least a  $4\times$  speed-up compared to a state-of-the-art distributed tensor factorization method
- **Transfer Learning:**
  - Focused on robust machine learning deployment for data-scarcity scenarios
  - Developed high-accuracy out-of-distribution data models using domain adaptation
  - Showed significant improvement on immunotherapy response prediction and time series classification

2018-2023 **Machine Learning for Healthcare.**

- **Longitudinal Electronic Health Records Phenotyping :**
  - Developed a semi-supervised tensor decomposition method that innovatively combines labeled and unlabeled data, enhancing the accuracy and relevance of patient phenotypes extracted from electronic health records
  - Developed a distributed tensor decomposition framework that not only scales efficiently with large datasets but also significantly improves the interpretability and robustness of health data analysis, setting a new standard in healthcare analytics.
- **Precision Medicine via ML-based Causal Inference**
  - Spearheaded the development of cutting-edge machine learning methodologies aimed at revolutionizing precision medicine. This involved the intricate use of advanced analytical techniques to tailor patient care based on longitudinal observational data.
  - Innovated in the realm of causal sparse tensor estimation, deriving novel theoretical results that provide deeper insights into the complex relationships within high-dimensional and sparse healthcare data.
  - Developed a novel generative model specifically for Electronic Health Record (EHR) digital twins simulation, paving the way for more accurate and personalized patient care models and simulations in medical research.

---

## Teaching Experience and Service

2024 Fall **Teacher**, *Math 2667: Linear Algebra (Honor)*, Auburn University.

2019-Present **Program Committee/Reviewer of NeurIPS, ICML, ICLR, AISTATS, AAAI, UAI, AMIA, KDD, TKDE, CHIL, ECCB.**

2023 **Organizer of ICLR Workshop**, *Topic: Generative AI and Foundation Models in Health*, Proposal Submitted.

- 2023 **Co-supervisor to two undergraduate students**, *Topic: Generative Models for EHRs*, Emory University, Accepted by ICLR 2024.
- 2023 Summer **Organizer and Instructor of NSF REU/RET Computational Mathematics for Data Science**, *Topic: Fair Generative Modeling in Healthcare*, Emory University. The work is accepted by 2023 AI In Health Conference
- 2022 **Organizer of SIAM Conference in Optimization**, *Topic: Optimization for Healthcare*, Seattle.
- 2022 **Organizer of SIAM Conference on Computational Science and Engineering**, *Topic: Acceleration methods for scientific and machine learning applications*, Amsterdam.
- 2022 **Technology Chair of Conference on Health, Inference, and Learning (CHIL)**, Cambridge, Massachusetts.
- 2021 Summer **Instructor of NSF REU/RET Computational Mathematics for Data Science**, *Topic: Learning From Images via Convex Normalizing Flow*, Emory University, Atlanta.
- 2019 Spring **Teaching Assistant of Natural Language Processing**, *Emory University*, Atlanta.
- 2018, 2019 Fall **Teaching Assistant of Machine Learning**, *Emory University*, Atlanta.

---

## Invited Talks

- 2023 Scientific Computing Meets Machine Learning (Invited Talk). In: *Clemson University*.
- 2023 Multi-Modal EHR Generative model (Invited Talk). In: *INFORMS Annual Data Mining*.
- 2023 Nonlinear Acceleration for minimax and nonlinear optimization (Invited Poster). In: *Brown University*.
- 2023 Optimization for Healthcare. In: *SIAM Conference in Optimization*.
- 2023 Acceleration methods for machine learning applications. In: *SIAM Conference on Computational Science and Engineering*.
- 2022 Acceleration Algorithms for Machine Learning Models (Invited Talk). In: *Weill Cornell Medicine*.
- 2021 Nonlinear Acceleration for Tensor Decomposition. In: *SIAM Conference on Computational Science and Engineering*.
- 2019 Tensor Decomposition for healthcare analytics. In: *2019 Georgia Scientific Computing Symposium*.

---

## Work Experience

- 2021 Summer **Efficient Ensembling Uncertain Quantification**, *Machine Learning Internship*, Lawrence Livermore National Laboratory, Livermore, CA.
- Aimed to compress deep learning models for better uncertainty quantification.
  - Used iterative pruning methods to train model efficiently.
  - Adopted snapshot learning rate scheme to obtain diverse ensemble models.
  - Adopted focal loss and meta-learning to further enhance uncertainty quantification.
- 01/2021-04/2021 **Continuous Glucose Monitors (CGM) Forecasting**, *Machine Learning Researcher*, SiBionics MedTech, Shenzhen, China.
- Calibrated model parameters of Continuous Glucose Monitors (CGM) system.
  - Developed a meal detection algorithm using real CGM data.
  - Benchmarked and developed glucose prediction algorithms using real CGM data.