

# Tong Si (she/her)

tong.si@slu.edu | (314)-696-9062 | St. Louis, MO

Homepage: <https://mathstat.slu.edu/~gong/tong.html>

## Academic Experience

### Culver-Stockton College

Canton, MO

- Assistant Professor of CS & Mathematics (tenure track)

Aug. 2024 – present

### Saint Louis University

Saint. Louis, MO

- Ph.D. in Statistics (GPA 3.96/4.0)

Aug. 2020- Aug. 2024

Advisor: Dr. Haijun Gong

Department of Mathematics and Statistics

- M.S. Candidate in Artificial Intelligence (GPA 3.96/4.0)

Aug. 2022-Present

Department of Computer Science

- M. A. in Mathematics

Aug. 2020 - May 2022

Department of Mathematics and Statistics

### Jilin University

Changchun, China

- B. S. in Mathematics and Applied Mathematics.

Sept. 2016 - Jun. 2020

- BEc in Actuarial Science(minor)

Sept. 2017 - Jun. 2019

## Research Publications

### Peer-Reviewed Papers

- Tong Si**, Yunge Wang, Lingling Zhang, Evan Richmond, Tae-Hyuk Ahn, and Haijun Gong. "Multivariate Time Series Change-Point Detection with a Novel Pearson-like Scaled Bregman Divergence." *Stats* 7, no. 2 (2024): 462-480.
- Tong Si**, Zackary Hopkins, John Yanev, Jie Hou, and Haijun Gong. "A novel f-divergence based generative adversarial imputation method for scRNA-seq data analysis." *Plos ONE* 18, no. 11 (2023): e0292792.
- Helen Richards, Yunge Wang, **Tong Si**, Hao Zhang, and Haijun Gong. "Intelligent Learning and Verification of Biological Networks." *Advances in Artificial Intelligence, Computation, and Data Science: For Medicine and Life Science* (2021): 3-28.

### Peer-Reviewed Abstracts

- Tong Si**, Yunge Wang, Lingling Zhang, Kate Cannell, and Haijun Gong. "Change-Point Detection Using Scaled Bregman Divergence." *F1000Research, 22<sup>nd</sup> International Conference on Bioinformatics (2023)*
- Tong Si**, Zackary Hopkins, John Yanev, Jie Hou, and Haijun Gong. "sc-fGAIN: An f-divergence-based Generative Adversarial Imputation Method for scRNA-sq Data Analysis." *F1000Research, 22<sup>nd</sup> International Conference on Bioinformatics (2023)*

## Research Projects

### Imputation of Time Series Data via Generative Models and GRU

Oct. 2023-Present

Team leader, Dr. Gong's group

- Conduct a thorough literature survey on time series data imputation to identify prevalent limitations and gaps in current methodologies.
- Develop a GRU-based time-series generative adversarial imputation network algorithm and investigate the mathematical theory underlying the algorithm.
- Implement the time series imputation algorithm based on different divergence functions using Python.
- Prepare a manuscript for submission to a peer-reviewed journal.

### Change-Point Detection for Time Series Data Using Scaled Bregman Divergence June 2023 - March 2024

Team leader, Dr. Gong's group

- Developed a Pearson-like Scaled Bregman Divergence Method [1] for Change-point Detection (CPD) of multivariate time series data; investigated the mathematical foundation of the algorithm and reinforce the algorithm's generality and reliability across a broader range of applications.
- Reproduced comparative methods in R and Python to benchmark our model against existing techniques. Compare the accuracy in identifying change-points, and performance across diverse datasets and conditions.

- A paper [1] was **published** in Stats in 2024.

### **Innovative Web-Based Library Management System**

*Sept. 2023 - Dec. 2023*

*Team leader of Course Project*

- Utilized SQL for robust database design and management, ensuring efficient data storage, retrieval, and manipulation; Implement the user interface using HTML, creating an intuitive and responsive web application.
- Built the core functionality of the system using Python, ensuring seamless integration with the database and frontend components; Apply GitHub for source code management and team collaboration, maintaining an organized and efficient development workflow.
- Used CircleCI for continuous integration, automating code testing and deployment processes, to enhance code quality and deployment efficiency
- Employed Docker Hub for containerizing the application, ensuring consistent deployment across different environments.

### **Imputation of sc-RNA Sequencing Data via Generative Adversarial Networks**

*Oct. 2022 - May. 2023*

*Team leader, Dr. Gong's group*

- Led the team to develop a novel single cell f-divergence based generative adversarial imputation network (sc-fGAIN) algorithm to impute the missing values in the single cell RNA sequencing data.
- Implemented the sc-fGAIN algorithm using Python and provide mathematical proofs to confirm its effectiveness and general applicability in imputation tasks.
- Managed a massive dataset with dimensions 10,164 by 3,918, ensuring efficient data preprocessing and algorithm application.
- Implemented and compared different state-of-the-art imputation methods as benchmarks using R, Python, and MATLAB to validate the superiority of our approach.
- Our paper [2] has been **published** in PLOS ONE in 2023, and receive a **Best Oral Presentation Award** [5] at 2023 International Conference on Bioinformatics, held in Brisbane, Australia.

### **Analytical Text Processing Using Machine Learning**

*Sept. 2022 - Nov. 2022*

*Course Project*

- Applied Python libraries Pandas for data manipulation and Scikit-Learn for machine learning model implementation, including using feature `sklearn.feature_extraction.text.CountVectorizer` for text preprocessing and feature extraction
- Processed raw text data using tokenization and lemmatization techniques.
- Implemented a variety of classification algorithms, including Naive Bayes, SVM, and Random Forest, to compare performance. Optimize models using cross-validation and grid search techniques.

### **Statistical Inference and verification of Regulatory Networks**

*Sept. 2020 - May 2021*

*Collaborative Research Project, Dr. Gong's group*

- Applied a weighted dynamic Bayesian network method to reconstruct gene regulatory network from time series microarray data with other team members.
- Implemented different model checking technique, including SMV and PRISM for the network verification.
- **A paper [3] was published in 2021.**

### **Teaching at Culver-Stockton College**

Computer Programming (Python); Elementary Statistics; Beginning Math; College Algebra;

### **Teaching at Saint Louis University**

**Instructor of College Algebra**

*Jan. 2022-Dec. 2022*

**Teaching Assistant: Regression Analysis; Bayesian Statistics & Statistical Computing**

*2023-2024*

**Teaching Assistant: Calculus I**

*Aug. 2021 - Dec. 2021*

### Professional Service

<b>Member of International Society for Computational Biology</b>	<i>Mar. 2024-present</i>
<b>Research Assistant, Dr. Gong's Group, Saint Louis University</b>	<i>Jan. 2023-Present</i>
<b>Reviewer of the following Journals:</b> BMC Bioinformatics; Heliyon; PLOS ONE ; Journal of Bioinformatics & Computational Biology; Journal of Theoretical Biology; Genomics ...	
<b>Treasurer of Association for Women in Mathematics (AWM) in St. Louis University</b>	<i>Aug. 2022- Jan. 2023</i>

### Conference Presentation

Poster Presentation at <b>16<sup>th</sup> Great Lakes Bioinformatics (GLBIO) conference, Pittsburgh, PA</b>	<i>May.2024</i>
Oral Presentation at <b>Annual Graduate Research Symposium, Saint Louis University</b>	<i>Apr. 2024</i>
Oral Presentation at <b>the Mathematical Association of America Missouri Section, Liberty, MO</b>	<i>Apr. 2024</i>
Oral Presentation at <b>the Danforth Plant Sciences Center, St. Louis, MO</b>	<i>Jan. 2024</i>
Oral Presentation, <b>22<sup>nd</sup> International Conference on Bioinformatics, Brisbane, Australia</b>	<i>Nov. 2023</i>
Poster Presentation, <b>International Conference on Intelligent Biology &amp; Medicine, Tampa, FL</b>	<i>Jul. 2023</i>

### Awards and Certificate

<b>Full financial support for Mathematical Problems in Industry (MPI) Workshop</b>	<i>Jun. 2024</i>
<b>Full financial support for Graduate Student Mathematical Modeling Camp (GSMMC)</b>	<i>Jun. 2024</i>
<b>2nd Place for Oral Presentation Award, Annual Graduate Research Symposium, SLU</b>	<i>Apr. 2024</i>
<b>GLBIO 2024 travel fellowship</b>	<i>Mar.2024</i>
<b>Dean's Travel Award, Saint Louis University</b>	<i>Nov. 2023, Mar.2024</i>
<b>Best Oral Presentation Award, 22<sup>nd</sup> International Conference on Bioinformatics, Australia</b>	<i>Nov. 2023</i>
<b>Travel Award, Forty Third Midwest Probability Colloquium</b>	<i>Oct. 2022</i>
<b>Teaching Certificate, Saint Louis University</b>	<i>Sept. 2024</i>

### TECHNICAL SKILLS AND CERTIFICATIONS

- Computer Languages: Python, R, MATLAB, SQL, HTML
- Skills: Data analysis for big data, Software development, Database skills, Website building skills

### References Lists

1. **Dr. Haijun Gong, (Ph.D. Advisor)**  
Associate Professor of Statistics  
Department of Math and Statistics, Saint Louis University  
<https://mathstat.slu.edu/~gong/>  
[haijun.gong@slu.edu](mailto:haijun.gong@slu.edu)
2. **Dr. Jie Hou, (Ph.D. Dissertation Committee, AI M.S. Committee Chair, collaborator)**  
Assistant Professor of CS  
Department of Computer Science, Saint Louis University  
<https://cs.slu.edu/~hou/>  
[jie.hou@slu.edu](mailto:jie.hou@slu.edu)
3. **Dr. Russell Blyth, (Teaching Observer)**  
Professor of Mathematics  
Department of Math and Statistics, Saint Louis University  
<https://blythrd.me/slu/>  
[russell.blyth@slu.edu](mailto:russell.blyth@slu.edu)
4. **Dr. Tae-Hyuk (Ted) Ahn, (Ph.D. dissertation committee, collaborator)**  
Associate Professor of CS  
Department of Computer Science, Saint Louis University  
<https://cs.slu.edu/~ahn/>  
[ted.ahn@slu.edu](mailto:ted.ahn@slu.edu)
5. **Dr. Bryan Clair, (Ph.D. dissertation committee, former department Chair)**  
Associate Professor of Mathematics  
Department of Math and Statistics, Saint Louis University  
<https://mathstat.slu.edu/~clair/>  
[bryan@slu.edu](mailto:bryan@slu.edu)