Yinan Bu

University of Science and Technology of China | ynbu1002@gmail.com | yinanbu.github.io
+86 151-8999-5102 | +1-734-548-2115 | github.com/byn1002

EDUCATION

School of the Gifted Young, Sep 2022 – Jul 2026
University of Science and Technology of China (USTC), Hefei, China
B.Sc. in **Statistics**

- **GPA: 4.02/4.30 (92.33/100) Major GPA: 4.12/4.30 (93.63/100)**
- **Ranking: 3/116** across **School of Management** and **School of the Gifted Young**;
  1st among female students (across **School of Management**, **School of the Gifted Young** and **School of Mathematical Sciences**).
  1st in statistics majors in the **School of Management**.

Research Interests: Statistical Network Analysis, Statistical Machine Learning, Biostatistics, Optimization

Skills: C, Python (Pytorch), R (Rcpp), LaTeX, Mathematica, MATLAB

AWARDS & HONORS

Guo Moruo Scholarship Nomination (most prestigious scholarship at USTC)	2025
China National Scholarship (highest scholarship from Ministry of Education of China)	2025
Yang Ya Alumni Fund Scholarship (top 5 female students in School of the Gifted Young)	2024
Excellent Student Scholarship – Silver (top 10%)	2023

PUBLICATION

- F. Jiang, **Y. Bu**, S. Wu, G. Xu, J. Zhu. (2025). *Efficient Synthetic Network Generation via Latent Embedding Reconstruction*. Under review.
- **Y. Bu***, D. Wu*, X. Wang. *Feature-Subspace Hyperdimensional Computing: Scalable and Robust Learning with Theoretical Guarantees.* In preparation.

RESEARCH EXPERIENCES

Efficient Synthetic Network Generation via Latent Embedding Reconstruction Jul 2025-Present
Advisor:
Prof. **Gongjun Xu** (Professor, Department of Statistics, University of Michigan, Ann Arbor)
Prof. **Ji Zhu** (Susan A. Murphy Collegiate Professor, Department of Statistics, University of Michigan, Ann Arbor);

- Developed **SyNGLER**, a general, efficient framework for generating synthetic networks by leveraging latent space network models with a distribution-free generator over learned latent embeddings. Extended this framework to attributed networks and proposed **SyNGLER-Attr**.
- Built scalable pipelines for a diffusion-based latent embedding generator and a bootstrap-based latent embedding resampler, preserving key network characteristics while enabling efficient training and sampling with lower computational cost than many existing deep architectures (GitHub).
- Conducted empirical studies on both simulated datasets and real-world datasets, showing that the proposed method efficiently generates networks that more faithfully preserve key characteristics than existing approaches.

Machine Learning and Hyperdimensional Computing Apr 2024-Present
Advisor: Prof. **Xueqin Wang** (Chair Professor, Department of Statistics and Finance , USTC)

- Derived asymptotic information loss in vanilla hyperdimensional computing (HDC) operations and developed Hoeffding-type bounds for both hypervector similarity and predictive accuracy.
- Designed **F**eature-**S**ubspace based **H**yper**d**imensional **C**omputing (**FSHDC**), a robust and highly scalable model for fast classification and interpretation. Applied to fMRI/MRI data from UK Biobank and achieved a **+0.20** AUROC improvement over vanilla HDC with strong robustness under class imbalance.
- Integrated **attention** mechanism into the HDC training pipeline, yielding a **30%** accuracy improvement on the HAR dataset over vanilla HDC and a **15%** improvement over an attention-only baseline.

Large Scale Optimization and GPU Acceleration Nov 2024-Feb 2025

Advisor: Prof. **Xueqin Wang** (Chair Professor, Department of Statistics and Finance , USTC)

- Worked on graph trend filtering (minimizing ℓ_1 norm of discrete graph differences) to recover piecewise-smooth signals; examined the ADMM trade-off between convergence speed and subproblem solvability.
- Proposed **D**ifferential **O**perator **G**rouping–bas**e**d ADMM (**Doge-ADMM**), grouping differential operators to get closed-form subproblems and parallel updates.
- Built a parallel implementation for first- and second-order cases and achieved up to **30**$\times$ speedup compared with standard ADMM solvers (GitHub).

ACADEMIC PROJECTS

Analysis of the Government Pension Fund of Norway Jun 2025-Jul 2025

Supervisor: Prof. Canhong Wen (Department of Statistics and Finance, USTC)

- Independently designed, implemented, and deployed an RShiny website for the Norwegian Government Pension Fund Global (NBIM) with interactive Plotly charts and a Leaflet world map (RShiny).
- Conducted a comprehensive analysis that integrated statistical summaries with mapped visualizations and annotated trends, including embedded figures and map snapshots (GitHub).

Uncertainty-Aware Time-Series Forecasting via Conformal Prediction Dec 2024-Jan 2025

Supervisor: Prof. Yu Chen (Department of Statistics and Finance, USTC)

- Reproduced the conformal prediction framework in Stankeviciute et al. (2021) for probabilistic time-series forecasting, implementing model-agnostic, distribution-free prediction intervals.
- Conducted experiments on simulated and real-world datasets (AR/ARIMA, sales, air quality, COVID-19), showing robust uncertainty quantification and competitive interval performance versus standard baselines.

CORE COURSES

Probability and Statistics:

Probability	91	Mathematical Statistics	91	Applied Stochastic Processes	94
Regression Analysis	98	Multivariate Analysis A	96	Time Series Analysis A	96
Non-parametric Statistics	95				

Mathematics:

Mathematical Analysis I	95	Mathematical Analysis II	92	Mathematical Analysis III	93
Linear Algebra I	93	Linear Algebra II	91	Differential Equations	93
Real Analysis	86	Complex Analysis	95	Functional Analysis	99

Machine Learning, Computing & Other Courses:

Machine Learning	92	Statistical Algorithms	94	A Primer in Game Theory	93
C Programming Language	95	Applied Statistical Software	96		

ADDITIONAL INFORMATION

Teaching Assistant:

- C Programming Language (Undergraduate), *Instructor: Prof. Lixiang Tan* Sep 2024-Feb 2025
- Linear Algebra I (Undergraduate), *Instructor: Prof. Junchao Shentu* Mar 2025-Jun 2025

Code & Software:

- **SyNGLER**: A synthetic network generation framework based on latent embedding reconstruction (GitHub).
- **Doge-ADMM**: Differential Operator Grouping–based ADMM (GitHub).
- **NBIM**: Interactive visualization of the Norwegian Government Pension Fund global holdings (GitHub, RShiny).

Standardized Tests:

- TOEFL: 108 (R: 28, L: 30, S: 23, W: 27)
- GRE: 322 (V: 152, Q: 170, AW: 4.0)

Leadership & Activities:

- Excellent member of the football team of School of the Gifted Young 2022-2025
- Winner of 3 gold medals in track and field University-wide Annual Sports Games 2023-2025
- Flute player at the school Chinese orchestra 2022-2025